CHAPTER ONE

1.0 INTRODUCTION

1.1 BACKGROUND TO THE STUDY

Viral hepatitis has been the main cause of hepatic pathology and its associated problems, which have evidently become a global burden (WHO, 2014). Hepatitis can be caused by other factors different from viral agents. These include autoimmune disorder, alcohol, fatty liver diseases and toxins. Viral hepatitis can be caused by one or combination of viral agents (A, B, C, D and E). Hepatitis is a human infection that spreads mainly through blood causing inflammation of the liver (WHO, 2014).

Hepatitis poses a serious global health challenge and it is a disease of high morbidity and mortality worldwide; and infection can range from a mild illness that lasts a few weeks to a serious lifelong illness (WHO, 2018). However, the level of awareness about HCV infection in the developing nations is poor; and majority of infected persons are ignorant of their infectious status. The disease is accompanied by late presentation of symptoms (Seeff, 2000). This may result in chronic infection that can result in permanent liver damage, liver cirrhosis and hepatocellular carcinoma, in chronically -infected patients (Acheson, 2007). Studies have shown an overlap of disease risk factors for HIV, HBV and HCV, due to transmission through same means, resulting in infection with different viruses in the same individual (Otegbayo *et al*., 2008).

Globally, there are about 180-185 million infected individuals, with estimated 3-4 million becoming newly infected every year (Ashfaq *et al*., 2011; Mohd*et al*., 2013), out of which about 350, 000 deaths occur annually (WHO, 2017).According to Kapoor *et al.,* (2011), more than 3% of the entire populace is chronically infected with HCV, with a large number of people dying each year as a result of complications emanating from damage of the liver by the virus(CDC, 2012). In all, about 75% of HCV- related deaths occurs among adults aged 45 to 64 (Mukherjee, 2012; Dhawan, 2016).

Hepatitis C virus is a Flavivirus as it belongs to *Flaviviridae* family. It is the only member of the genus Hepacivirus. It is a small enveloped, spherical virus with a positive sense, single-stranded RNA genome. It was first identified in the year 1989 and designated as a non-A,

non-B hepatitis (Choo *et al.,*1989), while the first clone of the virus is H77 prototype strain (Arai *et al.,* 2013).As a result of frequent changes in the genome of HCV in form of mutation, different forms of the virus known as quasispecies are produced (Batenschlager and Lohmann, 2000).

Hepatitis CVirus has marked genetic heterogeneity or variation with an estimated nucleotide substitution rate of between 1.44×10^{-3} and 1.92×10^{-3} substitutions per site per year, which has been observed in its life cycle (Penin *et al.*, 2004). Consequently, accumulation of these nucleotide substitutions during HCV replication has resulted in the emergence of seven major HCV genotypes (1–7), each divided further into closely related subtypes classified as 1a, 1b, 1c, 2a, 2b etc., (Simmonds *et al*. 2005; Smith *et al.* 2014); with more than 100 subtypes in existence. A difference of about 30 to 35% exists in the nucleotide sequences over a whole genome (Iles *et al.* 2014).

Hepatitis CVirus subtypes are further grouped into quasispecies of about 10% variation depending on the difference in their gene composition (Ohno *et al*., 2007).Typically, characterizing the 5'UTR region results in genotypes with high accuracy, while analysis of the core/envelope protein1(C/E1) and non-structural protein (NS5B) coding region which is relatively conserved is able to distinguish the genotypes and subtypes of the virus (Gededzha *et al*., 2012).

Hepatitis CVirus nonstructural protein 5B (NS5B) is an RNA-dependent RNA polymerase (RdRp) and the key enzyme for replication of HCV RNA (Han *et al.,* 2014). The NS5B gene which codes for the RNA dependent RNA polymerase enzyme catalyzes replication of new (*de novo*)HCV RNA (Sir *et al*., 2012). HCV is known for its vast genetic variability. Theanalysis of this diversity in the viral genome reveals different levels of variation. For example, the 5'UTR and Core regions of the viral genome are highly conserved, while non-structural part which includes NS2, NS3, NS5B and 3'UTR are relatively variable, whereas envelope regions E1 and E2 and the NS4 and the NS5A genes show the maximum variations in the sequences(Le Guillou-Guillemette *et al.,* 2007). These two glycosylated envelope glycoproteins, named E1 and E2 embedded in the viral (lipid) envelope are important in cell entry during infection (Op De Beeck and Dubuisson, 2003).

Hepatitis C virus genotypesare distributed differentially throughout the world (Frank *et al*., 2000). Furthermore, the occurrence of the different HCV genotypes varies according to

geographic location (Ramia and Eid-Fares, 2006). For instance, Gt 1 and 2 are known to be prevalent in West Africa, while genotype 3 is found mostly in South Asia. Genotype 4 has been reported in central Africa and Middle East while genotype 6 is mainly prevalent in the South-East Asia. The subtype 4a has its origin in Egypt (Markov *et al*., 2009). Genotype 5 however is known to be common in northern part of South Africa and Belgium while genotype 7 was isolated from central African immigrants in Canada only (Messina et al., 2015; Murphy et al., 2015; Naamani *et al*., 2013).

The diversity exhibited by the virus has been associated with antiviral failure, inability to design a universal vaccine against HCV infection, as well as influence on viral persistence and disease progression in infected individuals. In Nigeria, studies are required to understand HCV diversity and its influence on HCV pathogenesis and epidemiology, which have serious implications for therapy, control and eradication of the virus.

1.2 STATEMENT OF PROBLEM

Studies have indicated the endemicity of hepatitis C virus in Africa and dearth of information on the sequence data from those regions including Nigeria, where enormous viral diversity yet to be discovered has been noted (Markov *et al.,*2009; Agwale *et al*., 2004). Owing to scarce data from Nigeria, policies regarding therapy and vaccine design may be delayed. HCV forms vary in their distribution and susceptibility to treatment, as well as in the efficiency of any immune response that they produce in the course of infection and treatment in infected individuals (Kimura *et al.*, 2000). HCV infection remains very expensive to diagnose and treat, due to the fact that its treatment is dependent on the genotype of the virus. This study will provide useful information on HCV diversity and epidemiology in order to aid case management and clinical outcomes.

Asymptomatic HCV-infected persons constitute risk for HCV transmission. Blood donors who look apparently healthy fall into this category. For those already showing signs of clinical disease, an interplay between some viral and host factors is presumed to be responsible. This study also answers some pertinent questions about the reason(s) while some HCV-infected individuals progress to clinical diseases, while others remain asymptomatic for a longerperiod of time as in blood donors, by analysis of the viral proteins for markers of disease progression.It has been observed that some level of interaction exist between host-

cell factors and RNAand HCV protein, in addition to their functions in replication. This was detailed by the function of HCV proteins during translation and post-translational changes. Furthermore, viral proteins have been observed and alleged to be responsible for a good number of biological processes that result in liver disease and carcinogenesis. Such activities include elicited signals to cell components, modulation of mRNA transcription, cellular transformation, cell death, and reorganization of membranes, movement of vesicles, regulation and modification in translation of proteins (Levrero *et al.* 2006).

1.3 AIMOF STUDY

This present study was designed to ascertain HCV diversity and viral markers of disease progression among asymptomatic blood donors; individuals with HIV and patients with clinical hepatitis(symptomatic) infected with different HCV genotypes.

1.4 STUDY OBJECTIVES

The following are the specific objectives of the study:

Study 1: To determine the circulating HCV genotype among patients and asymptomatic blood donors based on the NS5B gene.

Study 2: To identify mutations (major and minor) at the NS5B region of the isolates in the study groups in relation to disease progression and antiviral resistance.

Study 3: To identify post-translational modifications in NS5B through protein phosphorylation analysis of these diverse HCV strains,that may affect HCV replication and disease progression in infected population.

1.5 RESEARCH QUESTIONS

This study seeks to provide answers to the following pertinent research questions.

1. Which genotypes/subtypes are responsible for infections among symptomatic and asymptomatic HCV-infected individuals?
2. Are there natural resistant mutations in NS5B gene region? Are the mutations different in the population in relation to disease chronicity or progression; and to antiviral drug resistance? What are the implications?

3. Are there post-translational modifications in form of conformational changes in NS5B proteins of these diverse HCV strains that can influence disease outcomes in blood donors, individuals with HIV and patients with clinical hepatitis? Do they differ or same in the study groups?

1.6 JUSTIFICATION/RATIONALE FOR THE STUDY

The purpose of this research is to provide information and advance knowledge on the diversity of circulating strains of the virus among patients and infected blood donors,as well as to identifyHCV genetic factors of disease progression to chronic symptomatic state in patients with clinical hepatitis or hepatic pathology, in comparison to infected blood donors and patients with HIV, who remain asymptomatic or non-progressive to clinical disease. The study seeks to answer pertinent research questions on why some HCV- infected persons remain infected for long periods without showing any symptoms or progressing to clinical disease while others infected with the same virus proceed even faster to liver diseases and complications.

Knowledge about HCV diversity and epidemiology in every infected population is important in making appropriate decision on HCV management and therapy. At the moment, no vaccine is available for prevention of HCV. HCV is known for its high degree of strain variation due to high rate of mutation, being an RNA virus; as such, this necessitates further research on circulating HCV variants and the extent of this variation which has a serious implication for development of effective vaccine and new antiviral molecules (Ashfaq *et al.,* 2011). Hence, information on HCV types in relation to their distribution and genetic features in infected donors and patients is the ultimate goal of this research, due to its implication in therapy and control of HCV infection.

Consequently, knowledge of HCV genetic diversityin a single or multiple infections is a major prognostic factor in treatment outcome, thus influences choice of treatment (Legrand-Abravanel*et al.*, 2004). Genotyping and sequence analysis are important tools required for the study and monitoring of patients who are chronically infected with HCV. Sequencing of the NS5B gene of HCV enables accurate determination of genotypes and subtypes circulating in a population. In resource limited settings, genotyping for every infected patient remains elusive due to the high cost of the technique. However, this technique is most effective (gold standard) for genotyping due to its high accuracy. In practical terms, genotyping is typically based on polymorphisms or substitutions in a gene region or section (5'NCor NS5B).

Hence, it is obviously impossible to approximate the frequency of recombining events in a single gene (Le Guillou-Guillemette et *al.,* 2007).Understanding HCV genetic factors that contribute to virulence or infectiousness of the variants becomes more important in the management of complicated cases, especially with reference to therapy and response to antiviral agents. Previous studies reported that the following populations are more susceptible to liver complications such as hepatocellular carcinoma. These include patients infected with genotype 1 or 4; those co-infected with HIV or HBV, with elevated HCV burden, or not responding to treatment with one antiviral agent; and individuals that have gone through organ transplant (Ayesh *et al*; 2009; Dhiman and Chawla, 2005).

In a situation of difficulty of treatment, a good knowledge of particular HCV sequence information in such patients would facilitate treatment response and good prognosis, as certain genotypes are associated with quicker disease progression and more severe hepatic pathology (Ramalho, 2003). The knowledge will enhance effective treatment to achieve sustained virological response (SVR) among infected patients undergoing therapy. SVR is achieved when HCV RNA is not detected with PCR in a patient's blood in aboutsix months or 24 weeks after treatment has stopped (Deutsch and Hadziyannis, 2008).

Differences exist in patients' response to antiviral therapy which may be due to several reasons ranging from some host to viral features. Ultimately, the genotypic properties of hepatitis C virus coupled with its replication capacity, determine to a large extent, the rate at which successful treatment can be achieved. Hence, overall, adequate knowledge of its gene constitution and epidemiology is vital when predicting disease course and progression, treatment options and response to antiviral therapy.

The study is needed to provide information necessary for care and management of infected individuals especially among chronically- infected patients. A good understanding of HCV epidemiologyand pathogenesis in any region remains important in informing useful decisions regarding the virus and its control. The usefulness also extends to patients who have developed resistance to monotherapy and current antiviral drugs. In addition, the result of this study will advance knowledge on the epidemiology of HCV which is required for recommendation of appropriate policies other than management of infection, towards prevention, control and eradication of the virus in the country.

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 HISTORICAL BACKGROUND

Hepatitis C Virus (HCV) affects humans but its exact origin lacks records to elucidate it (Mandal, 2015). Due to lack of historical records, the evolution of Hepatitis C Virus can only be understood throughphylogenetic analyses of the current circulating populations (Simmonds, 2001).

Blood tests to detect hepatitis A and hepatitis B infections were respectively developed between 1963 and 1973 (Choo *et al.*, 1989).However, during the tests, it was observedby Alter*et al*, that samples that actually infected those people were negative for types A and B. Based on this; the infectionwas classified as NANB hepatitis. HCVwas first discovered in 1989 by investigators from CDC (Choo *et al*., 1989), when the prototype strain HCV-1 genome was cloned.As a result of the findings above, it became obvious that more than 90% of cases that were previously classified as NANB hepatitis were actually due toinfection with HCV (Zinotto *et al*., 1996). Thefirst prototype of HCV is H77 clone (Arai *et al.,* 2013). Subsequently, in 1997, two molecularly-different clones of H77 were reported infectious in chimpanzee model, but could not be replicated *in vitro*. Before now, research involving HCV replication and pathogenesis was hindered by scarcity of a supportive cell culture system, which evidently became available and feasible only in the year 2005, and this led to the discovery of JFH1 (genotype 2a), a genome with whichHCV infection could be established in Huh7.5 (human hepatoma) cell lines.

2.2 THE HEPATITIS C VIRUS

Hepatitis C Virus is a small, enveloped, spherical virus with positive sense, single-stranded RNA genome of ~9.6 kb consisting of a single open reading frame (ORF) and 5′ and 3′ untranslated regions –UTR (Moradpour *et al*., 2007). The ORF encodes viral structural proteins (Core and envelope glycoproteins E1 and E2), a small membrane protein (p7), and six non-structural proteins (NS2, NS3, NS4A, NS4B, NS5A, and NS5B) (Moradpour*et al*., 2007;Smith *et al.,* 2014). HCV has been classified into 7 major genotypes differing in nucleotide and amino acid sequences by approximately 30% and numerous subtypes with sequence heterogeneity of 15 to 20% (Simmonds *et al.*, 2005; Li *et al.,* 2015).

Hepatitis C virus has chronically infected over 170 million people worldwide and is a leading cause of liver fibrosis, cirrhosis, and hepatocellular carcinoma (HCC). WHO reports that more than 350,000 deaths thatoccur annually are due to HCV-related liver diseases (WHO, 2014). It was estimated that about 75 to 85% of persons infected with HCV will progress to chronic infection and will be at risk for developing cirrhosis, with an increased risk of advancing to hepatocellular carcinoma (HCC), in the remaining 10 to 15% within 20 years of infection with HCV as shown in Fig 2.1 (Thomas *et al.,* 2000).

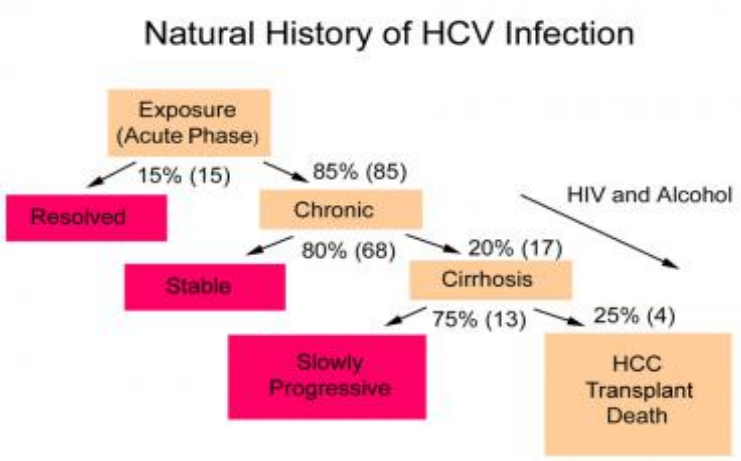

Fig. 2.1 Natural History of HCV Infection in Humans (Thomas *et al.*, 2000)

2.2.1 Classification

Hepatitis C virus belongs to the *Flaviviridae* familyand the only member of the Hepacivirus genus.The *Flaviviridae* family is divided into three genera: flavivirus, pestivirus, and hepacivirus. Flavivirus genus includes Yellow fever virus, dengue fever virus, Japanese encephalitis virus, and Tick-borne encephalitis virus. Pestiviruses include bovine viral diarrhea virus (BVDV), classical swine fever virus and Border disease virus. HCV, with seven genotypes and numerous subtypes and quasi-species is the only member of the hepacivirus genus. The virus shown in Fig 2.2 and 2.3,has close relationship with Tamarin virus and to human GB virus (GBV-C) which is less common (Lindenbach *et al,* 2001).

2.2.2 Physico-Chemical Properties

Hepatitis C virus is very unstable at room temperature and during repetitive freeze-thawing. Exposure of the virus to fat (phospholipids) as well as detergents, or even heating for about 10hrs at temperature of 60°C or 100°C for just 2 minutes in aqueous solution, exposure to concentrated formaldehyde (1:2000) at temperature of about 37°C for 72 hr, in ß-propriolactone and ultraviolet radiation(Purcell, 1994) can all inactivate it.

HCV particles possess a density of about 1.24 g/cm3 in some salts such as CsCl, and also a coefficient of sedimentation at 200S in sucrose slope which measures from 1.08 - 1.11 g/ml. Viral nucleocapsid in sucrose have a compactness (density) of 1.25 g/cm^3(Houghton, 1996; Purcell, 1994).

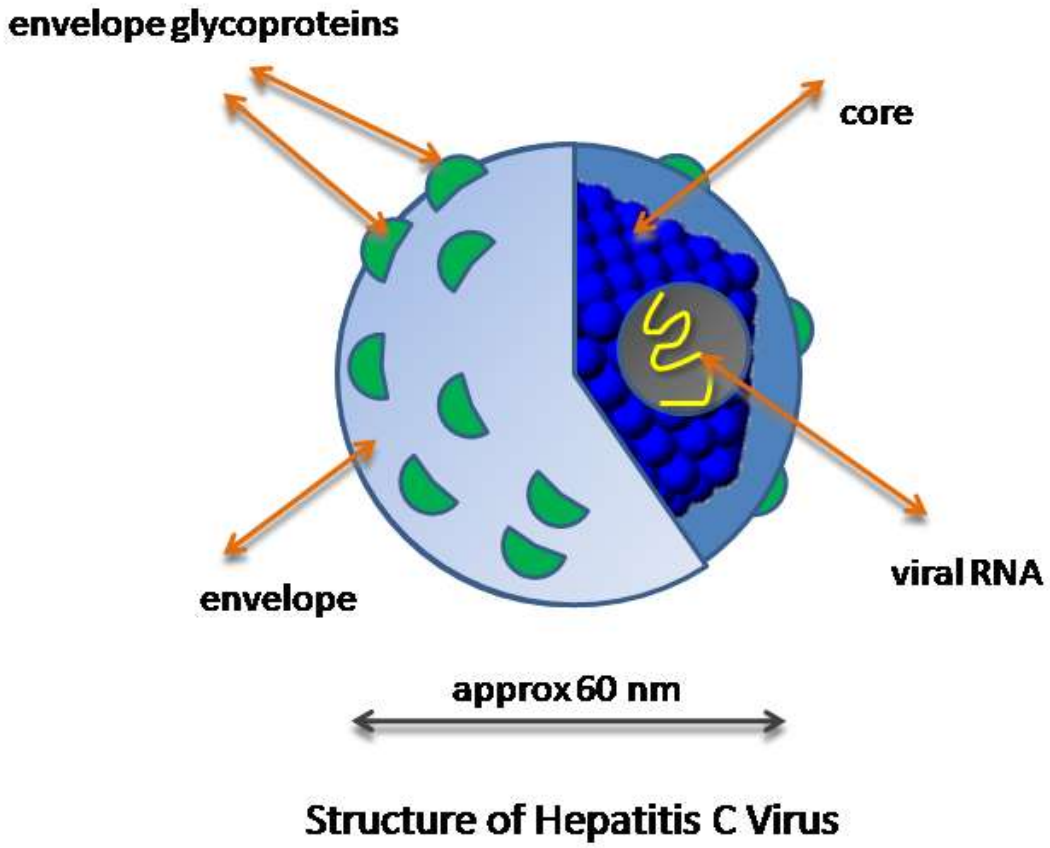

Fig. 2.2 Source: Graham *et al.,* (2007)

Electron Micrograph of Hepatitis C Virus

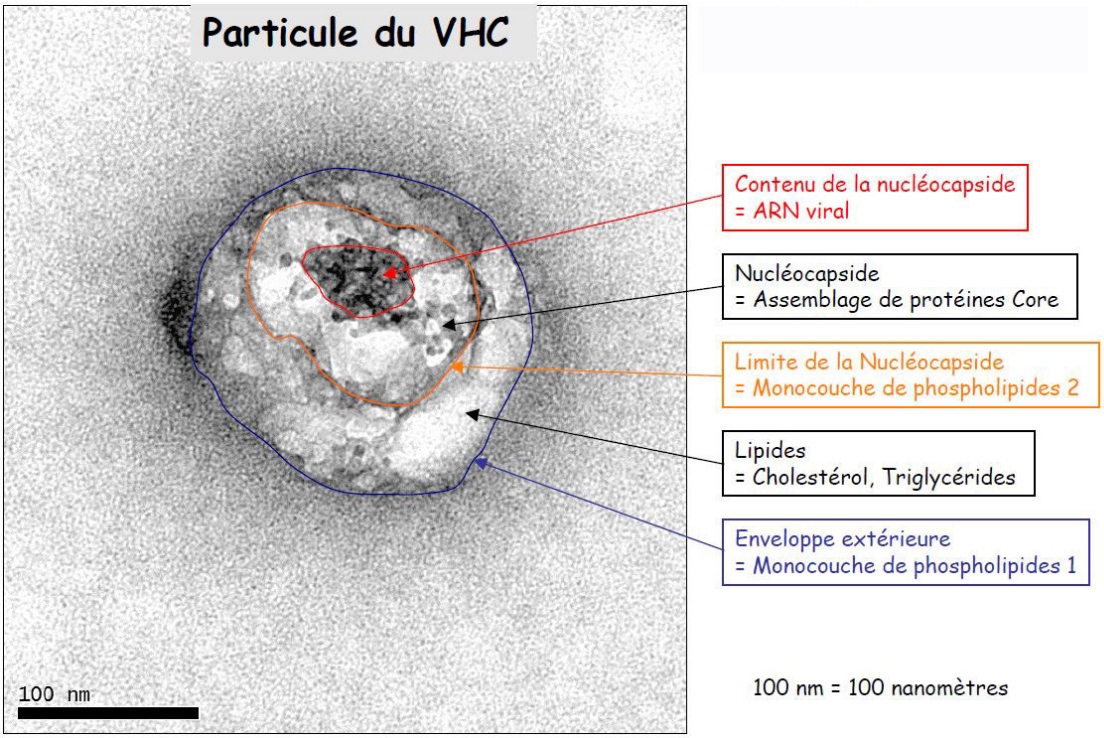

Fig. 2.3: Electron Micrograph of HCV

Source: British Society of Gastroenterology (2016)

2.2.3Hepatitis C Virus Genesand Proteins

The viral genome comprises small enveloped, spherical, single-stranded positive sense,RNA genome of approximately 9.6 kb (Moradpour *et al*., 2007; Smith *et al.,* 2014). HCV exhibits considerable sequence variation among individual isolates unlike HBV (Mori et al., 1994). Morphologically, HCV has a diameter of 55-65nm and consists of a core containing the viral RNA genome enclosed within an envelope, which has glycoproteins with short spikes. The genome consists of an open reading frame (ORF) which in turn encodes three structural proteins including core, E1, E2; and seven non-structural proteins including p7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins (Simmonds, 2004). Core, envelope proteins, E1 (position 192-383) and E2 (384-729/746) form the structural components. Structural proteins namely p22 are at the core or nucleocapsid, gp35 for E1, gp70 for E2, p7 for NS1, p23 for NS2 and non-structural proteins are p70 for NS3, p8 for NS4A, p27 for NS4B, p56/58 for NS5A and p68 for NS5B (Graham *et al.,*2007).

Notably, the E2 protein has two regions which possess high mutation rate. These are known as hypervariable zones or regions 1 and 2 (HVR 1 and 2). E2 contains binding sites (glycoproteins) that interact with CD-81 (cell surface molecules) receptors found on liver cells (hepatocytes) and B-lymphocytes (McLean *et al*., 1977). The envelope has sequence variation with amino acid similarity of 49-70% among different HCV isolates. The translation of these viral proteins (gene products) depends on the internal entry site of the ribosome in the 5'UTR. This contains a complex RNA which interrelates directly with the 40S ribosomal subunit initiation of translation in the cell (Simmonds, 2004).

Non-structural proteins include NS1, NS2 (codes for viral proteins and helicases), NS4A, NS4B (Interferon resistance proteins), and NS5B (codes for RNA dependent RNA polymerase) (Acheson, 2007). NS3 protein contains serine protease typically at the aminoterminal, while helicase is at the carboxyl end (Dubuisson *et al*. 1994).

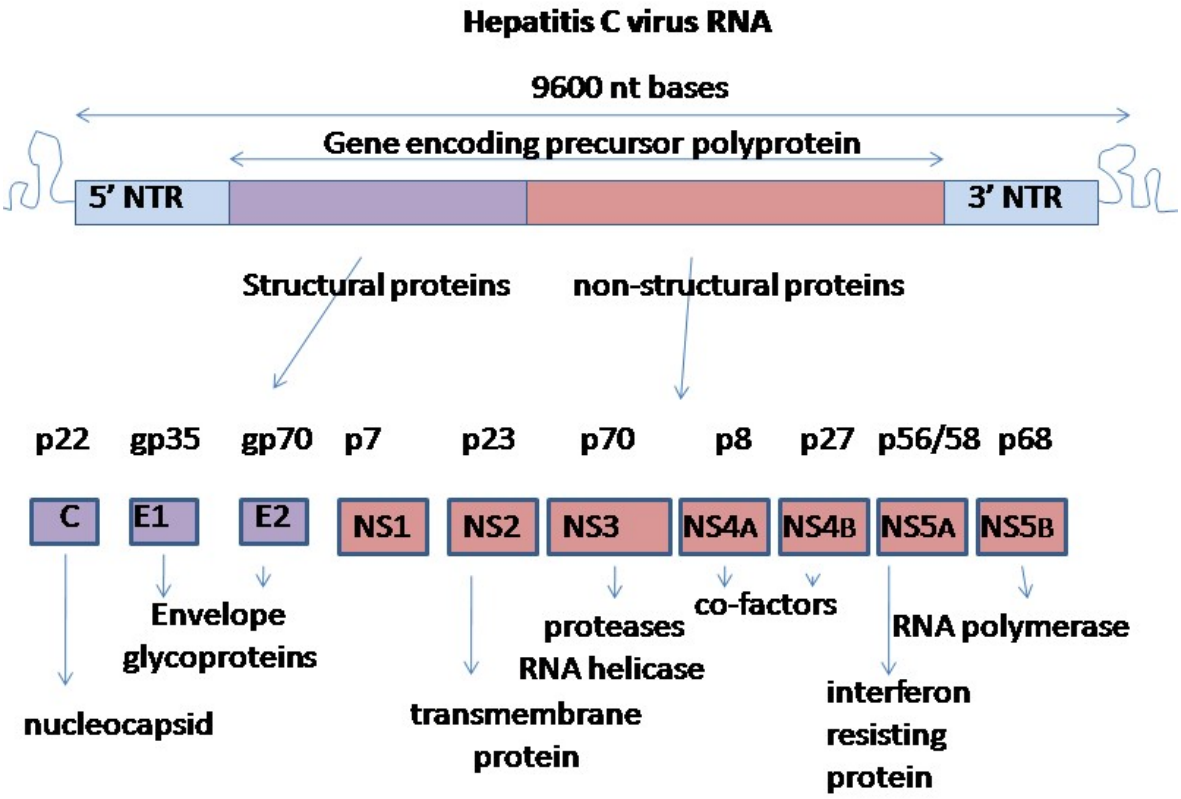

Fig 2.4 Source: Graham *et al*.,(2007)

THE NON -STRUCTURAL PROTEIN 5B (NS5B)

This protein codes for RNA-dependent RNA polymerase (RdRp), an enzyme, very crucial in HCV replication (Hans *et al.,* 2010).The enzyme ('life wire' of the virus) is the catalytic subunit of amplification (replicative machinery) responsible for synthesis of new or *de novo* HCV RNA and it is critically considered for designing new antiviral agents against HCV infection (Kolykhalov *et al.*, 2000). NS5B Polymerase lacks the proofreading ability to correct mismatches in copied sequences and so prone to error during replication;as a result,a large number of mutant strains (quasispecies) is generated. In most cases, these strains represent minor variations with about 1 or 2% nucleotide changes. The presence of these mutant strains has serious challenge in the control of HCV through immune mediation and this may explain the reason for varied infection with HCV, as well as the challenge in designing an efficacious broad vaccine (Kim *et al.,* 2004).

HCV NS5B (RdRp) has the typical closed right hand consisting of finger, palm, and thumb structure,which are common to all RdRps (Han et al., 2014). The palm domain contains the polymerase active site. An extensive finger-thumb interactions shield the palm domain to form a closed conformation, even in the absence of an RNA replication template (Chinnaswamy*et al*., 2008).

The NS5B protein forms an RNA replicase complex through interactions with some of the viral NS proteins. It is also able to form an oligomeric complex through its Δ1 loop and thumb subdomain which enables it to achieve the *de novo* RNA replication initiation-competent conformation (Chinnaswamy*et al.,* 2010).

Kang stated that activity of NS5B is regulated by its interaction with the viral capsid protein (Core proteins). Cellular proteins,in addition to viral proteins may also modulate the functions of NS5B (Kang *et al.,* 2009; Goh 2004).

In a study conducted by Han and colleagues,they demonstrated that protein kinase C-reactive protein 2(PRK2), which is a Serine and Threonine protein kinase, binds to and phosphorylates HCV NS5B polymerase at the N-terminal finger subdomain comprising amino acids 1 to 187,with regards to HCV RNA replication activated by PRK2. However, silencing or destabilizing of PRK2 expression by small interfering RNA (siRNA) molecules inhibiting PRK2 activity,reduced HCV replication and this demonstratesthe regulatory activity ofPRK2 in the entire existence of HCV (Kim *et al.,* 2009; Jakubiec *et al.,* 2006).

Structure of NS5B Gene

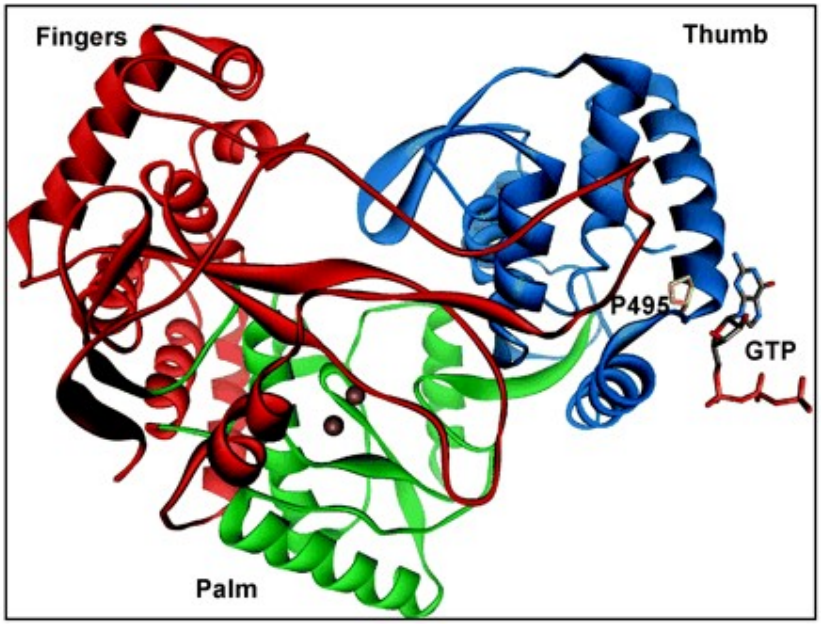

Fig. 2.5 Three Dimensional (3-D) Structure of NS5B
Source: Kolykhlov *et al.*, (2000)

2.2.4 STEPS INVOLVED IN HCV LIFE CYCLE

In general, replication of HCV takes place in the cytoplasm outside the nucleus typical of other RNA viruses. These include its penetration into the host cell, followed by uncoating of the viral genome, translation of viral proteins, processing of polyproteins, replication of viral genome, assembly of viral components and release of assembled virions (Bartenschlager *et al.,* 2013).

Typically, the virus enters a susceptible host, and this can be directly e.gthrough contaminated needle or transfused contaminated blood; or through a percutaneous means such as sexual or perinatal transmission (Alter, 1997). Next, the virus enters the hepatocytes or other susceptible cells such as the peripheral blood mononuclear cells (PBMC), utilizing the surface molecules (CD81 and scavenger ligands), as the viral receptor (Pileri *et al.,* 1998). The virus uncoats and releases the viral genome to initiate replication (Fig. 2.6). The genome (RNA as the mRNA being a positive sense), serves as the template for translation of the polyprotein and the processed non-structural proteins form a complex with the genome, forming a replication complex (Fig. 2.7), which resides in the cytoplasm derived from the endoplasmic reticulum, to initiate synthesis of negative strand RNA, which also serves as the template for positive RNA. The RNA replicative intermediate matures and begins to interact with the core and envelope proteins to assemble the virion. Non-structural proteins such as NS5B play a critical role in viral replication and productive infection by HCV (Liang *et al.*, 2000). The steps involved are outlined below, from attachment to release of newly-formed virion.

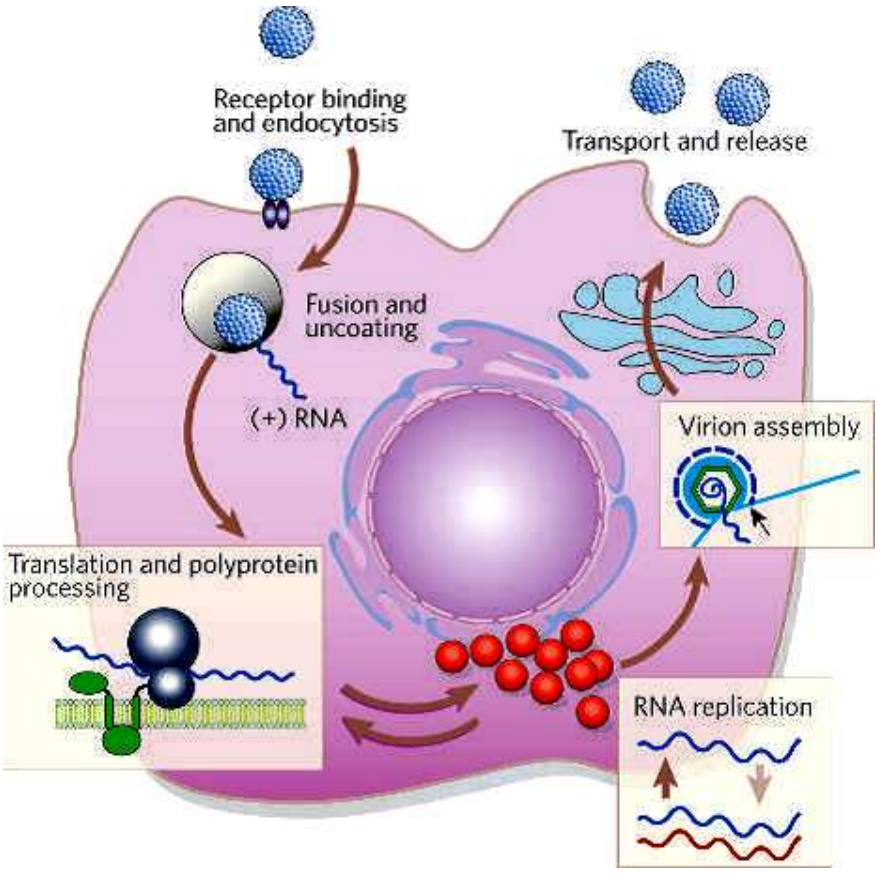

Fig. 2.6Life Cycle of HCV
Source: Lindenbach *et al.*, (2005)

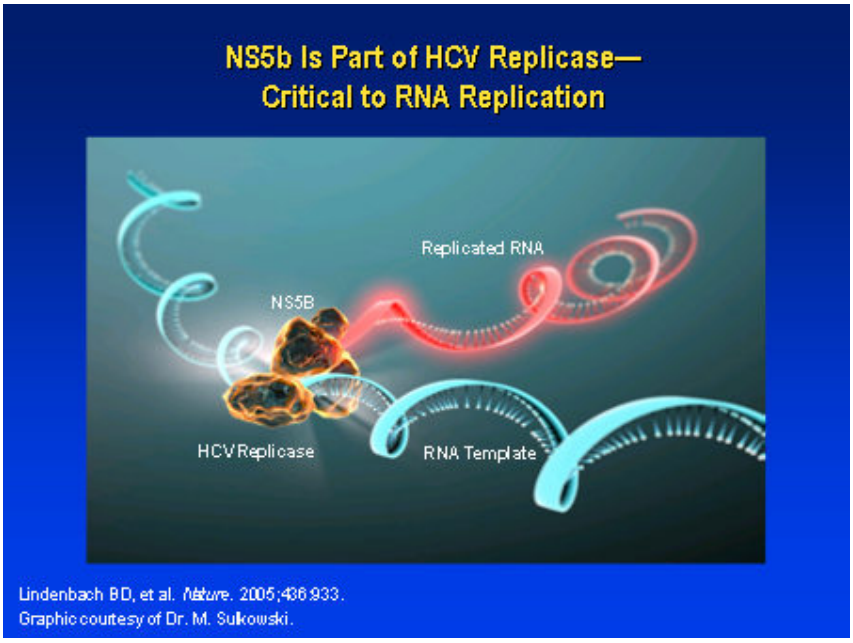

Fig. 2.7 Schematic diagram of steps involved in replication of HCV

Source: Lindenbach and Rice, (2005)

2.3. REPLICATION OF HEPATITIS C VIRUS

2.3.1 Attachment

Humans and experimental chimpanzee are the two models that mostly support the replication of HCV. Theenveloped virus particleattaches to specific surface receptors (CD81) on the host cells (hepatocytes) and becomes internalized in the endoplasmic reticulum (Lindenbach *et al.,* 2005)

2.3.2 Fusion and Uncoating

In the presence oflow pH of the endoplasmic reticulum, viral and cellular membranes of the virus tend to fuse together. This actually causesthe release of a single-strand (ss), positive-sense RNA genome into the cytoplasm of a newly infected cell (Lindenbach *et al.,* 2005).

2.3.3 Transcription

During the process of transcription, HCV RNA genome performsnumerous roles. In the first instance, it functions as mRNA used for translation of the HCV proteins. Secondly, it serves as a template for genome replication and finally it plays the role of a nascent genome which is wrapped within newly-generated viral particles (Moradpour *et al.,* 2002).

2.3.4 Translation

The single-stranded RNA serving as the mRNAscanbetranslated directly intoproteinsby activation of hostcell protein kinases (Moradpour *et al.,* 2007). Translated proteins are processed into polyproteins or polypeptide chains thereafter.

2.3.5 Assembly, Processing and Budding

The assembled virions (new viral particles) are released into the endoplasmic reticulum (ER) through the process of budding (Moradpour et al, 2002; Lindenbach et al, 2005).

2. 4 PATHOGENESIS OF HEPATITIS C VIRUS

Pathogenesis is the stepwise or sequence of biological mechanism(s) or events that lead to a diseased state. It involves step by step process beginning with exposure to a pathogen toestablishmentof disease in a host. There are two pertinent questions that are always raised when trying to look into the life cycle or natural history of HCV:

1) Why do some HCV-infected persons clear HCV within the first few months of infection–precisely during acute infection or have delayed manifestation of symptoms, while most people proceed even faster to chronic stage of infection?
2) What is the cause of different degrees of liver damage among chronically-infected individuals?

HCV pathogenesis is affected or controlled by the immune cells of the host, as well as some important metabolic modifications which can include hepatic steatosis(build-up of fats in the liver), insulin resistances andoxidative stress. In HCV infection, viral proteins may directly contribute to or aggravate some conditions, especially steatosis and hepatocellular carcinoma(Irshad, 2006).

It is of note that genetic variations that exist between HCV strains may produce changes in the actual proteins that they code for, while in some others, there may not be any resulting changes. According to Farci *et al.,*(2012), the disease process can extremely differ in individuals, and this has become a significant development in research involving the virus. As have been observed earlier, a few individuals remain asymptomatic for several years and may finally pass on due to other health conditions or reasons; however, a few other people may become chronically-infected, progressing into complications such as liver cirrhosis and cancer, which may eventually result into death within a short period. Over the years, this has remained a rather huge debate or puzzle that has stimulated great interest in HCV research. In their work, they categorized various observed changes noted in six HCV patients that were transfused with infected blood in the 70s, well before the virus was identified in 1989 (Farci *et al*., 2012).

Farci and his colleagues observed that individuals with rapid progression to disease possessed some specific detectable variations which date back to the early infection stage. The research team noted that those individuals had significant level of a particular protein called MCP-1. This important protein activates the pathway for liver fibrosis to set in, resulting in further complications. However, they reported that changes accumulated in the viral genome as it evolves were not responsible for observed changes in the proteins they encode. Hence, the team concluded that genetic and serum markers may in future enable medical practitioners to identify HCV- infected persons that are at risk for rapid progression to chronic stage and death; and that same information can be used for proper management of the disease in such individuals (Farci *et al*., 2012).

HCV, just like HIV is able to escape innate and adaptive immune responses in most infected persons, which enables it to establish chronic infection. In an attempt to eradicate the virus, cytotoxic -T-lymphocytes of the host cause liver injury (Irshad,2006).However, the presence of multispecific T-lymphocytes in blood of infected persons is associated with viral clearance.This simply means that immune status of the host is important during HCV pathogenesis. Particularly, cytotoxic-T- lymphocytes play important role during viral eradication from the host or its persistence. The processes involved are controlled by several factors including HCV quasispecies and extrahepatic regulatory pathway (Irshad *et al*., 2013).

2.4.1 Host Factors Associated with HCV Persistence and Progression

HCV persistence involves the way and manner HCV achieves chronicity in infected individuals. This has been a point of focus in most studies.Genetic, environmental as well as immune status of the host play significant role in viral persistence or progression. Such factors determine the progression of inflammation and fibrosis in chronic patients because of their impact on antigen recognition, processing and presentation (Gremion and Cerny, 2005).

These factors include: individual genetic differences, strength of immune response to HCV infection, age of host (HCV severity increases with age), gender (male are more prevalent than female gender in many studies), and race (people of African descent have higher chances of progressing to chronicity and resistance to therapy especially those infected with genotype 1). These can be as a result of genetic variations in the host (Viso, 2007) and alcohol abuse (Kohla and Bonacini, 2006).

Different findings have correlated increased alcohol consumption with HCV persistence in cells (Campollo, 2002). Again, a few other studies contradicted the report, stating that alcohol consumption influences host's immune responses, alters the inflammatory response of cytokines, which results in increased viraemia and can serve as essential co-factor in the development of hepatocellular carcinoma (Schiff, 1999). Notably, increased alcohol consumption results inincreased hepatic steatosis and this induces cell death in infected persons (Vento and Cainelli, 2002).

Numerous other host factors that can facilitate progression of chronic hepatitis C are smoking which can increase inflammation of the liver and fibrosis, promote direct liver injury or indirect liver damage by exerting toxic effect and also leads to production of some immunological components that can cause liver damage such as TNF-alpha , Interleukins-1

and 6 (El-Zayadi, 2006). Again, HCV when co-infected with HIV, HBV and human T-cell lymphotropic virus and intravenous HCV transmission are otherextrinsic factors involved in progression to complications like fibrosis, cirrhotic liver or hepatocellular carcinoma (HCC) in chronic patients (Viso, 2007; Kishihara, 2001).

Sugimoto and others observed that in a study involving several individuals who recovered from HCV infection,$CD4^+$ T-cell response was more in Caucasian-Americans with highest level of Th 2 responses(Sugimoto *et al.,* 2003).

According to Azocar *et al*(2003), hepatitis C evolution across various ethnic groups and race may be as a result ofsome hereditary causes such as presence of Human Leukocyte Antigen (HLA) class II alleles, in the hosts that can be associated with spontaneous elimination of the virus (Azocar *et al*., 2003). Patients with cirrhosis stand the chance of developing hepatocellular carcinoma as the disease progresses (Davis *et al.,* 1989; Kohla and Bonacini, 2006).

Studies have identified some host's genetic components such as MHC class II alleles e.g DR5 that are linked with reduced occurrence of liver cirrhosis (Thursz *et al.,* 1997) in chronic patients. There have been lots of assumptions, that MHC class II molecules are absent in cells infected by HCV, because viral proteins inhibit antigen presentation through induced interferon negative immune regulation (Taylor *et al.*, 2000). Interleukin-10 haplotypes when present are said to be predictors of HCV clearance (Viso, 2007).

2.4.2 Viral Factors associated with HCV Persistence

A few viral factors have been proposed and reported as the reasons for HCV persistence in a host. These may be increased rate of mutation in HCV, coupled with the effects of some of these viral proteins on the activities of infected cells, as well as, on the key arms of the immune structure responsible for host's response toHCV infection (Agnello, 2004).

HCV genotype distributed worldwide is HCV genotype 1. Some studies reported that this genotype has a link with serious hepatitis outcome, which may result in HCC. It has been reported also that these genotypes; 1b, 2 and 3 exhibit poor response to therapy (Levrero, 2006). Non-response to therapy could lead to most complications in infected individuals. Furthermore, in HIV/HCV co-infected patients with increased immune deficiency, it can also result in the acceleration of HCV-related liver damage and complications (Einav, 2002).

2.4.3 AcuteHCV Infection

This type of HCV infection poses a challenge due to the fact that few symptoms are presented at this stage, making it hard to diagnose infection in most infected individuals. However, the course of infection is divided into three phases which are: Preicteric or sub-clinical (asymptomatic or prodromal phase), Icteric (flu-like) and convalescent symptomatic phase. Fulminant HCV infection occurs in less than 1% of the population. The incubation period for HCV ranges from 15 to 150days (CDC, 2012).

It is very difficult to establish if an infection by HCV is acute or chronic. However, acute HCV infection can be defined as HCV infection less than six month following acquisition of Hepatitis C Virus (WHO, 2014), While recent infection is estimated duration of infection longer than six months but shorter than two years (Busch *et al*., 2003). In general, the gold standard for the laboratory diagnosis of acute HCV infection is negative HCV antibody result before suspected exposure (anti-HCV seroconversion) and positive anti-HCV following potential exposure, combined with a positive HCV RNA test and elevated alanine aminotransferase (ALT). Hence, to accurately diagnose an acute HCV infection, an individual or a population will have to be monitored from seronegative status to when exposure to HCV occurs (CDC, 2012).

According to the CDC, a level of ALT greater than 400IU/L is high. Use of anti-HCV testing therefore, to diagnose acute HCV infection is not reliable, since only about 50-70% of patients have detectable anti-HCV at the onset of symptoms (Gerlach *et al.*, 2003). Furthermore, when patients develop symptomatic acute HCV infection, they most often present with jaundice, influenza-like symptoms, dark urine, nausea, abdominal pain and malaise (Chung *et al.,*2005). According to Chung *et al*,(2005) still, polymorphisms in genes involved in innate immunity appear to influence the outcome of acute HCV infection. Positive antibody test does not differentiate acute from chronic HCV. Hence, with acute HCV, patients usually first have detectable HCV RNA, then elevated ALT and lastly positive antibody test when seroconversion has taken place (CDC, 2012).

2.4.4 HCV Clearance and Immune Functions

Despite HCV immune escape, the innate immune system strongly responds to HCV within the first days after infection (Heim, 2013). Typically, acute hepatitis develops 10-14weeks after infection and at this point, sharp increases in serum ALT levels and the appearance of HCV specific T-cells can be detected in the blood (Racanelli *et al*., 2003).

Upon virus entry through its attachment to receptor cells (CD 81) in the host, hepatocytes (liver cells) are stimulated to produce IFN- α and IFN-β, which in turn induce Kupffer cells to produce macrophage inflammatory proteins-1α and these recruit Natural Killer cells that in turn secrete IFN-γ. Interferon -gamma (IFN-γ) immediately up-regulates chemokines that direct liver-infiltrating lymphocytes into the hepatic parenchyma (Gerlach *et al.,* 1999). The effector cells identify the MHC (major histocompatibility complex) class 1-HCV peptide groups present on infected liver cells (hepatocytes). This recognition results in apoptosis, in the cells (cell death), which leads to clearance of infection. It has been observed that the presence of strong T cell ($CD4^+$and $CD8^+$) responses during acute HCV infection may be linked with recovery from HCV infection (Gerlach *et al.,* 1999). On the contrary, absence of T-cell response during this stage of infection may signify development of chronic stage of infection (Gerlach, 1999).

In a study involving chimpanzee model, the role of $CD4^+$ responses was demonstrated (Grakoui *et al*., 2003). In this study, chimpanzee animal models that recovered from HCV infection were depleted of $CD4^+$ cells and the animals re-challenged with HCV, which resulted in subsequent lingering infection. This confirmed the significance of $CD4^+$ cells in the resolution of an acute infection with the virus (Grakoui *et al.,* 2003; Busch *et al*., 2012).

Farci *et al*(2000),however, noted that reduction in number of HCV quasi-species population may be a way to regulate acute stage of HCV infection and increase immune response; whereas chronicity is associated with quasi-species increase (Farci *et al.*, 2000). In general, they noted that about 25% of individuals infected acutely experience natural recovery, mostly observed in those that possess IL28B genotypes, younger age less than 40 years, female sex and symptomatic illness (icteric) stage (Gerlach *et al*., 2003). Indeed, these host's factors that are associated with spontaneous clearance suggest that individuals who possess strong basal immune responsiveness and ability to produce jaundice and clinical illness (due to immune response mounted by host defense against infection), have a better likelihood of controlling infection with HCV. According to the study therefore, asymptomatic HCV -infected individuals have the higher chance of progressing to chronic infection than symptomatic persons (Hajarizadeh *et al*., 2013).

In one study, high rate of clearance was reported among young mothers who received tainted anti-D immune globin (Kenny-Walsh, 1999). Similarly, Ceci *et al*(2001),noted that infants

who contracted HCV infection had 75-100% rate of spontaneous clearance (Ceci *et al.*, 2001).

2.4.5 CHRONIC HCV INFECTION

The mechanism by which HCV establishes chronic infection remains elusive, but has stimulated numerous research activities to understand the process better. However, a combination of viral, genetic, and immunologic factors are involved in pathogenesis of chronic HCV infection (Busch, 2003).

Ordinarily, without considering gold standard for HCV testing, it can be very challenging to differentiate an acute HCV infection from chronic infection especially in individuals that have not previously undergone anti-HCV testing and are also asymptomatic (Busch, 2003).

Chronically-infectedHCV patients with liver disease can have non-specific symptoms or none at all, as this is one indication of liver disease in some individuals. Such people may progress with severe liver damage. It has be reported by Davis *et al,*(1989), that chronic liver disease (CLD) is the reason for orthotropic liver transplantation (OLT) in the United States (Davis *et al*., 1989).

Some immunologic factors are responsible for a range of extra hepatic manifestations observed in chronic HCV infection. These may include autoimmune and lymphoproliferative disorders including mixed cryoglobulinemia, Sjögren's syndrome, as well as B-cell non-Hodgkin's lymphoma (Agnello, 2004).

Prolonged infection with HCV can result in numerous hepatitis pathology which include fibrosis (scarring of the liver), cirrhosis (severe scarring that affects liver function and interferes with blood flow), steatosis (build-up of fats in the liver), and hepatocellular carcinoma (liver cancer). Infections with HAV, HBV, HEV or heavy alcohol consumption as well as hereditary disorders can give rise to these clinical manifestations as well (Agnelo, 2004; Viso, 2007).

2.4.3 HCV IMUNE RESPONSE AND PERSISTENCE

HCV establishes persistence and maintains chronicity in most infected persons. This distinct feature of the virus over the years is a phenomenon which aroused major concern about its pathogenesis.

According to the study done by Farci and colleagues (2000), human system reacts as soon as infection is established, however, in most patients the virus changes its surface proteins (mutation) once the patient's antibody start to respond to the infection, which makes it to slip into a disguised or masked state that enables it to continue to replicate without detection, thereby tricking the immune system (Farci *et al.*, 2000). This is a similar scenario observed in HIV and Influenza virus infections.

In the same study still, specific variations in the genes encoding special proteins that coat the surface (glycoproteins) explains why some HCV-infected persons fail to recover completely from infection with the hepatitis c virus. They found that among the patients studied, the virus in some patients did not change immediately after the initial immune response. In this category, HCV was completely eliminated in several weeks. However, in some, new genetic variants (quasispecies) began to appear in response to early immune response. Notably, in this group, the rapid evolution resulted in chronic infection (Farci *et al.*, 2000).

2.4.4 HCV EVOLUTION

HCV evolution is formed by distinct selection pressures associated with historical events responsible for HCV adaptation to its human host (Smith *et al.,* 2014). This has ensured sustenance of HCV transmission. The virus evolves rapidly, yielding adaptive changes that are linked to every new infection in individuals due to certain immune factors that relate to antiviral therapy. The observed variations in the sequences are accumulated by HCV over time. It is responsible for most observed differences between the viral genotypes (Simmonds, 2004). This significant genetic diversity is typical of HCV infection and a consequence of the lack of proofreading activity or error in RdRp (Le Guillou-Guillemette *et al.,* 2007; Asfaq*et al.,*2011) and increased viral replication during its life cycle (Sir *et al.,* 2012). Alterations in nucleotide frequency in HCV RNA are about 1400to 1900 switches per nucleotide annually (Nasu *et al.,* 2011). This eventually results in infection with HCV copies showing considerable divergent species, resulting in a collection of viruses in a population of closely related but non-identical genomes referred to as quasispecies (Pawlotsky, 2006; Martell *et al.,* 1992; Argentini *et al.,* 2009)), with the main viral population evolving as a result of its viral replicative fitness and immune selection pressure (Nasu *et al.,* 2011). Point insertion or deletion mutations in the genome within the seven genotypes and by recombination between HCV strains of different genotypes may have resulted in HCV evolution (Yun *et al.,* 1996).

2.5 GENETIC DIVERSITY OF HCV

2.5 .1 HCV Genotypes

Before 2017, there were seven known genotypes of HCV, with about 67 subtypes. The variations in genome resulted in the classification into seven distinct genotypes (Murphy *et al.,* 2015). A recent study by Borgia*et al* (2018), based on high degree of genetic diversity of HCV, classified it into eight genotypes (Smith *et al.,* 2017; Borgia *et al*, 2018) and 87 subtypes in alphabetical order 1a – 1n, 2a – 2u, 3a – 3k, 4a – 4w, 5a, 6a – 6xf, 7a – 7b, and 8a (Smith *et al.*, 2017; Borgia *et al*, 2018) .

Molecularly, an average of 35% nucleotide divergence has been observed between the genotypes, while the subtypes have nucleotide divergence of 25%. Only genotypes 5 and 7 are not yet subdivided into subtypes (Simmonds *et al.,* 1993).Genotyping is a molecular sequencing technique or tool consisting of amplification of partial or whole gene such as NS5B in the viral genome. It is used to analyze genetic variations in the genome. Genotype refers to the set of genetic variations that occur in HCV (Smith *et al.,* 2014).

HCV genotypes identification is based mainly on nucleotide sequence diversity (molecular differences) determined through phylogenetic analysis of some regions of HCV which can be the untranslated region (UTR) also known as non-coding region (NC) of the core, envelope and non-structural protein 5 -NS5A or NS5B (Murphy *et al*., 2007a; Van Doorm *et al.,* 1994). The distribution of each genotype has been well-noted over time. Genotypes 1, 2 and 3 have a worldwide distribution, while genotypes 4, 5 and 6 are restricted to some particular geographic regions. Genotype 7 has only been reported from Canada, in a Central African Republic immigrant (Murphy *et al.,* 2015).

2.6 EPIDEMIOLOGY OF HEPATITIS C VIRUS

2.6.1Pattern of HCV Circulation in the Globe

The virus has infected over 185million people globally (Ashfaq *et al*., 2011, Mohd *et al*., 2013), with an estimated 32 million in Africa (Mostafa *et al.*, 2010). In Egypt, the prevalence is highest in more than 10% of the general population while China has the highest number of people (29.8 million) with HCV (Hajarizadeh *et al.*, 2013).

Globally, HCV genotypes are distributed differentially but with restrictions to some regions. The first few genotypes are circulated all over the world but with little differences, while

some such as 4, 5 and 6 are restricted to specific geographical areas. The last genotype 7 was reported in Canada, only in a Central African Republic immigrant by Murphy *et al.,* (2015).

Specifically, genotype 1 is distributed globally while genotypes 2 are found mainly in Africa (Forbi *et al.*, 2012). Genotype 3 is found mostly in some Asian regions including Thailand, Singapore, and India etc. Subtypes of genotype 4 are mostly seen in Middle East particularly Egypt and Central Africa (Markov *et al.*, 2009; Njuoum *et al*., 2009). Genotype 5 has been found predominantly in South Africa whereas, genotype 6 has been found in some parts of South-east Asia. The last genotype (7) was reported in a Central African immigrant in Canada only (Murphy *et al.,* 2015).

It has been estimated that about 20% of infection around the United States is acute while about 40% referrals to major liver clinics are due to HCV infection, with an increase new infections and death in more than 30,000; and about 8000 to 10,000 respectively (Alter*,* 1988).

The overall incidence of HCV among various populations has been noted in various studies (Fig. 2.8). Sanaullah and colleagues (2011), observed that HCC resulting from HCV-related incidences was found to be on the increase in the United States, due to increasing number of chronically-infected patients (Sanaullah *et al.,* 2011).

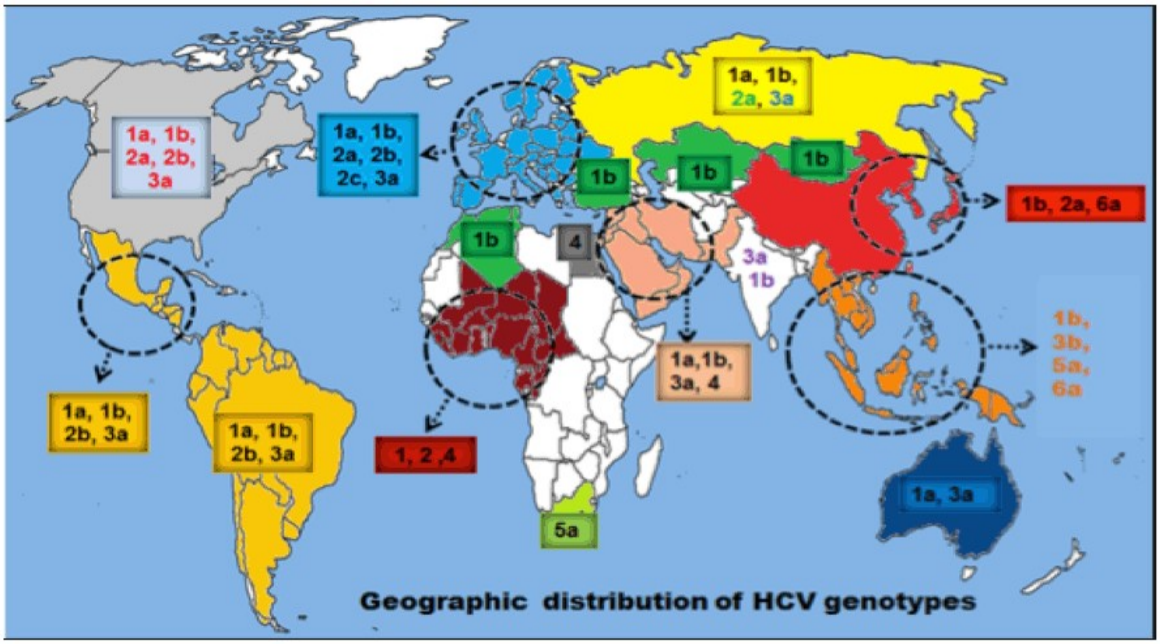

Fig. 2.8 Global Distribution of HCV Genotypes

Source: Lindenbach and Rice (2005)

2.6.2 Epidemiology of HCV in Nigeria

HCV infections have been reported in some parts of Nigeria but with varied prevalence and incidence. Most studies carried out involved seroprevalence studies. In any case, prevalence of HCV has been reported among different populations; including co-infected patients (Adewole *et al*., 2009).

In Nigeria, genotypes 1 and 2 have been found in a few molecular studies (Forbi *et al*., 2012; Oni *et al.*, 1996), with other genotypes such as 3,4, 6 identified but with low prevalence (Okwuraiwe *et al*., 2012).

A seroprevalence study by Afolabi (2012) among blood donors reported prevalence of 1.4%, which was lower than previous reports (Afolabi *et al*.2012). However, Fashola *et al.,* (2008) in their own study reported prevalence of 4.96% among blood donors in UCH Ibadan. In a study among married and unmarried participants, the prevalence of 2.1% was reported among the unmarried participants than in married individuals. It was believed that this high prevalence among unmarried persons may be due to the fact that the unmarried persons are more likely to engage in risky behaviours that predispose them to infections with HCV and other blood –borne infections than the married participants (Gao*et al.,* 2011).

For other studies that were carried out in some other locations in Nigeria, a higher prevalence was reported in blood donors than in married and unmarried participants. For instance, Chukwurah *et al.*(2005) and Ayolabi *et al.*(2006) also found higher prevalence of 7.6% among blood donors in the Southeastern Nigeria, in their respective study. In Lagos, Ayolabi*et al.*(2006) reported that 8.4% of blood donors were positive for HCVantibody. Low prevalence of 1.4% of HCV antibodies in one of the studies was the lowest prevalence ever reported from Nigeria. This probably was due to the positive impact of HIV /HCV education and public enlightenment over the years (Ayolabi *et al,* 2006; Balogun *et* al., 2010).

2.6.3 Race, Sex and Age Differences in HCV Infection

HCV infection has been associated largely with poor educational achievement and economic status in minority populations especially among people of African descent and Hispanics than in other populations (Alter *et al*., 1999). Studies have shown that prevalence of genotype 1 was higher among US population of African descent than in other race (Alter *et al*., 1999).

Sex and race have not been associated with HCV infection according to national surveys carried out, there is no sex dominance occurring with HCV infection (Alter *et al*, 1999). With reference to age, age group between 30 to 49 years has recorded the highest for HCV

infections (Nakashima*et al*., 1995). It has been observed that people became when younger fare better with treatment presenting better prognosis than those who got infected in older age. It was found in a study that infection was not common among persons aged 20 years and below but it was more prevalent in the ages 40 years and above(Osella *et al*., 1997). An association between age and route of transmission has also been established, especially transmission through iatrogenic means such as use of contaminated materials during medical procedures which may include parenteral routes, drug injection and vaccination (Kiyosawa *et al*., 1994).

2.7 TRANSMISSION AND RISK FACTORS FOR HCV INFECTION

Notably, before 1990 there was no test for hepatitis C prior to transfusion of blood in patients. Blood transfusion with infected blood was actually an important means of transmission (Niederau *et al*., 1998). But since 1990, this trend has changed as a result of frequent screening of every blood sample before transfusion. CDC estimated the risk factor for HCV contracted through transfusion during 1990-93 at a 5% and risk of infection actually came down after 1993(Shapiro *et al.*, 1996).

Easy detection of virus nucleic acid in early infection has increased due to screening with very sensitive methods like polymerase chain reaction (PCR), which can reduce HCV window period after to about 1 or 2wks (Kim, 2002). Risk of infection remains high among intravenous drug users, hemophiliacs, organ recipients and through blood –transfusion procedures. Intravenous drug users, who use contaminated needles or share straws for snorting cocaine, are at highest risk for HCV infection. Notably, intravenous drug abuse is the main source of new infections in developed countries (CDC, 2012).

Another important route of HCV transmission is through occupational exposures such as needle-stick and scratches in health facilities. Again nosocomial transmission may occur through the use of medical equipment during procedures such as surgery, dialysis as well as transplantation of organs and tissues (CDC, 2012). The use of body-piercing instruments during scarification, tattooing, sharing grooming objects such as blades, shaving sticks and contaminated needles and syringes in acupuncture increases the chance of infection with HCV, as with HIV and HBV. As a precaution, use of disposable needles or avoidance of re-use will eliminate this transmission route (Cooper *et al.*, 1992).

Several studies have implicated sexual contact as means of HCV transmission. Hence, high rates of infection of up to 18% have been recorded for homosexuals, 10% among heterosexuals and 12% among commercial sex workers (Mostafa *et al*., 2010).

Yeung and colleagues (2001),reported that transmission through sexual engagement and vertical mode are less efficient means of HCV transmission. However,both sexual and maternal-fetal transmission of HCV may be increased by co-infection with HIV (Yeung *et al.*, 2001; Graham, 2015). In addition, contact with a household member with HCV is another means of transmission and this has doubled over the years (Goldberge *et al*., 1998).

2.8 PROGRESSION AND PROGNOSIS OF HEPATITIS C VIRUS INFECTION

A number of factors such as host's immune response to infection during acute phase and quasispecies diversity may contribute to progressive characteristic of HCV infection, which has resulted in considerable burden of liver diseases across various nations of the world. According to Hajarizadeh *et al,* (2013), those countries with increased incidence of infection are bound to have accelerations in HCV-related complications including cirrhosis and hepatocellular carcinoma (HCC) in about twenty years to come. Generally, hepatitis C progresses slowly into chronic stage during the first twenty years of infection, however, it can be accelerated at this time due to some host's characteristics such as increasing age, consumption of alcohol and co-infection with HIV (Hajarizadeh *et al*., 2013; Viso, 2007). Co-infection with HBV, overload with iron and deficiency of alpha 1-antitrypsin may influence chronicity hence, increase the chances of occurrence of cirrhosis, which naturally develop in about twenty years of active infection in approximately 20% of chronically-infected individuals (Niedarau *et al.*, 1998).

Studies have shown that the danger of acceleration to cirrhosis and cancer is two times higher in those patients who got infected with HCV through blood transfusion. Progression to HCC in chronic patients is faster and influenced by presence of cirrhosis, alcohol consumption, HIV and HBV co-infection (Gordon, 1998). In a study involving 128,769 patients in HCV clinical registry, McCombs *et al.* (2014) reported that decrease in HCV RNA or viral load correlates with decrease in risk of hepatic pathology resulting in low morbidity and death (McCombs*et al*., 2014 ;Tucker *et al*., 2013). Also, in patients whose viral load were suppressed, the risk for future liver complications such compensated/decompensated cirrhosis, hepatocellular carcinoma etc., as well as risk of death were reduced by 45%, compared to patients whose viral load was not suppressed (McCombs *et al*., 2014, Tucker *et al*., 2013)). McCombs *et al.*(2014) also observed that race or ethnic affiliation; as well as genotypes can actually increase liver outcomes and death. In his studies, he observed that

these risks were higher in Caucasians than in patients of African descent, while patients with HCV genotype 3 had a higher risk for all liver events compared with patients who had HCV genotype 2, who had the lowest risk or genotype 1(McCombs *et al*., 2014; Tucker *et al*., 2013). In a similar study, Samsouk *et al.*(2003) found that immunosuppressed persons including HIV-infected persons or recipients of organ transplants experience higher rates of progression to chronicity after acute HCV infection (Samsouk *et al.*, 2003).

Emerging evidence supports the view that certain biological factors such as genetics may influence the outcome of HCV infection especially in individuals with non-transfusion sources of infection including injection drug use (Khakoo *et al*., 2004). Persons with homozygous polymorphism in the inhibitory NK cell receptor KIR2DL3 and its HLA- C1 ligand which are mediators of NK cell inhibition were more likely to experience spontaneous clearance than persons who were heterozygous for those alleles. This suggests that innate immunity to HCV is basally impaired in person possessing the heterozygous alleles (Khakoo *et al*., 2004).

2.9 DIAGNOSIS OF HCV INFECTION

2.9.1 Clinical Diagnosis

Most patients with HCV infection show no abnormal clinical findings on examination prior to developing hypertension or liver disease. But in those with clinical symptoms, porphyria cutanea tarda(finger infection) or necrotizing vasculitis are common (Dhawan, 2016).

2.9.2 Signs and Symptoms of HCV Infection

Although HCV infection can be asymptomatic initially, the first few symptoms are often extra-hepatic, with common symptoms of arthritis, skin problems and muscle aches and pains. Some other symptoms are paresthesias, myalgias, pruritus, sicca syndrome, althragia, and sensory neuropathy etc (Dhawan, 2016; Cacoub *et al*, 1999).

According to Dhawan, these symptoms are common when liver disease has set in and most present during hypertension as; changes in mental status involvinghepatic encephalopathy,ankle edema and abdominal distention known as ascites,melena or hematesis (variceal bleeding).

Before the onset of portal hypertension or decompensated liver disease, physical findings usually are not abnormal. Typically, signs in patients with decompensated liver disease are seen in different parts of the body ranging from;

a. Head: Symptoms such as icteric sclera, temporal muscle wasting, enlarged parotid gland, cyanosis may be observed
b. Hand: where palmar erythema, asterixis, leukonychia, clubbing can occur.
c. Fetor hepaticus, Gynecomastia ie small testes.
d. Abdomen: Abdominal signs can present as umbilical hernia, ascites, caput medusae, hepatosplenomegaly and abdominal bruit or extension.
e. Others can be ankle edema and scant body hair while skin signs can involve spider nevi, petechiae, excoriations due to pruritus.

A few other common extra hepatic signs of infection can present as cryoglobulinemia, involving the kidneys as proliferative glomerulonephritis (altered filtration), idiopathic thrombocytopenic purpura, and lichen planuscan be observed (Maticic, 2007). When the eye is involved, signs can present as keratoconjunctivitis. In many situations, Raynaud syndrome, cutaneous vasculitis, Sjögren syndrome, and lymphoma (Non-Hodgkin) have been reported (Cacoub*et al.* 2000).

In 9% of persons infected with HCV, sensory neuropathy has been observed and reported as extra-hepatic sign (Cacoub *et al.,* 2000). Furthermore, extra hepatic manifestation in chronic hepatitis C infected people presents with non-specific symptoms which are also seen as skin presentations. For examples, presence of pruritus and urticarial which are discreet dermatological features can give clues into underlying HCV infection (Cacoub *et al.,*).
This can be observed in post- transfusion infection, surgical infection, infection through injection of drugs, or wound onthe nasal mucosa from sniffing cocaine by sharing straws (Dhawan, 2016).Most often, skin problems are seen in those with ongoing disease and this may finally result in organ failure. Those with advanced (end-stage) liver disease may show pruritus, dryness, palmar erythema, and jaundice. Often times, the above symptoms are used as evidence for further examination in a person. In most cases, bowel symptoms are seen in patients already having mucosa-associated lymphoma tumors (MALT) (Dhawan, 2016).

\

2.9.3 Laboratory Diagnosis

2.9.2.1 SEROLOGY

Initial laboratory diagnosis of HCV can be done with serological screening which involves an enzyme immunoassay (EIA); including ELISA (Enzyme-linked immunosorbent assay) commonly used to detect HCV antigen and antibodies. ELISA procedures are 97% specific although it typically does not differentiate acute and chronic phases of infections (Dhawan, 2016). For confirmation of HCV infection, recombinant immunoblot assay can be used. An immunoblot assay positive result can be defined as the detection of HCV antibodies against two or even more antigens, while an assay result is indeterminate when antibody is detected against a single antigen or none. However, when a positive immunoblot assay results in two or more undetectable HCV RNA consecutively, it then suggests that HCV infection is resolved in that individual. Typically, when a positive HCV antibody immunoassay result is followed by a negative immunoblot assay result, this does not indicate a true positive (false-positive) immunoassay and based on this, additional test is not needed (USFDA, 2010).

Third-generation enzyme-linked immunoassays can lead to detection of antibodies against the core protein and non-structural proteins 3, 4, and 5. This can give rise to positive results about eight weeks after infection. Fourth generation ELISA is the most recent and it works with very high specificity. This can be used to diagnose infection within the first 6 weeks of exposure (Dhawan, 2016).

There are other baseline tests that can be done and these include

1. Analysis or determination of total blood cell number or counts including differential.
2. General liver function tests, which must include determination of AST and alanine aminotransferase (ALT) levels
3. Thyroid function tests
4. Genotyping of HCV which is an important process that guides treatment in almost all cases
5. Quantitative HCV RNA assay, which determines the viral load

In addition, other tests can be carried out to screen for co-infection with HIV, HCV, HTLV-1, alcohol abuse, drug abuse, and/or depression.

However, in individuals whose immune status have been compromised, e.g HIV-1 infection, cryoglobulinemia in HCV infection and renal failure,false-negative results can occur. However, the possibility of a false-positive result occurring can be higher in individuals without risk and in asymptomatic persons such as blood donors and health care givers (Mukerjee, 2012).

Recently, rapid antibody screening was authorized by FDA, which can be used for at risk persons or those with symptoms typical of HCV infection (USFDA, 2010).

However, polymerase chain reaction (PCR) is best suited for diagnosis of HCV infection in those with injury from needle-stick, involving infected patient, and can be repeated every 2 months for a period of 6 months (WHO, 2014).

2.9.2.1.1 Cryoprecipitation

Cryoprecipitate from infected individuals usually contains large amounts of HCV antigens and antibodies. Cryoglobulins are found in almost all infected persons. Mixed cryoglobulinemia is essentially occurs as a result of HCV infection. Della et al reported that 90% of affected persons have HCV viremia (Della *et al.*, 2001). When symptoms such as vasculitis, arterial hypertension, purpura, lichen planus, joint aches (arthralgias) and decreased thyroxine levels are observed, typically these are associated with titers positive for cryoglobulin.

2.9.2.2 MOLECULAR DIAGNOSIS

2.9.2.2.1 Qualitative Assays for HCV RNA Detection

These are mainly used to test for presence of HCV RNA in an individual. This is done by the use of amplification process such as PCR to detect HCV RNA in patient's blood. Current PCR-based qualitative assays approved by FDA for qualitative HCV RNA detection include:

1. Testing with automated Cobas Amplicor machine manufactured by Roche Molecular Systems.
2. Versant HCV RNA Qualitative Assay by Bayer HealthCare

2.9.2.2.2 Quantitative Assays for HCV RNA

These assays can detect HCV RNA (viral load) in blood. This method works based on amplification by signaling (through branched DNA ie bDNA assay) and target amplification using PCR, TMA, TaqMan).

Comparatively, reverse transcription PCR (RT-PCR) is more sensitive than branched DNA method. The quantityof HCV RNA in any blood sample can help to predict possibility of the virus responding to therapy (USFDA, 2010). For purpose of accuracy, it is important to be

consistent with a particularquantitative test throughout period of therapy. This is in order to monitor response to therapy effectively in relation to a standard testing method. As at 2010, Versant HCV RNA assay, version 3.0 made by Bayer HealthCare, was the only approved testing technique. This is based on bDNA technology which has a dynamic range of 615-7,700,000 IU/mL (USFDA, 2010).

2.9.2.2.3 Genotyping of HCV

HCV genotyping is an important tool to determine the different existing forms of a virus. It becomes very useful when trying to predict patients' response to treatment and also duration of therapy, since in most cases, successful treatment with any antiviral may depend ontheinfectinggenotype (Lok *et al,* 2012).

The method can be by directly sequencing the regions of the genome which may range from use of reverse hybridization method to use of oligonucleotide probes specific to each genotype. Another method that can be used is the restriction fragment length polymorphisms (RFLPs) (Dhawan, 2016). The Abbott Real Time HCV Genotype II test was approved by FDA in 2013. The method can distinguish HCV genotypes. Approval was based partly on theability to compare the accuracy of the assay with sequencing method (Crane, 2013; USFDA, 2013).

There are other tests available for genotyping HCV genome which isnot yet approved by the FDA. These include:

a. Use of Trugene HCV genotyping (produced by Visible Genetics; Toronto, Canada): This method is based on principle direct sequencing of a gene, before comparing with reference sequences.

b. The Line probe assay by Inno LiPA HCVII, Innogenetics; Ghent, Belgium. The principle is based on reverse hybridization of amplicons or PCR products on a nitrocellulose strip, which has been glazed with genotype-specific oligonucleotide probes (Dhawan, 2016).

2.9.2.2.4 Single-nucleotide Polymorphism (SNP)

It is very obvious that genetic composition of patients play important part in the responseto antiviral therapy. It is observed that single-nucleotide polymorphism (SNP) known as rs12979860, which is located close to the IL28B gene on chromosome 19, andalso encodes a type III interferon, is associated with increased variance in rate of virological response to antiviral such as pegylated interferon and ribavirin. Thus, SNP can be detected by initial PCR

procedure. The method can predict treatment outcome (sustained virologic response) or failure irrespective of infecting HCV genotype (Pineda *et al*, 2010).

2.9.2.3 Liver Biopsy

When trying to assess level of hepatic pathology in patients infected with HCV, liver biopsy is mainly used as a test procedure.However, biopsy is only recommended when there is d uncertainty of diagnosis; in co-infected patients; in patients with normal liver enzyme test results but have no extra-hepatic symptoms and are being considered for treatment or when the patient is immunocompromised(Boyer *et al.*, 2002).

2.9.2.4 Histologic (Pathological) Testing

During histological investigations, lymphocytic infiltration, inflammation and necrosis as well as fibrosis are observed. In most histological testing, reformative nodules appear in patients with cirrhosis, and again some patients may have features indicative of hepatocellular carcinoma (Dhawan, 2016).

Disease can be graded or staged such as in fibrosis. In histological testing, several scoring structures can be applied. These include the Ishak (6-point scale) and the Knodell histologic activity index (18-point score); METAVIR scale. They are suitable for evaluating progress in histologic examinations, although not practiced often in clinical use (Bedossa and Poynard in 1996).

2.9.2.5 Radiological Studies

Radiological assessments are used to test for liver stiffness (FibroScan) as a non-invasive procedure of diagnosing liver damage or disease in chronic hepatitis C patients. However, factors such as obesity, female sex and age >52 years may yield inunacceptable test results.There have been reports of high estimates of liver fibrosis and food intake resulting in acute inflammation. FDA approved Hepatiq radiologic image processing system in 2014(Boyer *et al,* 2002; Business wire, 2015).

The principle of Hepatiq software is based on use of measurable procedure of analyzing liver and spleen pictures in order to detect the level of liver disease and to also forecast outcome of disease (Hepatitis C Antiviral Long-term Treatment against Cirrhosis] trial (Business wire, 2015).

HALT-C trial is a randomized but controlled study which evaluates if long-term interferon therapy can subdue HCV infection, or prevent progression to cirrhosis and HCC or able to decrease chances of transplant (HALT-C, 2015).

2.9.2.6 Virus Isolation in Cell Culture

In past years, studies on HCV isolates was hindered by scarcity of suitable animal model and an infective cell culture system, which only became established in 2005, following the discovery of JFH1 (genotype 2a), which has a genome that could enable infection in Huh7.5 (human hepatoma 7.5) cells (Li *et al.,* 2015). HCV was first identified in 1989 when prototype strain HCV-1(H77) genome was cloned. Several attempts have been made previously, to successfully infect cells *in vitro* with HCV (Pawlotsky, 2006). However, two HCV clones in 1997, belonging to H77 prototype strain became infectious inchimpanzee butcould not infect cells*in vitro*.In a study conducted by Li et al.,(2015), they developed an infectious clone for genotype 1a (TN), 2a (J6), and 2b (J8, DH8, and DH10) strains. They discovered important changes in these strains. With these, they were able to infect cells *in vitro*. Their findings ushered in innovative ways isolation of HCV isolates of different genotypes in adapted culture.

Afterwards, several studies have shown the usefulness of hepatoma cells in the study of HCV life processes. Notably, infection with a single strand RNA virus of a positive sense results in reorganizations of intracellular membranes, which precedes the development of a replication medium or complex that links viral proteins, cellular constituents andnewly-formed RNA strands (Gretton, 2005).

Brimacombe*et al*. (2014) also reported migration of hepatoma and their invasion after theappearance of CD81 which was followed by a reduction in invasive potential upon CD81 quieting or suppression. This indicates that HCV infection disturbs CD81-dependent hepatoma spread (Brimacombe *et al.,* 2014; Farquhar 2012; Harris, 2013).

Hence, advancement in producing a healthy cell culture system for HCV infection unraveled great insight into successful replication of HCV. This in turn, facilitated antiviral drug production targeting different stages in the viral life cycle.

2.10 THERAPY AND MANAGEMENT OF HEPATITIS C INFECTION

There are different guidelines for treating chronic HCV infection successfully. These guidelines tend to vary from place to place, depending on the severity of infection. The guidelines were collectively put together by Infectious Diseases Society of America (IDSA) and the American Associations for the Study of Liver Diseases (AASLD), in collaboration with the International Antiviral Society-USA (IAS-USA) (AASLD/IDSA/IAS Guidelines, 2015) and World Health Organizations (2015).

According to the guidelines, therapy is to be provided with considerations of the extent of need by infected persons, with priority to be given to individuals with urgent need for treatment (AASLD/IDSA/IAS Guidelines, 2015).

a. Urgent treatment or precedence to be given to infected individuals who have advanced fibrosis, cirrhosis, recipients of liver through transplantation, and patients with severe extra-hepatic presentations.
b. This is followed by patients at danger of complications due to liver-related or severe extra-hepatic presentations.
c. Important steps should be taken to treat individuals at risk of transmission of infections to others. This group includes homosexuals, injection drug addicts and patients undergoing hemodialysis.
d. Treatment should be administered earlier to patients whose stage of fibrosis is still low in order to achieve aim of HCV cure.
e. There are basically two important objectives for treating HCV infection; the first is to bring about total HCV cure or sustained virologic response in patients. This can be achieved when HCV RNA cannot be detected in patient's sample 48 weeks or more after completion of therapy. Another objective is to avoid development of complications of liver disease such as cirrhosis, liver cancer and liver transplant.

Ideally, commencement of treatment should be based on individual patient severity of disease. Nonetheless, those having increasedserum aminotransferase levels orgreater than eighteen years, anti-HCV positive HCV test or detectable HCV RNA,compensated liver pathology ie no abdominal swelling (encephalopathy) or ascites, haematological and biochemical values which range from more than 13 g/dl for women and 12 g/dl for men. Other laboratory indicators can be neutrophil count of greater than 1500/mm^3, serumcreatinine of greater than1.5 mg/dl), readiness to receive and also stick to treatment regimen as well as absence of serious impediment to therapy (Ly *et al.,* 2007).

Suppression of viral load irrespective of the infecting genotype decreases the danger of hepatitis C liver disease and death(Tucker, 2013; USFDA, 2013).

It has been observed that racial or ethnic background of patients as well as HCV genotypes can determine outcome with liver pathology and death in patients. This was demonstrated in a study by Tucker*et al*. (2013), where the danger of liver disease complications and death was greater in Caucasians than in in patients with African descent. Again, it was reported that individuals infected with HCV genotype 3 had severe diseaseconsequence those infected with genotypes 1 and 2(Tucker, 2013; USFDA, 2013).

2.10.1Antiviral Treatment

The treatment of hepatitis C keeps evolving until recently when direct acting antiviral therapy became available. Initial studies were based on IFN monotherapy. Subsequently, combination of ribavirin which is an oral nucleoside analogueand IFN which has added polyethylene glycol (PEG) molecules, that is PEG-IFN, were introduced. This was followed by design of combination of protease inhibitors. Boceprevir (Victrelis) was the foremost protease inhibitor that was licensed for treatment of HCV infection. It was licensed by Food and Drug Administration in 2011. Another protease inhibitor after Boceprevir was Telaprevir. These two protease inhibitors however, were not endorsed because more effective options became available.In November 2013, Simeprevir (Olysio) was also licensed for use and on its recommendation, became part of combination therapy for treatment of chronic hepatitis C (USFDA, 2013).

Polymerase inhibitor Sofosbuvir (Sovaldi) was introduced last, and treatment with polymerase inhibitors can inhibit HCV replication. Furthermore, in2014, an oral Simeprevir plus Sofosbuvir regimen for treatment-naïve or treatment-experienced patients was made available for use.For patients without cirrhosisperiod of treatment is 4 monthsand 6 months for those with cirrhosis (Lawitz *et al.,*2014).

2.10.1.1 Polymerase Inhibitors

These antiviral agents target specific enzymes involved in viral RNA replication. HCV NS5B polymerase is involved actively in HCV replication.Most polymerase inhibitors target this enzyme and are termed direct-acting antiviral agents (DAAs). A good example is Sofosbuvir (Sovaldi).It was licensed in 2013 and it is used for treatment of chronic HCV infection

caused by genotypes 1, 2, 3, and 4. As a component of combination therapy, it is indicated for use by individuals with hepatocellular carcinoma or awaiting liver transplant (USDA 2013). Other infected category of people includes those with HIV/HCV co-infection and also usedto prevent HCV recurrence. Treatment with Sovaldi like other antivirals and periodof therapy also depend on HCV genotype and patient population (infected category). Generally, patients with genotype 1 or 4 are treated with combination of sofosbuvir, peginterferon alfa and ribavirin 3 months. Individuals infectedwith genotype 2 and 3 could follow oral regimen which comprises of sofosbuvir and ribavirin for 3 or 4 months respectively (Dhawan, 2016).

Numerous evaluation studies were conducted before sovaldi became available for use. Supporting data were gathered from several Phase 3 studies that assessedweeks (12 or 16) of treatment with different combination cocktails. Threeof the studies evaluated sofosbuvir plus ribavirin in genotype 2 or 3 patients who were either treatment-naïve(Lawitz *et al.,* 2013). Treatment-experienced persons or peginterferon intolerant persons and those not eligible or willing were assessed by Jacobson and colleagues (2013). In the fourth study, sofosbuvir in combination with peginterferon/ribavirin was assessed in treatment-naïve patients with genotypes 1, 4, 5 or 6.In all the studies, sofosbuvir was found to be more effective over all the controls used ((Lawitz *et al.,* 2013), to placebo (Jacobson *et al.,* 2013) and superior over current available regimen (Lawitz *et al.,* 2013). These results were rated on the percentage of patients who achieved sustained virological response after 12 weeks (SVR12) of finishing therapy. When patients achieve SVR12, proved by no detection of HCV RNA in sample, they are cured of HCV. For most patient participants on trial, taking sofosbuvir-based therapy, SVR12 rates of 50-90%, have been achieved, however no remarkable improvement has been recorded in patients with cirrhosis and HCC outside viral suppression (Lawitz *et al.,* 2013).

2.11 PREVENTION, CONTROL AND ERADICATION OF HCV INFECTION

There is no vaccine or potent immunoglobulin available for prevention or control of hepatitis c virus infection (Jawetz *et al.*, 1998). A few precautionary steps can be taken against transmission of HCV in a population. At the same time, intensified efforts are being made to

develop an effective vaccine against this virus. One of the major drawbacks to achieving this is the high genetic variability due to high mutation rate (Simmonds *et al.,* 2005).

So far prevention is targeted at the transmission route and risk factors. Infected persons are urged to desist from consumption of alcohol. Sexual contact is an established mode of transmission hence; those infected should practice barrier protection during sexual activity (Everson *et al.*, 2002).

More than four hundred million individuals are living with hepatitis B or C with no country unaffected, due to lack of symptoms at some point. Therefore, most people with hepatitis remain undiagnosed. This has resulted in death comprising 1.4 million people every year from complications of viral hepatitis and most of these death can actually be prevented (WHO, 2016).

Increasing global awareness about these infections, universal access to prevention, and availability of vaccines, testing, diagnosis, care and treatment which are fundamental human rights will control the spread and eventual elimination of infections. Hence, promoting access and affordability of these services to global communities remain the responsibility of all stakeholders including government, academia, non-governmental organizations, and health care providers etc., to eradicate and offer a lasting solution to the menace of viral hepatitis in the globe.

CHAPTER THREE

3.0 MATERIALS AND METHODS

3.1 STUDY LOCATION

This study was carried out in Ibadan, the capital city of Oyo State, Nigeria from 2012 to 2018 (Fig 2.9). The city has a population of about 5.592 million according to 2006 population census (Salami *et al.,* 2016).It is inhabited by a combination of indigenes, people from other ethnic groups in Nigeria,and other foreign nationals, especially from neighboring countries including, Niger Republic, Republic of Benin, and Cameroun.

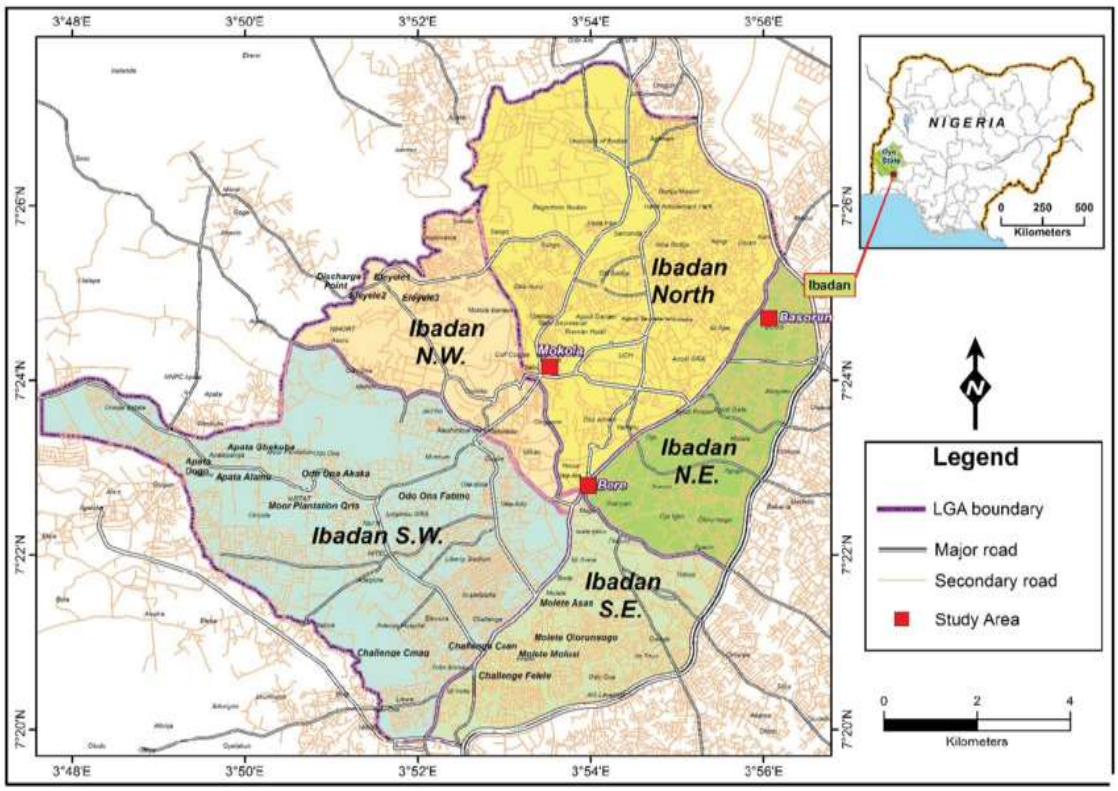

Fig. 2.9 Location of study sites in Ibadan metropolis. Source: Salami *et al.,* (2016)

3.2 STUDY DESIGN AND POPULATION

This study was cross- sectional and the participants were three subgroups including:

HCV-positive patients with clinical hepatitis (symptomatic) or liver disease; HCV-positive blood donors and HCV-positive patients with HIV (co-infected patients). The participants included both gender; and age range between 3 months to 83 years and a mean of 42 years.

3.2.1.1 Inclusion Criteria

All HCV-infected individuals (blood donors, patients and HIV co-infected persons) with positive HCV antibody test.

3.2.1.2Exclusion Criteria

Blood samples with indeterminate HCV antibody test results and HCV negative samples were excluded from the study.

3.3 ETHICAL CONSIDERATION

Ethical approval for the study was received from:

University of Ibadan/ University College Hospital Ethical Committee or institutional review board (IRB), with the assigned number UI/EC/ 14/0019; and the Oyo State Ministry of Health.

3.4 SAMPLING MATERIALS

The materials used during sampling include: personal protective equipment (PPE) and other materials such as nose masks and powdered disposable gloves, high quality cold boxes for carriage of samples, Ice gel, ice packs and ice blocks for maintaining sample cold chain, Cryovial tubes, covered and transparent racks for storage of samples, Pasteur pipettes, small and medium-sized biohazard bags, 70% alcohol, hypochlorite and absorbent paper.

3.5 SAMPLING METHOD

3.5.1 Sample Collectionand Processing

The samples used for this work wereobtained from patients referred from different clinics to the Department of Virology, College of Medicine, University of Ibadan and blood donors at Blood Bank, University College Hospital, Ibadan (UCH). Samples were collected for a period of three years (May 2014- June 2017). Three hundred and one (301) HCV antibody – positive samples were collected for the study comprising 99 patients diagnosed with clinical

hepatitis or onset of liver diseases such as patients withindications for fibrosis, cirrhosis and hepatocellular carcinoma etc, 125 HIV/HCV co–infected patients and 77 blood donors.

Plasma obtained after separation was dispensed into cryovials and immediately transferred into -80^{o} C ultralow freezer for preservation. The samples from blood bank were also transported in cold insulated box on ice each time to the laboratory where they were processed immediately or stored at -80^{o} C.About 5ml of plasma was collected from each positive blood sample after anti-HCV testing into appropriately labelled sterile cryovial tubes.

3.6 ISOLATION OF VIRAL RIBONUCLEIC ACID (VIRAL RNA)

MATERIALS USED FOR RNA EXTRACTION

These includedJena Bioscience RNA purification kit, 2-Mercaptoethanol (2-ME), 99% Ethanol, Isopropanol, Eppendorf tubes, cold centrifuge, Pipette, pipette aid and pipette tips (30µl, 100µl, 200µl,1000µl).

The components of Jena Bioscience RNA purification kit included:

Lysis Buffer
Activation Buffer
Primary wash buffer
Elution Buffer
Spin Columns
2mL Collection tubes.

3.6.1 Extraction of HCV RNA

Viral RNA was extracted from each of the plasma samples using total RNA purification kit (Jena Bioscience, Jena, Germany) according to instructions from the manufacturers. The extraction process was accomplished when the extracted nucleic acid trapped in the spin column was eluted as described below.

3.6.1.1 Reconstitution of Extraction Reagents

Before commencement of RNA extraction, the following components (not included in the kit) were added to the respective bottles according to manufacturer's instructions:

a. 1.3ml of 2-Mercaptoethanol was added to 250ml of Lysis buffer (mercaptoethanol is a reducing and antioxidant agent that removes polyphenols, and stabilizes the RNA or DNA preventing it from oxidation).
b. 40ml of 99% ethanol was added to blood washing buffer
c. 160ml of 99% ethanol was added to each of primary and secondary wash buffer to obtain 200ml for 250 preparations, respectively. This enables precipitation of RNA or DNA out of solution as a white cottony mass that can be seen with naked eye.

PRINCIPLE OF SPIN COLUMN-BASED EXTRACTION

The principle of the Spin column-based nucleic acid purification is based on the binding of the nucleic acid to the silica membrane under well controlled conditions, allowing proteins and other contaminants such as inhibitors to be removed from samples. This method is based on use of membrane composed of silica gel and the rapidity of a micro spin technique which involves three stages:Lysing of the virus capsid with lysis buffer, followed by precipitation of nucleic acid by washing to get rid of contaminants in the membrane while leaving only the nucleic acid bound to the silica gel; and lastly elution. During elution, the bound nucleic acid is detached by the buffer from the membrane, and it is collected in a fresh tube.

3.6.1.2 Procedure or Protocol for Extraction

(i) Work Area Preparation

Biosafety cabinet and work areas were sterilized with 70% alcohol. Then all the reagents and extraction materials were arranged in the biosafety cabinet. The cabinet was then switched on making sure that all the necessary conditions including airflow direction are maintained.

(ii) Activation of Spin Column

Eppendorf tubes were labelled with each sample identification number and arranged in a rack. Spin columns were aseptically inserted into the accompanying collection tubes and arranged in another rack. Thereafter, 100 µl of activation buffer was dispensed into the spin column, after which each column was centrifuged at 12,000g for 1min. The resulting supernatant was discarded.

(iii) Lysis of cell membrane

Three hundred microliter (300 µl) of lysis buffer was dispensed into the labelled Eppendorf tubes, then 100 µl of plasma from each sample was added to the corresponding tubes, after which the tubes were allowed to stay in the biosafety cabinet for 2 to 3 minutes for lysis to take place.

(iv) Precipitation of nucleic acid

Three hundred microlitre of isopropanol was added to the prepared lysate (this precipitates the DNA by engaging with the water molecule so that DNA does not dissolve in water), followed by brief mixing by vortexing the tube. About 900 µl of resulting mixture was then transferred into the spin column (the spin column traps the lysed viral RNA in the sample) and centrifuged at 12,000g for 30 sec, using cold centrifuge to preserve the integrity of the viral RNA (high temperature denatures the RNA). Supernatant (flow through) in each tube was discarded into safe collection trash and the columns fixed back into the tubes ready for washing.

(v) Washing

During the wash procedure, 700µl of primary wash buffer was added into the columns in the tubes, centrifuged at 12,000g for 30 sec and supernatant discarded, while columns were also fixed in the tubes again. Into the tubes was added 700 µl of secondary wash buffer and centrifuged at 10,000g for 30sec. After, fresh RNase-free capped Eppendorf tubes were arranged in a separate rack. The columns were then transferred into the capped Eppendorf tubes and centrifuged at 10,000g for 1min. The flow through was then discarded and the columns re-centrifuge at 10,000g for 2 minutes to let out the residual alcohol and salts in the sample.

(vi) Elution of viral nucleic acid

Columns were transferred into another set of 2ml fresh RNase-free Eppendorf tubes, then into the center of the column, 50 µl of elution buffer was added without letting the tip touch the column base. Next, these tubes containing elution buffer were incubated or left in the biosafety cabinet for 2mins, centrifuged at 10,000g for 1min to elute the RNA. Then used columns were safely discarded while the RNA that has been eluted was stored immediately at -80^{o}C in 20 µl aliquots in different and appropriately labelled PCR tubes for further use.

3.7 REVERSE TRANSCRIPTIONOF ORSYNTHESIS OF COMPLEMENTARY DNA (cDNA)

3.71 Principle of Reverse Transcription

The principle of reverse transcription or Complementary DNA (cDNA) synthesis is based on the generation of first strand DNA fromgenomic RNA, used as the template in the presence and action of reverse transcriptase enzyme. Reverse transcription PCR (RT-PCR) is used when the starting material is RNA. First, RNA is transcribed into complementary DNA (cDNA) by using reverse transcriptase, a genetically-engineered version of M-MLV Reverse Transcriptase with eliminated RNase H activity and increased thermal stability. The enzyme is a RNA-directed DNA polymerase that synthesizes a complementary DNA strand initiating from a primer using single-stranded RNA as template. The resulting cDNA is then used as the template or first strand cDNA in the synthesis of second strand DNA before amplification process of PCR.

3.72 Materials used for cDNA Synthesis

Jena Bioscience script cDNA synthesis Kit containing all components for first strand synthesis (table below), was used. Other materials used include micro amp tubes (PCR tubes), Pipette tips (10 µl, 30 µl), Pipette, thermal cycler, random hexamers (primers).

3.73 Protocol for cDNA Synthesis

The cDNA synthesis was done using a two-step SCRIPT cDNA synthesis kit (Jena Bioscience, Germany) according to the manufacturer's instructions. The components calculated and dispensed are used based on the number of reactions and assay. For a 20 µl assay, 7µl RNase-free water was aseptically drawn into a micro amp tube (PCR tube). Random hexamer (0.5µl) was added into same tube, mixed by pipetting or by gentle vortexing. Thereafter, 5µl of extracted viral RNA was dispensed and mixed, changing tips after each tube. Next the other components of the Script cDNA synthesis kit were dispensed as outlined in the table 3.1, yielding 20 µl cDNA in each tube.

Table 3 .1: Protocol for first strand cDNA synthesis

Component	**Conc. In Stock**	**Final Conc.**	**Volume in**
			20µl assay (for each reaction)
RNase -free water			Fill up to 20 µl (7 µl)
RNA Template		Total RNA:10pg	5 µl
Primer(Random hexamer)	100 µM	50pmol (100ng)	0.5 µl
Script Complete Buffer	5X	1X	4 µl
dNTP Mix	10mM each	500mM each	1 µl
DTT Stock Solution	100mM	5mM	1 µl
RNase Inhibitor	40unit/ µl	40 units	1 µl
Reverse Transcriptase	200unit/µl	100units	0.5 µl

The reaction tubes were incubated in a thermal cycler (Applied Biosystems, California, USA) using the following conditions: 42^0C for 10min and 55^0C for 60min

The resulting cDNA product was used immediately or stored at - 20^0C until subjected to further analysis.

3.8 POLYMERASE CHAIN REACTION (PCR)

3.8.1 Principle

Polymerase chain reaction is based on the principle of*in vitro* method of synthesis of nucleic acid by specifically amplifying or replicating a particular segment or region of DNA. The technique imitates an *in vivo* DNA replication that involves the Okazaki fragments (Saiki *et al.,* 1988). It involves two complementary oligonucleotide primers, that flank the DNA segment (template) to be amplified through repeated cycle of PCR steps under varying temperatures, which includes heat denaturation of the DNA (separation or unwinding of the double helix DNA strand at 94 -96^{o}C), annealing or binding of the primers to their complementary sequences or target regions at 52^{o}C -56^{o}C and extension or elongation of the annealed primers by DNA polymerase at 72^{o}C-74^{o}C. The primers hybridize to opposite strands of the target sequence and are oriented so that DNA synthesis by the polymerase proceeds across the region between the primers. Extension products are equally complementary to and capable of binding the primers, hence, successive cycles of amplification essentially double the amount of the target DNA synthesized in the previous cycle, resulting in an exponential accumulation of the specific target fragment approximately 2^n where n is the number of cycles of amplification performed (Saiki *et al*., 1988; Innis and Gelfand, 1990).

3.8.2 Materials used for PCR

1. Jena Bioscience high yield PCR Kit containing Taq polymerase in master mix and RNase-free water
2. Thermal cycler
3. Pipette tips (10µl)
4. Pipette
5. Flat-covered Micro amp (PCR) tubes
6. Sterile 2ml micro centrifuge tubes and K-tubes
7. 50µl stock of Primers

8. Sterile nuclease-free water
9. First strand cDNA (synthesized) as template

3.8.3 Procedure

NS5B gene segment of HCVlocated at positions 8275-8618 was amplified through a nested PCR protocol thus:

In the first round of PCR, 2.5µl of PCR master mix containing the Taq was dispensed into the reaction tube. Next, 0.5 µl each of the forward (F1) and reverse (R1) gene –specific primers were added, 7 µl of nuclease-free water was then added and mixed well. Next, 2.5 µl of cDNA (template) was dispensed into the tube, in a final volume of 12.5 µl. This was repeated for the number of reactions that were carried out. The tubes were mixed and spun by gentle vortexing and immediately transferred to the thermal cycler using the PCR conditions in table 3.2. The nested PCR (second round) involved the use of 2.5µl of PCR master mix containing the Taq in a reaction tube, followed by addition of 0.5 µl each of inner primers - forward (F2) and reverse (R2) and 7 µl of nuclease-free water and mixed well. Next, 2.5 µl of PCR product from first round (template) was dispensed into the tube, in a final volume of 12.5 µl, with the tubes mixed, vortexed and transferred to the PCR machine (thermal cycler) using the cycling conditions in the table below.

Table 3.2: Primer Sequences and cycling conditions

Nested PCR	Primer Sequences	Cycling Conditions	Expected Size of Amplicons
1st Round	Forward: 5` TGGGGATCCCGTATGATACCCGCTG CTTTGA Reverse: 5`GGCGGAATTCCTGGTCATAGCCTC CGTGAA	95°C for 5min, 94°C for 30sec, 50°C for 30sec, 72°C for 45sec,72° C for 10min, for 30cycles	400bp
2nd Round	F2- CTCAACCGTCACTGAGAGAGACAT R2- GCTCTCAGGCTCGCCGCGTCCTC	95°C for 5min, 94°C for 30sec, 50°C for 30sec, 72°C for 45sec,72° C for 10min, for 45cycles	300bp

3.8.4 Agarose Gel Electrophoresis of PCR Products

The amplified products from the second round PCR were visualized and identified using 2% agarose gel electrophoresis.

3.8.4.1 Principle of Agarose Gel Electrophoressis

The gel principle is based on the separation of DNA using agarose gel electrophoresis, in which DNA is loaded into pre-cast wells and currentapplied.DNA being negatively charged due to its phosphate backbone, will migrate towards the positive electrode (anode) when current is applied. Agarose gel electrophoresis enables separation of DNA fragments by size in agarose gel.

DNA molecule has an unvarying mass /charge ratio. This enables DNA fragments to be parted based on size in agarose gel, in such a manner that makes the distance each fragment moves becomes inversely-proportional to the log of its molecular mass (Lee *et al.,* 2012). Hence, rate of migration of DNA molecule within the gel is determined by size of the DNA molecule(smaller fragments of DNAmove faster), agarose concentration, DNA conformation, voltage applied, type of agarose, PH of the buffer. Addition of ethidium bromide (a chelating agent that binds DNA and glows under UV light), shows the DNA in the gel.

Materials used in gel preparation and gel electrophoresis

The gel was prepared using Agarose, Tris- base, boric acid,Ethylene Diamine Tetracetic Acid EDTA), Evergreen gel stain or ethidium bromide, loading buffer and sterile de-ionized water, Gel Tank orgel chamber**,** gel combs,Jena Bioscience DNA Ladder (Mid-range DNA Ladder 100bp to 3kb linear scale ready-to-use), sterile aluminum foil, magnet and magnetic stirrer, microwave, Pipette and Pipette tips.

3.8.4.2 Preparation of Reagents

The most common gel running buffers are (a) TAE consisting of 40mM Tris-acetate and 1mM EDTA when prepared and (b) TBE consisting of 45mM Tris-borate and 1mM EDTA when prepared. Either of these can be used in gel preparation and as running buffer in electrophoresis apparatus.

1a. Preparation of 50X Tri-acetate-EDTA (TAE Buffer)

Tri-acetate-EDTA (TAE) buffer can be used as a flowing or running buffer in gel electrophoresis and as a component in agarose gel preparation. To make 50X stock solution, 242g Tris –base is dissolved in deionised water, then 1ml of glacial acid and 100ml of 50mM EDTA(with pH of 8.0) were added. The final volume is made up to 1 litre by adding the appropriate volume of deionised water. The stock solution was then diluted in ratio of 1:50 with deionised water to make a working solution of 1X. This working solution contains 40mM Tris-base, 20mM acetic acid and 1mM EDTA.

1b. Preparation of 10X Tris-borate EDTA (TBE Buffer)

Tris-base –EDTA (TBE) buffer has the combination of Tris-base (maintains pH), boric acid and EDTA (chelating agent that chelates Ca and Mg ions which might affect the nucleic acid). Thebuffer (prevents DNA or RNA from degradation) was prepared by dissolving 10.8g Tris- base, 5.5g boric acic and 4ml EDTA(pH of 8.0) in 1L of deionized or sterile distilled water, in a sterile conical flask. A small magnetic bar was gently dropped into the flaskand stirred on the magnetic stirrer for 60mins. The conical flask was then covered with a sterile aluminum foil, ready for use.

2. Preparation of Evergreen Gel Stain

Evergreenstain is a non-carcinogenic dye used to enhance DNA bands by making them more visible under UV light in agarose gel electrophoresis. As with most dyes, it is sensitive to light hence stored under a dark shield during and after use. The stain was prepared by pipetting 5µl of evergreen stain obtained from Jena Bioscience into a sterile eppendorf tube, then 100µl of deionized water was dispensed into same Eppendorf tube and mixed. This yielded a 20X evergreen stain ready for use in agarose gel preparation.

3. Loading Buffer

The loading buffer contains a dye that enables the PCR productor DNA to settle at the base of the well and glows in the gel as it runs through during electrophoresis; while the in-built dye makes PCR productdenser than the running buffer. In some non-premixed PCR products (without coloured master mix), it is usually mixed with DNA samples before loading in agarose gel. The buffer was prepared by mixing 3ml of 30% glycerol, 25mg of bromophenol blue dye and 10ml of deionised water. When needed, 1µl of the loading buffer was mixed

with 5µl of each PCR product before loading in gel. Dyes and reagents containing dye must be protected from light due to their sensitivity to light, in order to obtain accurate results.

4. **Preparation of 2% Agarose Gel**

Agarose powder (molecular grade) used is isolated from high quality seaweed genera *Gelidium* and *Gracilaria*, consisting of repeated agarobiose (L-and D-galactose) subunits(Lee *et al.,* 2012).

Two grams of agarose powder was diluted in 100ml of 1x working solution of TAE buffer to prepare a 2% Agarose gel. The two were mixed by swirling gently and the agar melted in microwave for 2 minutes and allowed to cool to 45^0C before gelation. To the melted agar was added 5µl of evergreen gel stain,before pouring it into a gel plate containing template combswhich are mounted vertically to make wells into the gel. The agar was allowed to solidify for about 20 to 30 minutes before the gel comb was carefully removed without allowing the solid gel to crack.

5. **Loading of PCR Products (DNA) in Agarose Gel**

The gel was placed in theelectrophoresis tank with TAE buffer poured overit to submerge the gel.Five microliter of each PCR product was loaded into thewells and 5µl of standard DNA marker was also loaded to the first or last lanes of the gel. Electrical current of400 milliamperes (mA) and voltage of120 volts were applied for 30 minutes to allow the DNA to migrate towards the positively-charged anode. Due to the fact that phosphate backbone of the DNA (and RNA) molecule is negatively charged, therefore, when it is placed in an electric field, the DNA fragments will migrate to positively charged anode. At the completion of the electrophoresis, DNA bands were visualized and photographed under the ultraviolet light using the Trans-illuminator. DNA bands seen at the expected band position with reference to DNA ladder are positive. Those wells without DNA bandsor with bands but not on the expected band position were regarded as being negative.

3.9 PURIFICATION OR CLEAN-UP OF AMPLIFIED PRODUCTS

At the end of gel electrophoresis, allamplicons from the nested or second round PCR with expected band size of 300bp were purified for sequencing using EXOSAP purification method to remove contaminating primers and dNTPs, eliminate spin columns and conserve the PCR sample for direct sequencing, according to manufacturer's instructions (Werle *et al*., 1994).

3.9.1 Components of EXOSAP kit

Exonuclease 1

Fast Ap^{TM} Thermosensitive Alkaline Phosphate or Shrimp Alkaline Phosphate

3.9.2 Procedure:

In a single-step, one-tube clean-up method, 10µl of PCR product (nested PCR) was dispensed into a 2ml reaction tube. Next, 2 µl of shrimp alkaline phosphatase (SAP) and 0.5µl of Exonuclease 1 were added into the same tube. The contents of the tube were gently mixed well by vortexing . Then the tube was incubated at 37^{o}C for 15mins. At the end of 15mins, the reaction was halted by heating the mixture at 85^{o}C for 15mins. The tube was allowed to cool and then kept at 4^{o}C prior to sequencing.

3.9.3 Quantification of the Purified PCR Products before sequencing

Quantification was carried out on purified products to determine the concentration of DNA in the sample, especially for PCR products with low quality bands during gel electrophoresis. This helps to determine the suitability of amplified products for sequencing and minimizes wastage of reagents and other resources. Nano drop curette spectrophotometer was used for quantification allowing for fastmeasurement of DNA in micro-volume. The equipment was calibrated by blanking with elution buffer. 1µlof elution buffer was dispensed into the Nano-drop curvet. Thereafter, 1µl of each purified PCR product was dispensed into the Nano-drop. Theconcentration of DNA in the product is given by the value displayed on the facility.

3.10 SEQUENCING OFHCV NS5B GENE

3.10.1 Principle of Sanger Sequencing (Population Sequencing)

DNA sequencing is a technique used to determine the exact order of nucleotides within a DNA molecule. It includes any method or technology that is used to determine the order of the four bases- adenine, guanine, cytosine, and thymine in a strand of DNA.

Sanger sequencing is a method of DNA sequencing first commercialized by Applied Biosystems. The principle is based on the selective incorporation of chain-terminating dideoxynucleotides by DNA polymerase during*in vitro*DNA replication. During sequencing reaction, fluorescent dyes are added to the reactions and a laser within an automated DNA sequencing machine is used to analyze the DNA fragments produced. It was developed by

Frederick Sanger and colleagues in 1977 (Sanger et al., 1977). This classical chain-termination procedure uses a single-stranded DNA template; primer, DNA polymerase, deoxynucleosidetriphosphates (dNTPs), as well as modified di-deoxynucleotidetriphosphates (ddNTPs) (terminate DNA elongation or extensionduring reaction). These chain-terminating nucleotides lack a 3'-OH group required for the formation of a phosphodiester bond between two nucleotides, causing DNA polymerase to cease extension of DNA when a modified ddNTP is incorporated. The ddNTPs may be radioactively or fluorescently labeled for detection in automated sequencing machines.

3.10.2 Cycle Sequencing of Purified PCR Products

3.10.2. 1 Principles of Cycle Sequencing

The principles of DNA cycle sequencing are similar to PCR. Cycle sequencing is a linear amplification process and that makes it less susceptible to contamination. The key difference is that only one primer is used in each cycle sequencing reaction, making the amplification linear and not exponential. Secondly, dideoxynucleotides are used which interrupts the extension of the DNA strands when incorporated. Cycle sequencing is a modification of the traditional Sanger sequencing method. The dideoxynucleotides are used in polymerization reaction to create a nested set of DNA fragments with dideoxynucleotides at the 3'terminus of each fragment. It employs a thermostable DNA polymerase which can be heated to 95^{o}C and still retains its activity unlike in Sanger method.In cycle sequencing, two master mixes are used but with only one primer.

3.10.2.2 Procedures

Cycle sequencing was initiated using the ready –to-use DNA cycle sequencing kit by Jena Bioscience which is based on the use of dye terminator (di-deoxynucleotide chain termination using fluorescent-labelled primers). The reaction mixtures for sequencing and the volumes used are shown in table 3.3.Two master mixes were made with the following in separately well-labelled tubes for each primer excluding the template, 2µl of sterile distilled water, 4µl of sequence buffer, 1µl of primer, 1µl of di-deoxynucleotide (dye terminator). Forward primer was added in one tube and reverse primer in another. Each was appropriately dispensed in the wells (micro amp plate) or strips labelled F and R (Forward and Reverse) using the same primers used for nested PCR (NS5B-F2 and NS5B –R2). To each well, template or purified PCR product was added, changing pipette tips after each use and closing the top of each well after adding template to keep track of each well where template was last

added. Plate was visually inspected for bubbles, and centrifuged briefly if bubbles were present and plate inserted in a DNA thermal cycler (Gene Amp). Cycle sequence program was run for:

95 ^{o}C, 5min (initial denaturation step), then 45 cycles of the following 3 steps:

94^{o}C, 30sec (denaturation)

50^{o}C, 30 secs (primer annealing)

72^{o}C, 45 secs (polymerization), then72^{o}C, 10mins (final denaturation step), 4^{o}C hold.

Table 3.3: Preparation of master mix for sequencing

Kit Components	**concentration In Stocks**	**Concentration In PCR**	**Volume in Sample(µl)**
Sterile distilled Water			2µl
Sequencing buffer	10^x	1^x	4µl
Sequencing polymerase		4units/ µl	1µl
primer	10µl	0.25µM	1µl
Dye Terminator	80µM each	8µM	1µl
purified PCR Product		5ng/ µl	8µl

3.10. 4 Post-Cycle DNA sequencing cleaning -up

Before injecting on ABI 3500 XL (ABI V3.1 Big dye terminator) genetic analyzer, the cleaning of cycle sequenced products was done using Zymo research (ZR) DNA sequencing clean-up kit for rapid removal of post-cycle sequencing reaction contaminants which include unincorporated fluorescent dyes, residual salts, dNTPs, primers and enzymes from DNA extension products which can interfere with the quality and signal strength of sequencing data. Particularly, unincorporated dyes can result in dye picks which may obscure portions of the sequencing chromatogram and interfere with the accuracy of sequencing analysis software. Thus, DNA post cycle clean-up or purification ensures complete elimination of dye blobs for high quality Phred scores and long read lengths.

3.10.5 Zymo Research (ZR) Kit components

Sequencing Binding Buffer

Sequencing wash buffer

Zymo-spin^{TM} IB columns

Collection tubes

3.10.6 Procedure

At the end of the cycle, the micro amp plate was removed from the thermocycler and the seal (foil) was removed. The sequenced samples were carefully transferred to Zymo spin columns for removal of contaminants. Three steps are involved in purification of Sequenced DNA sample + sequence binding buffer. These include spinning, washing and elution. Five hundred microliters of 0.1%Hydrochloric acid (HCL) was carefully dispensed into spin column without touching the base and incubated at room temperature for 30 minutes. The column was centrifuged for 30 seconds at 13,000rpm (15,000-16,000 xg) for thorough washing. The column was then allowed to dry overnight at 37^{o}C.DNA was eluted with a small volume of water or loading dye containing formamide before analyzing with a genetic analyzer (equipped with capillary electrophoresis instrument, 8-capillary-3500Dx system or 24-capillary-3500xL Dx system array and POP 6 or 7 polymer, DNA sequencing reagents and consumables for system qualification, integrated software for instrument control, data collection, quality control and auto-analysis of sample files and linked to a flat- screen monitor computer workstation.

3.11 EDITING OF HCV SEQUENCES

The ABI sequence data generated from sequencing (ABI format) were opened with Finch TV version 1.4 software.The chromatogram or electropherogram of each sequence was inspectedand edited onCLC sequence reader and cross-checked with Bio Edit software version 7.0.5. Sequences were imported into CLC workbench and instructions on editing were followed on how to excise, insert and delete nucleotide bases (ACGT) from sequences where needed to obtain pure, wobble-free sequences for bioinformatics analysis. The consensus sequences obtained were blasted in National Center for Biotechnology Information (NCBI) to determine HCV reference sequences with closest matching identity or relatedness to our sequences.

3.12 ANALYSES OF SEQUENCES USING BIOINFORMATIC SOFTWARES

Each bioinformatics software (CLC workbench, MEGA 7.0, BIOEDIT 7.0.5, and NETPHOS 3.1) was downloaded and saved for use in corresponding analysis.

3.12. 1 Phylogenetic Analyses

Phylogenetic analysis of the 42 isolates was performed to evaluate the genetic relatedness of the isolates to other sequences from other parts of the world. Molecular Evolutionary Genetic Analysis version 7(MEGA 7.0) software (Kumar *et al.*, 2015) was used for phylogenetic analysis. All the HCV sequences from this work together with HCV reference prototype sequence H77 and GeneBank sequences from different continents including Africa, North America, Europe, Australia etc., spanning 300-310nt of the NS5B gene {HCV H77 position, 8275-8618 (GenBank, NC.004102.1)} obtained from HCV sequence Database (Markov *et al.*, 2009), were aligned onMEGA 7.0 . The phylogenetic tree was constructed using the Neighbor-Joining method of reconstruction (Saitou and Nei, 1987).

Viral diversity as well as the ancestral relationships between study sequences and the reference sequences was investigated by running a Neighbor-Joining phylogeny with Maximum Composite Likelihood model. Test of Phylogeny that is, the percentage replicate trees in which the associated taxa clustered together was performed with Bootstrap replication of 1000 and branch support values of >60%.

3.12. 2 Alignment of Nucleotide Sequences

3.12.2.1 Nucleotide Substitutions

Alignment of the nucleotide sequences with all GenBank reference sequences was done on MEGA. From the aligned nucleotide sequences, nucleotide substitution matrix was calculated using Tamura-Nei model (Tamura and Nei, 1993), to estimate base substitution pattern in the sequences. Evolutionary analyses were conducted in MEGA7.0. The estimation was based on a total of 414 positions with 86 nucleotides on MEGA 7.0.

3.12.2.2 Evolutionary Divergence between Sequences (Pairwise)

In order to estimate the genetic distances between the isolates and their level of divergence, a pairwise distance analysis of all the sequences to determine the p-distance was performed.Each sequence was compared in pairs with prototype HCV reference H77 to determine the P- distance. The analysis compares frequencies of p-distance within and between genotypes using both transition and transversion substitution types. The evolutionary divergence was computed using Maximum Composite Likelihood method (Tamura et al., 2004). Evolutionary divergence was estimated still on MEGA 7.0.

3.12.3 Amino Acid and Protein Translation

The nucleotide sequences were translated into proteins made up of different amino acids translated from each codon (e.g codons TCT, TCC,TCA,TCG when translated give serine(S) amino acid), which can be present with numerous other amino acids in a protein sequence. . Translation was done on MEGA 7.0. Amino acid composition and frequencies were generated on MEGA 7.0 in each isolate and for all the study populations. Mutations which occur at amino acid level were sought in the symptomatic and asymptomatic groups, represented by insertions and deletion especially in variable sites. Each major mutation was recorded in the groups. Conserved sites were also noted in the two groups.

3.12.4 Statistical Analysis

Statistical differences between the two groups were evaluated by the Fisher's Exact test, according to a computerized SPSS version 20 program at Exact Chi^2 = 0.05. Values with $p<0.05$ are significant.

3.13 POST-TRANSLATION MODIFICATION ANALYSIS

HCV NS5B is a phosphoprotein because it traps phosphate or phosphoryl group ($PO3^-$) catalyzed by a host of enzyme kinases. Phosphorylation of amino acid residues (mainly serine, threonine and tyrosine) is a post-translational modification undergone by HCV proteins which may influence disease progression or response to antiviral drugs (Toivola, 2004). It alters NS5B structure, and this in turn affects expression and function of that particular protein during HCV RNA replication and general life cycle.

3.13.1 NS5B Protein Phosphorylation

3.13.2 Principle of Protein Phosphorylation

Phosphorylation of proteins is catalyzed specifically by protein kinase and phosphatases (enzymes) by transferring phosphate group from adenine triphosphate (ATP) to serine, threonine or tyrosine amino acid residues on protein peptide substrates, directly affecting the activity of the protein (Han et al., 2014). Phosphorylation plays important role in regulating the structures and functions of proteins (regulatory mechanism). A major modification process undergone by HCV NS5B is phosphorylation or addition of covalent bond phosphate group making HCV NS5B a phosphoprotein. Hepatitis C viral protein is usually processed by most cellular and viral proteases into the structural (C, E1, E2 and p7) and non-structural (NS2, NS3, NS4A, NS4B, NS5A and NS5B) proteins. HCV NS5B protein is about 68-kDa and it is the viral RdRp, a key component of the RNA replicase complex linked to other non-structural proteins (Bartenshchlager et al., 2000). RdRp is phosphorylated at its N-terminal finger sub-domain (of amino acid position 1 to 187) (Kim *et al*., 2004) and this results in conformational changes that affect replication of HCV RNA.

3.13.3 Analysis

NetPhos 3.1 web server software was used to predict phosphorylation sites in NS5B proteins (P68) of all the HCV isolates by protein kinases. The analysis was carried out to determine phosphorylated sites in each isolate, pattern of this biochemical process in the isolates responsible for infections in the three subgroups, which might have affected protein function and expression during HCV replication and consequently outcome of HCV infection in the infected individuals. Each protein sequence was imported into Netphos 3.1 software and submitted for analysis. Predictions were based on output scores that crossed the threshold of

0.500. The predictions were made for amino acids serine, threonine and tyrosine. The phosphorylated sites were noted at the Serine (Ser -S), Threonine (Thr-T) and Tyrosine (Tyr-Y) sites in all the HCV isolates. Output scores ranged from0.000-1.000, with 0.500 as the threshold. The higher the output scores the higher the confidence of the prediction and the higher the chance of site being the true phosphorylation site (Blom *et al*., 1999; Blom *et al*., 2004). Amino acid residue with highest score above the threshold of 0.500 is the predicted phosphorylated site for the isolate.

CHAPTER FOUR

4.0 RESULTS

4.1 SAMPLE COLLECTION

A total of 301 HCV antibody positive samples were collected from 99 patients with clinical hepatitis (symptomatic) or indications for liver diseases (patients with indications for fibrosis, steatosis, cirrhosis and hepatocellular carcinoma etc); 125 HIV –infected individuals; and 77 asymptomatic blood donors over a period of three years (May 2014 – June 2017). There were 168 males and 133 females, age 3 months to 83 years and a mean of 42 years.

4.2 MOLECULAR CHARACTERIZATION OF HCV NS5B GENE

4.2.1 Amplification of HCV NS5B gene and Sequencing Results

The NS5B genes were amplified in 60 HCV isolates (20%) of the 301 samples that were analyzed. Forty–two isolates (13.95%) yielded readable sequences. The remaining 18 (5.98%) either had very poor reads or excessively wobbled chromatograms, hence did not yield readable consensus sequences. A representation of gel pictures indicating the DNA bands are shown in Figure 4. 1.

HCV sequences obtained were aligned in Molecular Evolutionary Genetic Analysis 7 version (MEGA 7.0).

OBJECTIVE 1: To determine the diversity of hepatitis C virus isolates among patients and blood donors in Southwest Nigeria.

4.2.2 Diversity of study HCV NS5B Sequences (Phylogeny)

Analysis of the 42 NS5B genes revealed co-circulation of six subtypes including; 1a, 1b, 2b, 2c, 3a and 5a of four genotypes 1, 2, 3 and 5. Table 4.1 shows the distribution of all the subtypes among the general study population. Twelve (28.6%) of the 42 isolates were subtype 1a, of which 8(66.7%) were found among blood donors and 2 (16.7%) each was found among patients with HIV and Patients with clinical hepatitis, respectively. Five (11.9%) of the 42isolates were subtype 1b, found only among patients with clinical hepatitis. Four isolates (9.5%) were subtype 2b, found only among patients with clinical hepatitis. One isolate(2.4%) was subtype 2c, found only in a blood donor, while 1(2.4%) was subtype 3a

found only in a patient with HIV. Nineteen (45.2%) of the isolates were subtype 5a, of which 10 (52.6%) were among patients with HIV, 5(26.3%) found among patients with clinical hepatitis and 4(21.1%) among blood donors. This result shows that predominant strain among the entire study population was subtype 5a (45.2%), followed by 1a (28.6%).

Fisher's Exact test statistics (Exact Chi^2 = 0.05) showed significant difference between the subtypes circulating in patients with HIV and Clinical hepatitis (p= 0.0005), patients with clinical hepatitis and blood donors (p= 0.0005; p< .05), HIV patients and blood donors (p= 0.0005). Subtypes 1b and 2b were found only in patients with hepatitis, while subtypes 1a and 5a were predominant in blood donors and individuals with HIV respectively (p=0.0004). This means that specific subtypes of HCV were responsible for infection in the different subpopulations of the study. There were also significant differences in the distribution of each subtypes (1a, 1b, 2b, 2c, 5a)withinstudy populations (p=0.007, 0.003, 0.003, 0.003 and 0.041; p< .05), respectively.

Table 4.2 shows the distribution of subtypes within individual study subpopulations.

a. Out of 13 isolates obtained from 13 individuals with HIV, 10 were subtype 5a (76.7%), while 2 isolates were subtype 1a (15.4%) and 1subtype 3a (7.7%). This result shows that among patients with HIV, the predominant HCV strain circulating was subtype 5a, followed by 1a and least was subtype 3a. This means thateven though multiple HCV subtypes circulate amongpersons with HIV, subtype 5a was implicated more than other subtypes.
b. Among patients with clinical hepatitis, there were sixteen isolates, 2 of which were subtypes 1a (12.5%), 5 (31.5%) were 1b, 4 (25%) were 2b and 5 (31.5%) were 5a. This result also indicates multiple circulations of subtypes but with equal distribution of subtypes 1b and 5a (31.5% each),which were also the predominant strains among patients with clinical hepatitis, followed by subtype 2b (25%).Furthermore, subtypes 1b and 2b were found only among this group, meaning that there is an association between these subtypes and symptomatic infection with HCV, since these subtypes were not found among other subpopulations studied.
c. There were 13 isolates in blood donors, of which 8 were subtype 1a (61.5%) followed by 4 isolates of subtype 5a (30.8%) and 1 subtype 2c (7.7%). Blood donors had HCV subtype 1a (61.5%)as the predominant subtype circulating among them, followed by 5a (30.8%). This implies that among multiple subtypes circulating among blood

donors (asymptomatic infection), subtype 1a was more in circulation. Overall, the results obtained indicated diverse HCV strains, but with significant differences in the subtypes circulatingin all the study population.

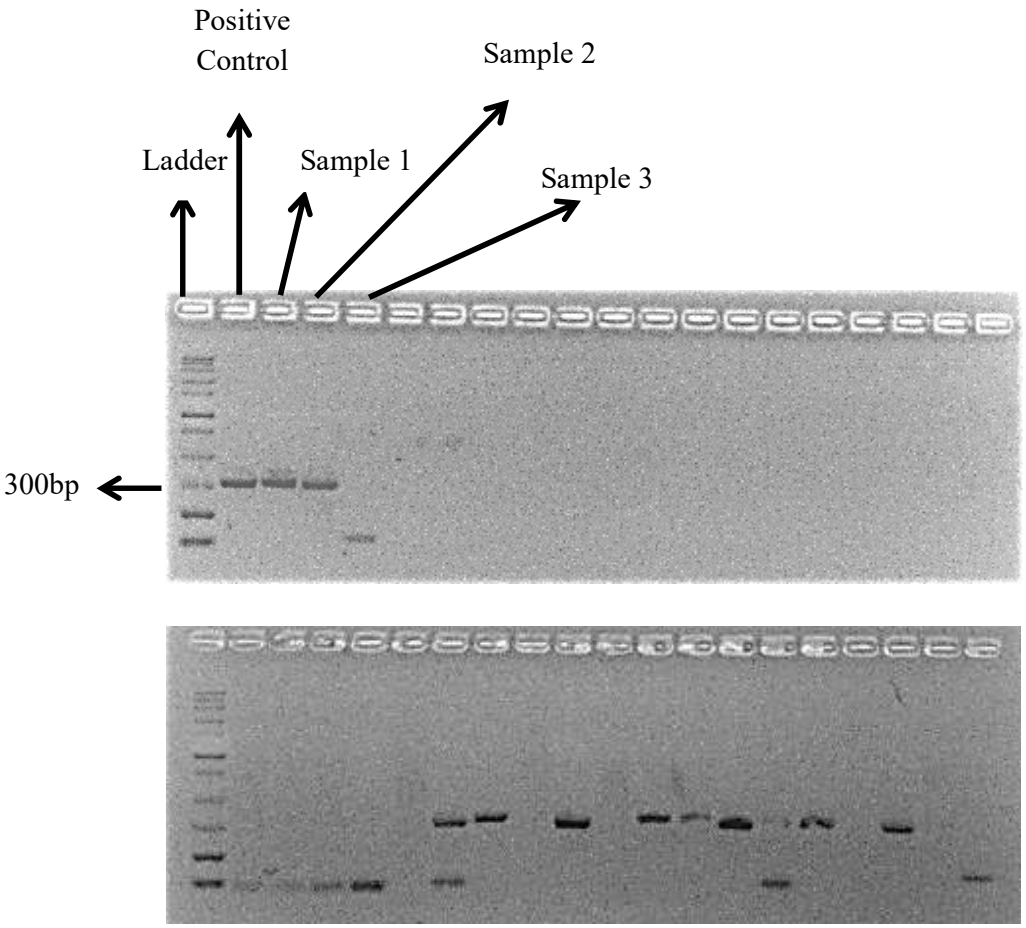

Figure 4.1: .Gel electrophoresis showing amplified Hepatitis CVirus NS5B gene with expected band sizes at 300bp

Table 4.1: Distribution of Hepatitis CVirus subtypes circulating in study populations

Subtypes	Patients with HIV No.(**%**)Occurrence	Patients with Clinical Hepatitis No. (**%**)Occurrence	Blood Donors No. (**%**) Occurrence	Total No. (%)	p< .05 **p= 0.0005**
1a	2 (16.7)	2 (16.7)	8 (66.7)	12 (28.6)	.007
1b	0 (0)	5 (100)	0 (0)	5(11.9)	.003
2b	0 (0)	4 (100)	0 (0)	4(9.5)	.003
2c	0 (0)	0 (0)	1 ((100)	1(2.4)	.003
3a	1(100)	0 (0)	0 (0)	1(2.4)	.003
5a	10 (52.6)	5 (26.3)	4 (21.1)	19(45.2)	.041
Total>	13 (31.0)	16 (38.0)	13(31.0)	**42 (100)**	

Table 4.2: Distribution of HCV subtypes within subpopulations

Study Subpopulation	Subtype 1a (%)	Subtype 1b (%)	Subtype 2b (%)	Subtype 2c (%)	Subtype 3a (%)	Subtype 5a (%)	Total Isolates
HIV patients	2(15.4)	0	0	0	1 (7.7%)	10(76.7%)	13
Patients with hepatitis	2(25)	5(31.5)	4 (25)	0	0	5 (31.5)	16
Blood donors	8(61.5)	0	0	1 (7.7)	0	4 (30.8)	13
Total	12(28.6)	5(11.9)	4(9.5)	1 (2.4)	1(2.4)	19 (45.2)	**42 (100)**

4.2.3 Hepatitis CVirus Sequence BLASTs for Identity

Basic Local Alignment Search Tool (BLAST) was used to query the evolutionary relatedness of the isolates. The consensus sequences blasted on NCBI were found to have 97%- 100% percentage identity with reference sequences obtained from GeneBank and HCV database. The different genotypes obtained are shown in table 4.3. BLAST results showed that the closest HCV strains to isolates from this study are from South Africa with 99% identity. On the other hand, previously identified Nigerian isolate closest to isolates in this study has accession number JQ679065 and belongs to genotype 1(not subtyped) with 89% identity. The result of the BLAST homology for some of the gene sequences are shown in tables 4.3 and 4. 4 for patients and blood donors respectively.

Phylogenetic Analyses

Figure 4. 2 show the phylogenetic tree of all the NS5B gene sequences obtained in this study. The major groups are genotypes 1 to genotype 7. Each group of genotype contains the subtypes clustering along the tree in order of relatedness. Subtypes 1a of genotype 1 from this study clustered mostly with subtype 1a sequences from Canada, Argentina and USA, and closely align with subtype 6a of genotype 6; whilesubtype 1b forms a different cluster with Brazilian strains but close to Nigerian strains from previous studies. Genotype 2 to which NGR-JS-BD30 (subtype 2c)belong has in that group, HCV strains from Guinea Bissau. It shares a common node with and closely related to a separate cluster where other genotype 2 strains and genotype (gt) 4 and 7 from Central African Republic and Democratic Republic of Congo respectively belong. All HCV genotype 5 (all subtype 5a) form a long cluster with South African strain as their ancestor. Subtype 3a closely aligned with a mutant 3a gene from Sri Lanka forming the last group.

Figure 4.3 is the phylogenetic tree showing alignment of all the isolates obtained from patients with clinical hepatitis aligned with sequences from other countries. All subtype 1b clustered together with strains from Spain and Brazil. The second group to which LD15 isolate belongs aligned with previously isolated Nigerian strains. Other genotype 1 sequences previously isolated from Nigeria were farther away on the tree. Subtype 1a formed a different cluster and close to those from Argentina and Canada. Subtype 2b clustered with genotypes 2 from Brazil and USA; and has 94% closeness to those from Amsterdam and Guinea Bissau. All subtype 5a formed a single cluster with strains from South Africa, Japan, Belgium and China.

Figure 4.4 shows the phylogenetic tree of the isolates from blood donors. There are five clusters. Strains originating from the same ancestor clustered together with short branches while those distantly related branched further away from the tree. Isolates that belong to subtype 1a clustered with those from Argentina, Brazil and UK. Isolate BD30 belonging to subtype 2c (genotype 2) clustered with those from Guinea Bissau and Argentina. NGR-JS-BD1 belongs to genotype 5 but has 90% identity to the rest in that group.

Figure 4.5 shows the phylogenetic tree of all the isolates from HIV–infected persons. There are two major groups. Group 1 has all the genotype 5 in one cluster, followed by genotypes 2 in the second cluster, 1b in third cluster, 1a in the fourth cluster with strains from Canada and Argentina. The second group consists of genotype 3 which forms a group with Sri Lanka isolate.

Figure 4.6 shows the phylogenetic tree showing the relatedness among all the isolates from this study differentiated by respective colours assigned to them.

Table 4.3: Basic Local Alignment Search Tool (BLAST) Homology results of some NS5B gene sequences in Patients (: NCBI GeneBank)

Isolates	Max Score	Total Score	Query	% Identity	Accession	Country
NGR-JS-P03	529	529	99%	98%	KR85599	Australia
NGR-JS-P06	534	534	99%	99%	KC767832	South Africa
NGR-JS- P10	503	503	99%	97%	KC 767832	South Africa
NGR-JS-LD5	484	484	95%	97%	EU155373	Switzerland
NGR-JS-LD15	488	488	91%	89%	JQ679065	Nigeria

Table 4.4: Basic Local Alignment Search Tool (BLAST) Homology results of some NS5B gene sequences in Blood Donors (Source: NCBI Genbank)

Isolates	Max Score	Total Score	Query	% Identity	Accession	Country
NGR-JS-BD2	448	448	96%	97%	JX463640	USA
NGR-JS-BD9	676	676	100%	100%	DQ333676	South Africa
NGR-JS-BD28	601	601	96%	99%	GU131382	USA

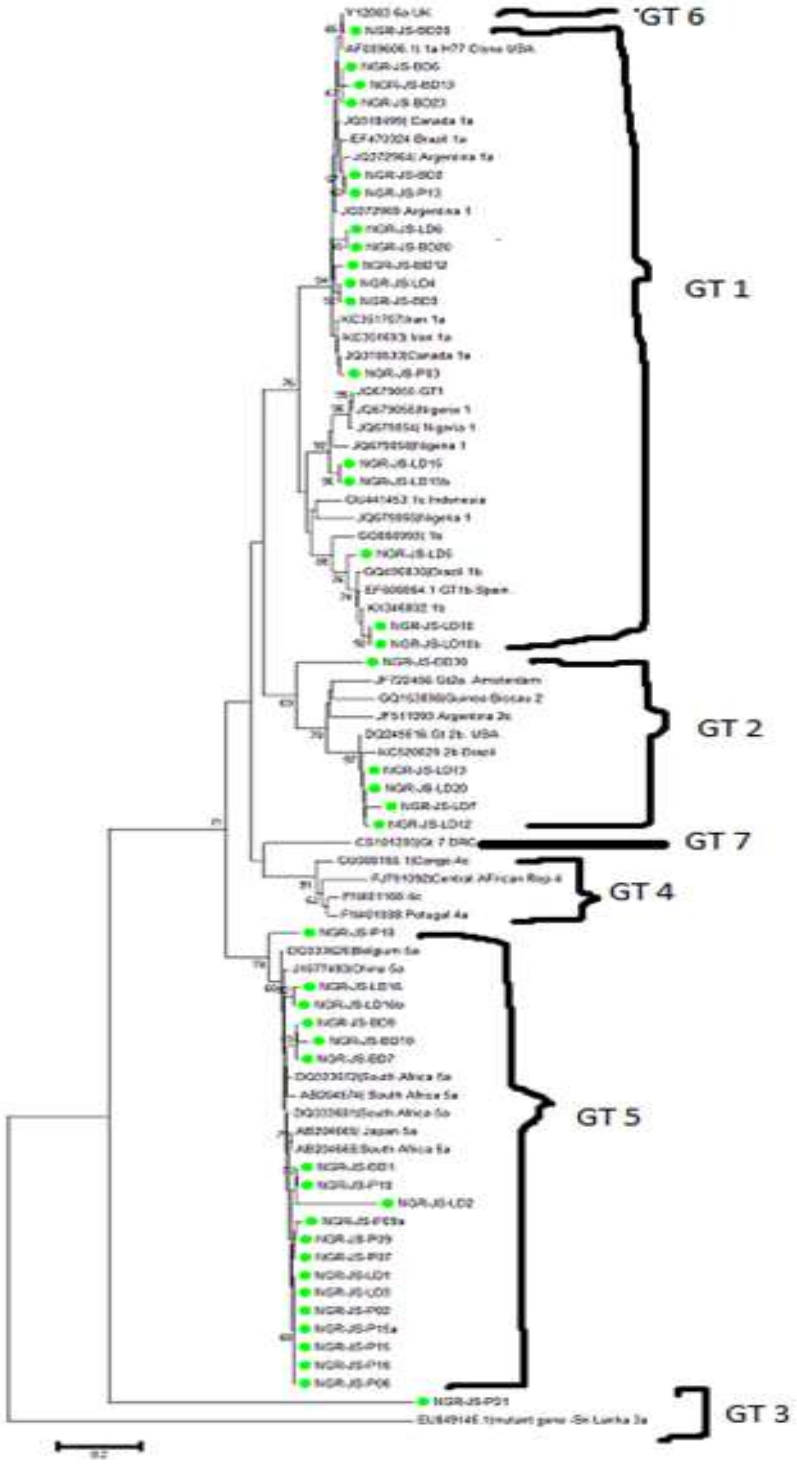

Fig 4. 2: Molecular Phylogeny of study HCV NS5B sequences (marked in green) along with reference sequences constructed on MEGA 7.0 using Neighbor-Joining method. Arrows indicate position of the different genotypes with reference sequences

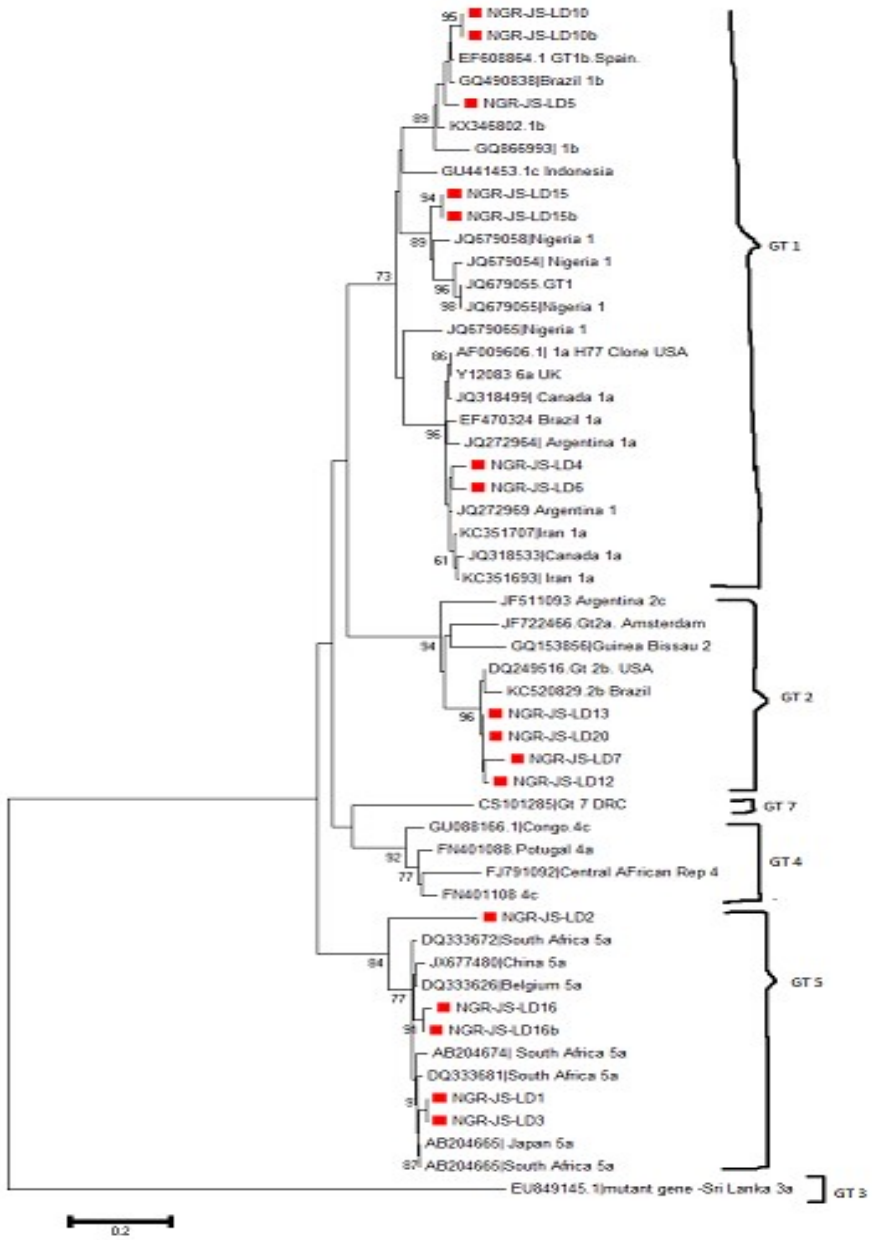

Fig. 4. 3: Phylogenetic tree of HCV NS5B sequences (1a, 1b, 2b 5a) among Patients with clinical hepatitis or liver disease (red blocks) along with reference sequences. Bar indicates the nucleotide substitution rate.

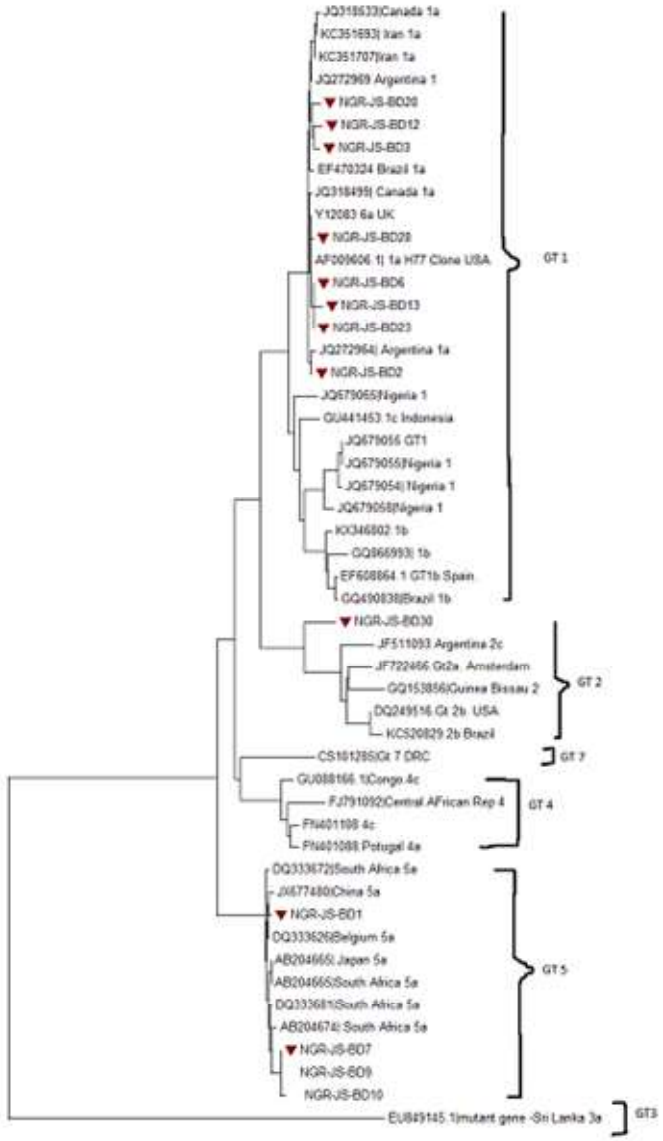

Fig.4.4: Molecular Phylogeny of HCV NS5B sequences (red arrows) among Blood Donors. Study HCV sequences are shown with arrows from the top of tree as genotypes /subtypes 1a, 2c, 5a

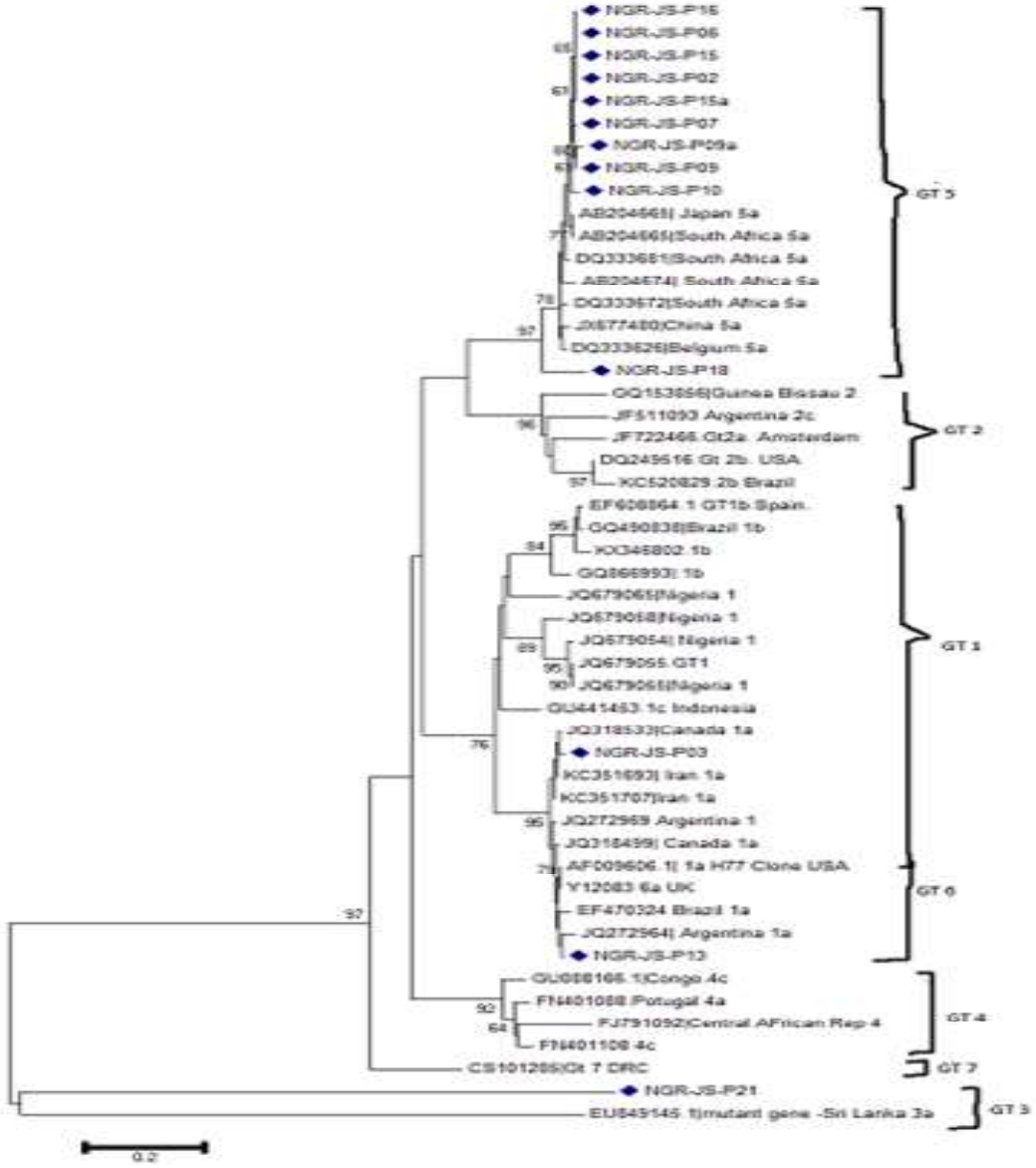

Fig 4.5: Molecular Phylogeny of HCV NS5B genes sequences among Individuals with HIV. Study sequences are shown in blue blocks as genotype /Subtypes 5a, 1a, 3a from the top of the tree. Tree was constructed on MEGA 7. 0

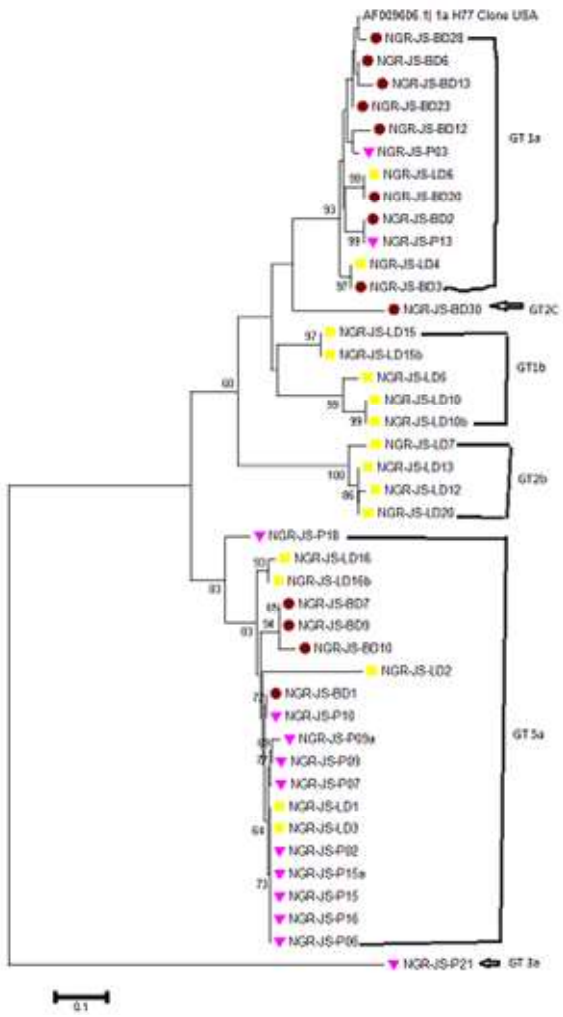

Fig. 4.6: Molecular Phylogeny of NS5B gene of study HCV sequences constructed using Neighbour- joining method in MEGA 7.0

Key: RED Blocks: Blood Donors; Purple arrows: Individuals with HIV; Yellow: Patients with clinical hepatitis

OBJECTIVE 2: **To identify mutations at the NS5B gene of HCV isolates.**

4.2.4 Alignment of Nucleotide Sequences

There were major substitutions in the nucleotide sequences, with some occurring as single nucleotide substitution and others as double or complete codon changes. Substitutions from Adenine to Guanine; Cytosine and Thymine and vice versa were observed in most variable sites than conserved positions where there were no changes in the bases compared to reference sequences. Nucleotide frequencies were calculated in MEGA 7.0 based on number of occurrence of eachbase AGCT with values obtained as 21.29%, 29.28%, 28.90%, 20.47%respectively. In all, Guanine has the lowest frequency while Thymine has the highest frequency among the sequences. The nucleotide frequency chart is shown in Figure 4.7, while the nucleotide sequences of all the study isolates are shown in figure 4.8.

4.2.5 Amino acid Translation and Protein Alignment

Table 4.5 shows the translated amino acid of all the NS5B sequences that constitute complete proteins for each sequence on MEGA 7.0. All the nucleotide sequences are translated to amino acids which were compared with the 77 prototype strain. A few amino acid positions have insertions or deletions while others have some stop codons (termination sites). The following position 5, 17, 25, 32, 40, 41, 43, 53, 63, 64, 69) have their amino acids conserved (.) when compared with HCV prototype H77. The complete protein sequences are shown in table 4.5.

Table 4.6 indicates each amino acid substitution site in the isolates. Complete Conserved sites across all the isolates with reference to H77 prototype are represented as red dots. The amino acids in ‘letters’ are the variable sites where there were substitutions in the viral protein. Isolates from blood donors have more conserved amino acids site, compared to H77 prototype, followed by patients with HIV while isolates in patients with clinical hepatitis have more variable amino acid positions or amino substitutions.

Table 4.7 shows amino acid alignment of HCV isolates and all other sequences from other countries in MEGA. In comparison to other references, HCV isolates from this study have similar amino acid positions with others. Specifically amino acid positions 1 to 10 (TNVVTWTRKP) in study isolates are same as in most reference sequences. Such similarities are more in isolates of blood donors and HIV patients, while isolates in patients with clinical hepatitis have lots of deletions.

4.2.6 Amino Acid Composition and Frequencies

Table 4.8 shows the frequencies of amino acid of the NS5B sequences as determined using MEGA 7.0, indicating the composition of each amino acid in each sequence and their percentage. In all, percentages of the different amino acids in the isolates were given in comparison with H77 amino acid frequencies. Most of the amino acids occurred frequently in the isolates while some such as Histidine which was absent in H77 prototype, remained absent in most of the study isolates. In addition, most showed substantial variation from the prototype strain H77. Amino acids Alanine, Glycine, Leucine, Proline, Serine, Threonine and Valine (14.32%, 6.15%, 8.15%, 10.03%, 9.82%, 9.82% and 9.45% respectively) had higher frequencies among the isolates. The composition indicates the number of times each amino acid occurs in any protein sequence of the isolates.

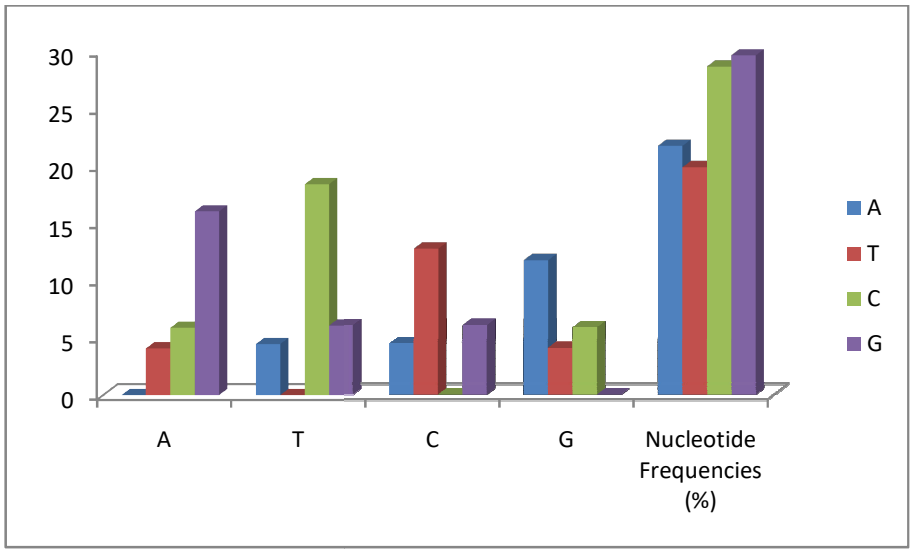

Fig 4.7: Maximum Likelihood Estimate of Base Substitution among the NS5B nucleotide sequences indicating frequency of occurrence of the bases (ATCG

>NGR-JS-BD1
TCTCAACCGTCACTGAGAGAGACATAATTACGGAAGAGACCATTTACCAATCAT
GTGACTTGCAGCCTGCAGGCGCGCGCGGCCATTACAGGTCACTCACCCAACGCCT
GTACTGTGGAGGCCCTATGTACAACAGCAAGGGGCAACAATGTGGTTATCGTAG
ATGCCGCGCCAGCGGCGTCTTCACCACTAGTATGGGCAACACCATAACGTGCTAC
GTTAAGGCTTTGGCCTCCTGTAGAGCTGCAAAGCTCCGGGACTGCACGCTCTTGG
TGTGTGGTGATGATCTTGTGGTCATCTGCGGAAGTCAGGAAACCCACGAGGACG
CGGCGAGCCTGAGAGCCTCAACCGTCACTGAGAGAGACATAATTACGGAAGAGA
CCATTTACCAATCATGTGACTTGCAGCCTGCAGGCGCGCGCGGCCATTACAGGTC
ACTCACCCAACGCCTGTACTGTGGAGGCCCTATGTACAACAGCAAGGGGCAACA
ATGTGGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACCACTAGTATGGGCAAC
ACCATAACGTGCTACGTTAAGGCTTTGGCCTCCTGTAGAGCTGCAAAGCTCCGGG
ACTGCACGCTCTTGGTGTGTGGTGATGATCTTGTGGTCATCTGCGGAAGTCAGGA
AACCCACGAGGACGCGGCGAGCCTGAGAGC

>NGR-JS-BD2
GGAGGAGGCGATCTACCAATGTTGTGACCTGGACCCCCAGGCCCGCGTGGCCAT
CAAGTCCCTCACTGAAAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGG
GAGAACTGCGGCTACCGCAGGTGCCGCGCAAGCGGCGTACTGACGACTAGCTGT
GGTAACACCCTCACTTGCTATATCAAGGCCAAGGCAGCCTGTCGAGCCGCAGGG
CTCCAGGACTGCACCATGCTCGTGTGTGGCGACGACCTAGTCGTTAATTTATGGA
TGGAGGAGGCGATCTACCAATGTTGTGACCTGGACCCCCAGGCCCGCGTGGCCA
TCAAGTCCCTCACTGAAAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGG
GGAGAACTGCGGCTACCGCAGGTGCCGCGCAAGCGGCGTACTGACGACTAGCTG
TGGTAACACCCTCACTTGCTATATCAAGGCCAAGGCAGCCTGTCGAGCCGCAGG
GCTCCAGGACTGCACCATGCTCGTGTGTGGCGACGACCTAGTCGTTAATTTATGG
AT

>NGR-JS-BD3
CTCTCAACCGTCACTGAGAGAGACATCCGTACGGAGGAGGCAATCTACCAATGT
TGTGACCTGGACCCCCAAGCCCGTGTGGCCATCAGGTCTCTCACCGAGAGGCTTT
ATGTCGGGGGCCCTCTTACCAACTCAAGGGGGGAAAACTGCGGCTATCGCAGGT
GCCGCGCAAGCGGCGTGCTGACAACTAGCTGTGGTAACACCCTCACTTGCTACAT
CAAGGCCCGAGCAGCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGT
GTGTGGCGACGACTTGGTCGTTATCTGTGAAAGTCAGGGGGTCCAGGAGGACGC
GGCGAGCCTGAGAGGTCTCAACCGTCACTGAGAGAGACATCCGTACGGAGGAGG
CAATCTACCAATGTTGTGACCTGGACCCCCAAGCCCGTGTGGCCATCAGGTCTCT
CACCGAGAGGCTTTATGTCGGGGGCCCTCTTACCAACTCAAGGGGGGAAAACTG
CGGCTATCGCAGGTGCCGCGCAAGCGGCGTGCTGACAACTAGCTGTGGTAACAC
CCTCACTTGCTACATCAAGGCCCGAGCAGCCTGTCGAGCCGCAGGGCTCCAGGA
CTGCACCATGCTCGTGTGTGGCGACGACTTGGTCGTTATCTGTGAAAGTCAGGGG
GTCCAGGAGGACGCGGCGAGCCTGAGAG
>NGR-JS-BD6
TCTCAACCGTCACTGAGAGAGACATCCGTACGGAGGAAGCAATCTACCAATAGT
TGTGACCTGGACCCCCAAGCCCGCGTGGCCATCAAGTCTCTCACTGAGAGGCTTT
ATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAAAACTGCGGCTACCGCAGGT
GCCGCGCGAGCGGCGTACTGACTACTAGCTGTGGTAACACCCTCACTTGCTACAT
CAAGGCCCGAGCAGCCTGTCGAGCCGCAGGGCTCCGGGACTGCACCATGCTCGT
GTGTGGCGACGACCTAGTCGCTTATCTGCGAAAGTGCGGGGGTCCAGGAGGACG

CGGCGAGCCTGAGATCTCAACCGTCACTGAGAGAGACATCCGTACGGAGGAAGC
AATCTACCAATAGTTGTGACCTGGACCCCCAAGCCCGCGTGGCCATCAAGTCTCT
CACTGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAAAACTG
CGGCTACCGCAGGTGCCGCGCGAGCGGCGTACTGACTACTAGCTGTGGTAACAC
CCTCACTTGCTACATCAAGGCCCGAGCAGCCTGTCGAGCCGCAGGGCTCCGGGA
CTGCACCATGCTCGTGTGTGGCGACGACCTAGTCGCTTATCTGCGAAAGTGCGGG
GGTCCAGGAGGACGCGGCGAGCCTGAGA

>NGR-JS-BD7
TTCTCAACCGTCACTGAGAGAGACATTATGACTGAAGAGTCCATCTACCAATACA
TGTGACTTACAACCTGAGGCGCGCACGGCAATACGGTCACTCACCCAACGCCTGT
ACTGTGGAGGCCCTATGTATAACAGCAAGGGGCAGCAATGTGGTTATCGTAGAT
GCCGCGCCAGCGGCGTCTTCACTACCAGTATGGGCAACACCATGACGTGTTACAT
TAAGGCTTTAGCCTCCTGTAGAGCTGCAAAGCTCCGGGATTGTACGCTCCTGGTG
TGTGGTGACGACCTTGTGGCCATCTGCGAGAGCCAAGGGACACACGAGGACGCG
GCGAGCCTGAGAGCATTCTCAACCGTCACTGAGAGAGACATTATGACTGAAGAG
TCCATCTACCAATACATGTGACTTACAACCTGAGGCGCGCACGGCAATACGGTCA
CTCACCCAACGCCTGTACTGTGGAGGCCCTATGTATAACAGCAAGGGGCAGCAA
TGTGGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACTACCAGTATGGGCAACA
CCATGACGTGTTACATTAAGGCTTTAGCCTCCTGTAGAGCTGCAAAGCTCCGGGA
TTGTACGCTCCTGGTGTGTGGTGACGACCTTGTGGCCATCTGCGAGAGCCAAGGG
ACACACGAGGACGCGGCGAGCCTGAGAGCA

>NGR-JS-BD9
TGGGGATCCCGTATGATACCCGCTGCTTTGACTCAACCGTTACCGAACATGACAT
TATGACTGAAGAGTCCATCTACCAATCATGTGACTTACAACCTGAGGCGCGCACG
GCAATACGGTCACTCACCCAACGCCTGTACTGTGGAGGCCCTATGTATAACAGCA
AGGGGCAGCAATGTGGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACTACCA
GTATGGGCAACACCATGACGTGTTACATTAAGGCTTTAGCCTCCTGTAGAGCTGC
AAAGCTCCGGGATTGTACGCTCCTGGTGTGTGGTGACGACCTTGTGGCCATCTGC
GAGAGCCAAGGGACACACGAGGATGAAGCGAGCCTGAGAGCCTTCACGGAGGC
TATGACCAGGAATTCCGCCATGGGGATCCCGTATGATACCCGCTGCTTTGACTCA
ACCGTTACCGAACATGACATTATGACTGAAGAGTCCATCTACCAATCATGTGACT
TACAACCTGAGGCGCGCACGGCAATACGGTCACTCACCCAACGCCTGTACTGTG
GAGGCCCTATGTATAACAGCAAGGGGCAGCAATGTGGTTATCGTAGATGCCGCG
CCAGCGGCGTCTTCACTACCAGTATGGGCAACACCATGACGTGTTACATTAAGGC
TTTAGCCTCCTGTAGAGCTGCAAAGCTCCGGGATTGTACGCTCCTGGTGTGTGGT
GACGACCTTGTGGCCATCTGCGAGAGCCAAGGGACACACGAGGATGAAGCGAGC
CTGAGAGCCTTCACGGAGGCTATGACCAGGAATTCCGCCA

>NGR-JS-BD10
TTCTCAACCGTCACTGAGAGAGACATTATACCTGAAGAGTCCATCTACCAATCAT
GTGACTTACAACCTGAGGCGCGCACGGCAATGCGGTCACTCACCCAACGCCTGT
ACTGTGGAGGCCCTATGTATAACAGCAAGGGGCAGCAATGTGGTTATCGTAGAT
GCCGCGCCAGCGGCGTCTTCACTACCAGTATGGGCAACACCATGACGTGTTACAT
TAAGGCTTTAGCCTCCTGTAGAGCTGCGAAGCTCCGGGATTGCACGCTCCTGGTG
TGTGGTGACGACCTTGTGGCCATCTGCGAGAGCCAAGGGACACACGAGGACGCG
GCGAGCCTGAGAGCATTCTCAACCGTCACTGAGAGAGACATTATACCTGAAGAG
TCCATCTACCAATCATGTGACTTACAACCTGAGGCGCGCACGGCAATGCGGTCAC
TCACCCAACGCCTGTACTGTGGAGGCCCTATGTATAACAGCAAGGGGCAGCAAT

GTGGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACTACCAGTATGGGCAACAC
CATGACGTGTTACATTAAGGCTTTAGCCTCCTGTAGAGCTGCGAAGCTCCGGGAT
TGCACGCTCCTGGTGTGTGGTGACGACCTTGTGGCCATCTGCGAGAGCCAAGGG
ACACACGAGGACGCGGCGAGCCTGAGAGCA

>NGR-JS-BD12
TCTCAACCGTCACTGAGAGAGACACCGGTACGGAGGAGGCAATCTACCAGTGTT
GTGACCTGGACCCCCAAGCCCGTGTGGCCATCAAGTCTCTCACCGAGAGGCTTTA
TGTCGGGGGCCCTCTTACCAATTCAAGGGGGGAAAACTGTGGCTATCGCAGGTG
CCGCGCGAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTCACTTGCTACATC
AAGGCCCAGGCAGCCTGTCGAGCCGCAGGGCTCCGGGACTGCACCATGCTCGTG
TGTGGCGACGACCTAGTCGTTATCAGCGAAAGTCAGGGAGTCCAGGAGGACGCG
GCGAGCCTGAGAGCATCTCAACCGTCACTGAGAGAGACACCGGTACGGAGGAGG
CAATCTACCAGTGTTGTGACCTGGACCCCCAAGCCCGTGTGGCCATCAAGTCTCT
CACCGAGAGGCTTTATGTCGGGGGCCCTCTTACCAATTCAAGGGGGGAAAACTG
TGGCTATCGCAGGTGCCGCGCGAGCGGCGTACTGACAACTAGCTGTGGTAACAC
CCTCACTTGCTACATCAAGGCCCAGGCAGCCTGTCGAGCCGCAGGGCTCCGGGA
CTGCACCATGCTCGTGTGTGGCGACGACCTAGTCGTTATCAGCGAAAGTCAGGG
AGTCCAGGAGGACGCGGCGAGCCTGAGAGCA

>NGR-JS-BD13
TCTCAACCGTCACTGAGAGAGACATCCGTACGGAGGAGGCAATCTACCAATGTT
GTGACCTGGACCCCCAAGCCCGCGTGGCCATCAAGTCTCTCACTGAGAGGCTTTA
TGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAAAACTGCGGCTACCGCAGGTG
CTGCGCGAGCGGCGTACTGACTACTAGCTGTGGTAACACCCTCACTTGCTACATC
AAGGCCCGAGCAGCCTGTCGAGCCGCAGGGCTCCGGGACTGCACCATGCTCGTG
TGTGGCGACGACCTAGTCGTTATCTGCGAAAGTGCGGGGGTCCAGGAGGACGCG
GCGAGCCTGAGAGCATCTCAACCGTCACTGAGAGAGACATCCGTACGGAGGAGG
CAATCTACCAATGTTGTGACCTGGACCCCCAAGCCCGCGTGGCCATCAAGTCTCT
CACTGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAAAACTG
CGGCTACCGCAGGTGCTGCGCGAGCGGCGTACTGACTACTAGCTGTGGTAACAC
CCTCACTTGCTACATCAAGGCCCGAGCAGCCTGTCGAGCCGCAGGGCTCCGGGA
CTGCACCATGCTCGTGTGTGGCGACGACCTAGTCGTTATCTGCGAAAGTGCGGGG
GTCCAGGAGGACGCGGCGAGCCTGAGAGCA

>NGR-JS-BD20
TCTCAACCGTTCACTGAGAGAGACATCCGTACGGAGGAGGCAATCTACCAATGT
TGTGACCTTGACCCCCAAGCCCGTGTGGCCATCAAGTCCCTCACCGAGAGACTTT
ACGTTGGGGGCCCTCTTACCAACTCAAGGGGGGAGAACTGCGGCTACCGCAGGT
GCCGCGCGAGCGGCGTACTGACAACCAGCTGTGGTAACACCCTCACTTGCTACAT
CAAGGCCCAGGCAGCCTGTCGAGCCGCAGGGCTCCGGGACTGCACCATGCTCGT
GTGTGGCGACGACTTGGTCGTTATCTGTGAAAGTGCGGGGATCCAGGAGGACGC
GGCGAGCCTGAGAGCATCTCAACCGTTCACTGAGAGAGACATCCGTACGGAGGA
GGCAATCTACCAATGTTGTGACCTTGACCCCCAAGCCCGTGTGGCCATCAAGTCC
CTCACCGAGAGACTTTACGTTGGGGGCCCTCTTACCAACTCAAGGGGGGAGAAC
TGCGGCTACCGCAGGTGCCGCGCGAGCGGCGTACTGACAACCAGCTGTGGTAAC
ACCCTCACTTGCTACATCAAGGCCCAGGCAGCCTGTCGAGCCGCAGGGCTCCGG
GACTGCACCATGCTCGTGTGTGGCGACGACTTGGTCGTTATCTGTGAAAGTGCGG
GGATCCAGGAGGACGCGGCGAGCCTGAGAGCA

>NGR-JS-BD23
TTCTCAACCTTTCACTGAGAGAGCCATCCGTACGGAGGAGGCAATTTACCAATGT
TGTGACCTGGACCCCCAAGCCCGCGTGGCCATCAAGTCTCTCACTGAGAGGCTTT
ATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAAAACTGCGGCTATCGCAGGT
GCCGCGCGAGCGGCGTACTGACTACTAGCTGTGGTAACACCCTCACTTGCTACAT
CAAGGCCCGAGCAGCCTGTCGAGCCGCAGGGCTCCGGGACTGCACCATGCTCGT
GTGTGGCGACGACCTAGTCCTTATCTGCGAAAGTGCGGGGTCCAGGAGGACGCG
GCGGACCTGAAAGCAATTCTCAACCTTTCACTGAGAGAGCCATCCGTACGGAGG
AGGCAATTTACCAATGTTGTGACCTGGACCCCCAAGCCCGCGTGGCCATCAAGTC
TCTCACTGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAAAAC
TGCGGCTATCGCAGGTGCCGCGCGAGCGGCGTACTGACTACTAGCTGTGGTAAC
ACCCTCACTTGCTACATCAAGGCCCGAGCAGCCTGTCGAGCCGCAGGGCTCCGG
GACTGCACCATGCTCGTGTGTGGCGACGACCTAGTCCTTATCTGCGAAAGTGCGG
GGTCCAGGAGGACGCGGCGGACCTGAAAGCAA

>NGR-JS-BD28
TCTGAACCGTTCACTGAGAGAGACATCCGTACGGAGGAGGCAATCTACCAATGT
TGTGACCTGGACCCCCAAGCCCGCGTGGCCATCAAGTCCCTCACTGAGAGGCTTT
ATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGCTACCGCAGGT
GCCGCGCGAGCGGCGTACTGACTACTAGCTGTGGTAACACCCTCACTTGCTACAT
CAAGGCCCGGGCAGCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGT
GTGTGGCGACGACTTAGTCGTTATCTGTGAAAGTGCGGGGGTCCAGGAGGACGC
GGCGAGCCTGAGAGCATCTGAACCGTTCACTGAGAGAGACATCCGTACGGAGGA
GGCAATCTACCAATGTTGTGACCTGGACCCCCAAGCCCGCGTGGCCATCAAGTCC
CTCACTGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACT
GCGGCTACCGCAGGTGCCGCGCGAGCGGCGTACTGACTACTAGCTGTGGTAACA
CCCTCACTTGCTACATCAAGGCCCGGGCAGCCTGTCGAGCCGCAGGGCTCCAGG
ACTGCACCATGCTCGTGTGTGGCGACGACTTAGTCGTTATCTGTGAAAGTGCGGG
GGTCCAGGAGGACGCGGCGAGCCTGAGAGCA

>NGR-JS-BD30
TCTCAACCGTCACTGAGAGAGACATACGTACCGAGCAGCCTATCTATCGATTGTG
CACTGTCTCGTGCCAAGCTCGCGATGGCCATCAAGTCCCTGACTGAGAGACTATT
ATGTGGGAGGCCCTATGACCAACACAAAGGGGGAATACTGCGGCTACAGCAGGT
GCCGCGCGAGCGGAGTGCTCACTACTAGCATGGGTAACACCCTCACGTGCTACG
TCAAGGCCCAGGCAGCCTGTAGAGCCGCGGGCATCCAGGCCCCCACCATGCTGG
TATGTGGCGACGACCTAGTCGTCATCTCAAAGAGTGCGGGGCTCCAGGAGGACG
CGGCGAGCCCTGAAAGCCTCTCAACCGTCACTGAGAGAGACATACGTACCGAGC
AGCCTATCTATCGATTGTGCACTGTCTCGTGCCAAGCTCGCGATGGCCATCAAGT
CCCTGACTGAGAGACTATTATGTGGGAGGCCCTATGACCAACACAAAGGGGGAA
TACTGCGGCTACAGCAGGTGCCGCGCGAGCGGAGTGCTCACTACTAGCATGGGT
AACACCCTCACGTGCTACGTCAAGGCCCAGGCAGCCTGTAGAGCCGCGGGCATC
CAGGCCCCCACCATGCTGGTATGTGGCGACGACCTAGTCGTCATCTCAAAGAGTG
CGGGGCTCCAGGAGGACGCGGCGAGCCCTGAAAGCC

>NGR-JS-LD1
TCTCAACCGTCACTGAGAGAGACATAATGACCGAAGAGTCCATTTACCAATCAT
GTGACTTGCAGCCTGAGGCGCGCGCGGCAATACGGTCACTCACCCAACGCCTGT
ACTGTGGAGGCCCTATGTATAACAGCAAAGGGCAACAATGTGGTTATCGTAGAT
GCCGCGCCAGCGGCGTCTTCACCACTAGTATGGGCAACACCATAACGTGCTACGT
TAAGGCTTTGGCCTCCTGTAGAGCTGCAAAGCTCCGGGACTGCACGCTCTTGGTG
TGTGGTGATGATCTTGTGGCCATCTGCGAAAGCCAGGGGACACACGAGGACGCG
GCGAGCCTGAGAGCTCTCAACCGTCACTGAGAGAGACATAATGACCGAAGAGTC
CATTTACCAATCATGTGACTTGCAGCCTGAGGCGCGCGCGGCAATACGGTCACTC
ACCCAACGCCTGTACTGTGGAGGCCCTATGTATAACAGCAAAGGGCAACAATGT
GGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACCACTAGTATGGGCAACACCA
TAACGTGCTACGTTAAGGCTTTGGCCTCCTGTAGAGCTGCAAAGCTCCGGGACTG
CACGCTCTTGGTGTGTGGTGATGATCTTGTGGCCATCTGCGAAAGCCAGGGGACA
CACGAGGACGCGGCGAGCCTGAGAGC

>NGR-JS-LD2
TCTCAACCGTCACTGAGAGAGACATCCGTACCGACGAGTCTATGTACCAATCATG
TGACCTGGACCCGCAGGCGCGCGCGGCCATTACAAGTCACCTCACCCAAAAGGC
TGTATGGTGGAGGCCCTCTTACCAATTCAAGGGGGGAGAAATGGGGCTTCCGCA
GATGCCGCGCCAGCGGCGTCTTCACCGACTAGCTGGGGTAACACCCTCACGTGCT
ACTTCAAGGCCTTGGCATCCTGTAGGCTGCAAAGCTCCAGGACTGCACCCTGTTG
GTGTGTGGTGACGATCTTGTGGCCATCTGCGAAAGTGAGGAAACACACGAGGAC
GCGGCGAGCCTGAGAGCTCTCAACCGTCACTGAGAGAGACATCCGTACCGACGA
GTCTATGTACCAATCATGTGACCTGGACCCGCAGGCGCGCGCGGCCATTACAAGT
CACCTCACCCAAAAGGCTGTATGGTGGAGGCCCTCTTACCAATTCAAGGGGGGA
GAAATGGGGCTTCCGCAGATGCCGCGCCAGCGGCGTCTTCACCGACTAGCTGGG
GTAACACCCTCACGTGCTACTTCAAGGCCTTGGCATCCTGTAGGCTGCAAAGCTC
CAGGACTGCACCCTGTTGGTGTGTGGTGACGATCTTGTGGCCATCTGCGAAAGTG
AGGAAACACACGAGGACGCGGCGAGCCTGAGAGC

>NGR-JS-LD3
TCTCAACCGTCACTGAGAGAGACATAATGACCGAAGGGTCCATTTACCAATCAT
GTGACTTGCAGCC
TGAGGCGCGCGCGGCAATACGGTCACTCACCCAACGCCTGTACTGTGGAGGCCC
TATGTATAACAGCAAAGGGCAACAATGTGGTTATCGTAGATGCCGCGCCAGCGG
CGTCTTCACCACTAGTATGGGCAACACCATAACGTGCTACGTTAAGGCTTTGGCC
TCCTGTAGAGCTGCAAAGCTCCGGGACTGCACGCTCTTGGTGTGTGGTGATGATC
TTGTGGCCATCTGCGAGAGCCAGGGGACACACGAGGACGCGGCGAGCCTGAGAG
CTCTCAACCGTCACTGAGAGAGACATAATGACCGAAGGGTCCATTTACCAATCAT
GTGACTTGCAGCCTGAGGCGCGCGCGGCAATACGGTCACTCACCCAACGCCTGT
ACTGTGGAGGCCCTATGTATAACAGCAAAGGGCAACAATGTGGTTATCGTAGAT
GCCGCGCCAGCGGCGTCTTCACCACTAGTATGGGCAACACCATAACGTGCTACGT
TAAGGCTTTGGCCTCCTGTAGAGCTGCAAAGCTCCGGGACTGCACGCTCTTGGTG
TGTGGTGATGATCTTGTGGCCATCTGCGAGAGCCAGGGGACACACGAGGACGCG
GCGAGCCTGAGAGC

>NGR-JS-LD4
TCTCAACCGTCACTGAGAGAGACATCCGTACGGAGGAGGCAATCTACCAATGTT
GTGACCTGGACCCCCAAGCCCGTGTGGCCATCAGGTCTCTCACCGAGAGGCTTTA
TGTCGGGGGCCCTCTTACCAACTCAAGGGGGGAAAACTGCGGCTATCGCAGGTG

CCGCGCAAGCGGCGTGCTGACAACTAGCTGTGGTAACACCCTCACTTGCTACATC
AAGGCCCGAGCAGCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGTG
TGTGGCGACGACTTGGTCGTTATCTGTGAAAGTCAGGGGGTCCAGGAGGACGCG
GCGAGCCTGAGAGGCTCTCAACCGTCACTGAGAGAGACATCCGTACGGAGGAGG
CAATCTACCAATGTTGTGACCTGGACCCCCAAGCCCGTGTGGCCATCAGGTCTCT
CACCGAGAGGCTTTATGTCGGGGGCCCTCTTACCAACTCAAGGGGGGAAAACTG
CGGCTATCGCAGGTGCCGCGCAAGCGGCGTGCTGACAACTAGCTGTGGTAACAC
CCTCACTTGCTACATCAAGGCCCGAGCAGCCTGTCGAGCCGCAGGGCTCCAGGA
CTGCACCATGCTCGTGTGTGGCGACGACTTGGTCGTTATCTGTGAAAGTCAGGGG
GTCCAGGAGGACGCGGCGAGCCTGAGAGGC

>NGR-JS-LD5
CTCAACCGTCACTGAGAGAGACATCCGTGTTGAGGAGTCAATTTACCAATGTTGT
GACTTAGCCCCCGAAGCCAGACAGGCCATACGGTCGCTCACAGAGCGGCTTTAT
ATCGGGGGCCCCCTGACTAATTCAAAAGGGCAGAACTGCGGTTATCGCCGGTGC
CGCGCAAGCGGCGTGCTGACGACTAGCTGTGGCAATACCCTCACATGCTACTTGA
AGGCCGCTGCGGCCTGTCGAGCTGCAAAGCTCCAGGACTGCACGATGCTCGTGA
ACGGAGACGACCTTGTCGTTATCTGTGAGAGCGCGGGAACCCAAGAGGACGCGG
CAAGGCTGAGAGCACTCAACCGTCACTGAGAGAGACATCCGTGTTGAGGAGTCA
ATTTACCAATGTTGTGACTTAGCCCCCGAAGCCAGACAGGCCATACGGTCGCTCA
CAGAGCGGCTTTATATCGGGGGCCCCCTGACTAATTCAAAAGGGCAGAACTGCG
GTTATCGCCGGTGCCGCGCAAGCGGCGTGCTGACGACTAGCTGTGGCAATACCCT
CACATGCTACTTGAAGGCCGCTGCGGCCTGTCGAGCTGCAAAGCTCCAGGACTG
CACGATGCTCGTGAACGGAGACGACCTTGTCGTTATCTGTGAGAGCGCGGGAAC
CCAAGAGGACGCGGCAAGGCTGAGAGCA

>NGR-JS-LD6
TCTCAACCGTTCACTGAGAGAGACATCCGTACGGAGGAGGCAATCTACCAATGT
TGTGACCTTGACCCCCAAGCCCGTGTGGCCATCAAGTCCCTCACCGAGAGACTTT
ACGTTGGGGGCCCTCTTACCAACTCAAGGGGGGAGAACTGCGGCTACCGCAGGT
GCCGCGCGAGCGGCGTACTGACAACCAGCTGTGGTAACACCCTCACTTGCTACAT
CAAGGCCCAGGCAGCCTGTCGAGCCGCAGGGCTCCGGGACTGCACCATGCTCGT
GTGTGGCGACGACTTGGTCGTTATCTGTGAAAGTGCGGGGATCCAGGAGGACGC
GGCGAGCCTGAGAGCACTCTCAACCGTTCACTGAGAGAGACATCCGTACGGAGG
AGGCAATCTACCAATGTTGTGACCTTGACCCCCAAGCCCGTGTGGCCATCAAGTC
CCTCACCGAGAGACTTTACGTTGGGGGCCCTCTTACCAACTCAAGGGGGGAGAA
CTGCGGCTACCGCAGGTGCCGCGCGAGCGGCGTACTGACAACCAGCTGTGGTAA
CACCCTCACTTGCTACATCAAGGCCCAGGCAGCCTGTCGAGCCGCAGGGCTCCG
GGACTGCACCATGCTCGTGTGTGGCGACGACTTGGTCGTTATCTGTGAAAGTGCG
GGGATCCAGGAGGACGCGGCGAGCCTGAGAGCAC

>NGR-JS-LD7
CCTTCTCAACCGTCACTGAGAGAGACATAAGAACAGAAGGGCAAGAATGCCAAG
GCTAGAACCTGCCCCAAGAGGCCAGAACTGCCATACACTCGCTCACTGAGAGAC
TTTACGTAGGAGGGCCCATGACAAACAGCAAAGGGCAATCCTGCGGTTACAGGC
GTTGCCGCGCAAGCGGCGTCTTCACCACCAGCATGGGGAATACCCTGACATGTTA
CATCAAAGCCCTTGGAGCGTACAAAGCTGCGGGGCCTGCGGACCCTGTTATGCT
GGTGTGTGGCGGAATCCTGGTCATCGCCTCCGCAAGCCAAGGCAGCAGGGAGGA
CGCGGCGAGCATAAGGGCATTCCTTCTCAACCGTCACTGAGAGAGACATAAGAA
CAGAAGGGCAAGAATGCCAAGGCTAGAACCTGCCCCAAGAGGCCAGAACTGCC

ATACACTCGCTCACTGAGAGACTTTACGTAGGAGGGCCCATGACAAACAGCAAA
GGGCAATCCTGCGGTTACAGGCGTTGCCGCGCAAGCGGCGTCTTCACCACCAGC
ATGGGGAATACCCTGACATGTTACATCAAAGCCCTTGGAGCGTACAAAGCTGCG
GGGCCTGCGGACCCTGTTATGCTGGTGTGTGGCGGAATCCTGGTCATCGCCTCCG
CAAGCCAAGGCAGCAGGGAGGACGCGGCGAGCATAAGGGCATT

>NGR-JS-LD10
CTCTCAACCGTCACTGAGAGAGACATCCGCATTGAGGAGTCAATCTACCAATGTT
GCGACTTGGCCCCCGAGGCCAGACAGGCTATAAGGTCGCTCACAGAGCGGCTTT
ATATCGGGGGTCCCCTGACTAATTCAAAAGGGCAGAACTGCGGTTATCGCCGGT
GCCGCGCAAGCGGCGTGCTGACGACTAGCTGCGGTAATACCCTCACATGTTACTT
GAAAGCCACTGCGGCCTGTCGAGCTGCAAAGCTCCAGGACTGCACAATGCTCGT
GAACGGGGACGACCTTGTCGTCATATGTGAGAGCGCGGGAACCCAAGAGGACGC
GGCGAGCCCTGAGACGGCTCTCAACCGTCACTGAGAGAGACATCCGCATTGAGG
AGTCAATCTACCAATGTTGCGACTTGGCCCCCGAGGCCAGACAGGCTATAAGGT
CGCTCACAGAGCGGCTTTATATCGGGGGTCCCCTGACTAATTCAAAAGGGCAGA
ACTGCGGTTATCGCCGGTGCCGCGCAAGCGGCGTGCTGACGACTAGCTGCGGTA
ATACCCTCACATGTTACTTGAAAGCCACTGCGGCCTGTCGAGCTGCAAAGCTCCA
GGACTGCACAATGCTCGTGAACGGGGACGACCTTGTCGTCATATGTGAGAGCGC
GGGAACCCAAGAGGACGCGGCGAGCCCTGAGACGG

>NGR-JS-LD10b
TCTCAACCGTCACTGAGAGAGACATAATTACGGAAGAGACCATTTACCAATCAT
GTGACTTGCAGCCTGCAGGCGCGCGCGGCCATTACAGGTCACTCACCCAACGCCT
GTACTGTGGAGGCCCTATGTACAACAGCAAGGGGCAACAATGTGGTTATCGTAG
ATGCCGCGCCAGCGGCGTCTTCACCACTAGTATGGGCAACACCATAACGTGCTAC
GTTAAGGCTTTGGCCTCCTGTAGAGCTGCAAAGCTCCGGGACTGCACGCTCTTGG
TGTGTGGTGATGATCTTGTGGTCATCTGCGGAAGTCAGGAAACCCACGAGGACG
CGGCGAGCCTGAGAGCTCTCAACCGTCACTGAGAGAGACATCCGCATTGAGGTG
TCAATCTACCAATAGTTGCGACTTGGCCCCCGAGGCCAGACAGGCTATAAGGTC
GCTCACAGAGCGGCTTTATATCGGGGGTCCCCTGACTAATTCAAAAGGGCAGAA
CTGCGGTTATCGCCGGTGCCGCGCAAGCGGCGTGCTGACGACTAGCTGCGGTAA
TACCCTCACATGTTACTTGAAAGCCACTGCGGCCTGTCGAGCTGCAAAGCTCCAG
GACTGCACAATGCTCGTGAACGGGGACGACCTTGTCGTCATCCTTGAGAGCGCG
GGAACCCAAGAGGACGCGGCGAGCCTGAGCGCCTCTCAACCGTCACTGAGAGAG
ACATAATTACGGAAGAGACCATTTACCAATCATGTGACTTGCAGCCTGCAGGCG
CGCGCGGCCATTACAGGTCACTCACCCAACGCCTGTACTGTGGAGGCCCTATGTA
CAACAGCAAGGGGCAACAATGTGGTTATCGTAGATGCCGCGCCAGCGGCGTCTT
CACCACTAGTATGGGCAACACCATAACGTGCTACGTTAAGGCTTTGGCCTCCTGT
AGAGCTGCAAAGCTCCGGGACTGCACGCTCTTGGTGTGTGGTGATGATCTTGTGG
TCATCTGCGGAAGTCAGGAAACCCACGAGGACGCGGCGAGCCTGAGAGCTCTCA
ACCGTCACTGAGAGAGACATCCGCATTGAGGTGTCAATCTACCAATAGTTGCGA
CTTGGCCCCCGAGGCCAGACAGGCTATAAGGTCGCTCACAGAGCGGCTTTATATC
GGGGGTCCCCTGACTAATTCAAAAGGGCAGAACTGCGGTTATCGCCGGTGCCGC
GCAAGCGGCGTGCTGACGACTAGCTGCGGTAATACCCTCACATGTTACTTGAAA
GCCACTGCGGCCTGTCGAGCTGCAAAGCTCCAGGACTGCACAATGCTCGTGAAC
GGGGACGACCTTGTCGTCATCCTTGAGAGCGCGGGAACCCAAGAGGACGCGGCG
AGCCTGAGCGCC

>NGR-JS-LD12
GGAGGAGGCGATCTACCAATGTTGTGACCTGGACCCCCAGGCCCGCGTGGCCAT
CAAGTCCCTCACTGAAAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGG
GAGAACTGCGGCTACCGCAGGTGCCGCGCAAGCGGCGTACTGACGACTAGCTGT
GGTAACACCCTCACTTGCTATATCAAGGCCAAGGCAGCCTGTCGAGCCGCAGGG
CTCCAGGACTGCACCATGCTCGTGTGTGGCGACGACCTAGTCGTTAATTTATGGA
TCTCTCAACCGTCACTGAGAGAGACATAAGAACAGAAGAATCCATATATCAGGC
TTGTTCCCTGCCCCAAGAGGCCAGAACTGCCATACACTCGCTTACTGAGAGACTT
TACGTAGGAGGGCCCATGACAAACAGCAAAGGGCAATCCTGCGGTTACAGGCGT
TGCCGCGCAAGCGGCGTCTTCACTACCAGCATGGGGAATACCATGACATGTTAC
ATCAAAGCCCTTGCAGCGTGCAAAGCTGCGGGGATCGTGGACCCTGTTATGCTG
GTGTGTGGAGATGACCTGGTCGTTATCTCAGAGAGCCAAGGTAACGGGGAGGAC
GCGGCGAGCCCTGAGAGCAGGAGGAGGCGATCTACCAATGTTGTGACCTGGACC
CCCAGGCCCGCGTGGCCATCAAGTCCCTCACTGAAAGGCTTTATGTTGGGGGCCC
TCTTACCAATTCAAGGGGGGAGAACTGCGGCTACCGCAGGTGCCGCGCAAGCGG
CGTACTGACGACTAGCTGTGGTAACACCCTCACTTGCTATATCAAGGCCAAGGCA
GCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGTGTGTGGCGACGAC
CTAGTCGTTAATTTATGGATCTCTCAACCGTCACTGAGAGAGACATAAGAACAGA
AGAATCCATATATCAGGCTTGTTCCCTGCCCCAAGAGGCCAGAACTGCCATACAC
TCGCTTACTGAGAGACTTTACGTAGGAGGGCCCATGACAAACAGCAAAGGGCAA
TCCTGCGGTTACAGGCGTTGCCGCGCAAGCGGCGTCTTCACTACCAGCATGGGGA
ATACCATGACATGTTACATCAAAGCCCTTGCAGCGTGCAAAGCTGCGGGGATCG
TGGACCCTGTTATGCTGGTGTGTGGAGATGACCTGGTCGTTATCTCAGAGAGCCA
AGGTAACGGGGAGGACGCGGCGAGCCCTGAGAGCA

>NGR-JS-LD13
TCTCAACCGTCACTGAGAGAGACATCCGTACGGAGGAGGCAATCTACCAATGTT
GTGACCTGGACCCCCAAGCCCGTGTGGCCATCAGGTCTCTCACCGAGAGGCTTTA
TGTCGGGGGCCCTCTTACCAACTCAAGGGGGGAAAACTGCGGCTATCGCAGGTG
CCGCGCAAGCGGCGTGCTGACAACTAGCTGTGGTAACACCCTCACTTGCTACATC
AAGGCCCGAGCAGCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGTG
TGTGGCGACGACTTGGTCGTTATCTGTGAAAGTCAGGGGGTCCAGGAGGACGCG
GCGAGCCTGAGAGGTTTTTTTAAGCCTCCCCCCTCCCTCTCAACCGTCACTGAGA
GAGACATAAGAACAGAAGAATCCATATATCAGGCTTGTTCCCTGCCCCAAGAGG
CCAGAACTGCCATACACTCGCTCACTGAGAGACTTTACGTAGGAGGGCCCATGA
CAAACAGCAAAGGGCAATCCTGCGGTTACAGGCGTTGCCGCGCAAGCGGCGTCT
TCACCACCAGCATGGGGAATACCATGACATGTTACATCAAAGCCCTTGCAGCGT
GCAAAGCTGCGGGGATTGTGGGCCCGTGTTATGCGTGGTGTGTCGAGCTGCACG
CTCGTCGTCATGCCAGAGAGCCAAGGTAACGAGGAGGACGCGGCGAGCCTGAAA
GATTGTCTCTCTCAGTGACGAGAGAGGTCTCAACCGTCACTGAGAGAGACATCC
GTACGGAGGAGGCAATCTACCAATGTTGTGACCTGGACCCCCAAGCCCGTGTGG
CCATCAGGTCTCTCACCGAGAGGCTTTATGTCGGGGGCCCTCTTACCAACTCAAG
GGGGGAAAACTGCGGCTATCGCAGGTGCCGCGCAAGCGGCGTGCTGACAACTAG
CTGTGGTAACACCCTCACTTGCTACATCAAGGCCCGAGCAGCCTGTCGAGCCGCA
GGGCTCCAGGACTGCACCATGCTCGTGTGTGGCGACGACTTGGTCGTTATCTGTG
AAAGTCAGGGGGTCCAGGAGGACGCGGCGAGCCTGAGAGGTTTTTTTAAGCCTC
CCCCCTCCCTCTCAACCGTCACTGAGAGAGACATAAGAACAGAAGAATCCATAT
ATCAGGCTTGTTCCCTGCCCCAAGAGGCCAGAACTGCCATACACTCGCTCACTGA

GAGACTTTACGTAGGAGGGCCCATGACAAACAGCAAAGGGCAATCCTGCGGTTA
CAGGCGTTGCCGCGCAAGCGGCGTCTTCACCACCAGCATGGGGAATACCATGAC
ATGTTACATCAAAGCCCTTGCAGCGTGCAAAGCTGCGGGGATTGTGGGCCCGTGT
TATGCGTGGTGTGTCGAGCTGCACGCTCGTCGTCATGCCAGAGAGCCAAGGTAA
CGAGGAGGACGCGGCGAGCCTGAAAGATTGTCTCTCTCAGTGACGAGAGAGG

>NGR-JS-LD15
TCTCAACCGTCACTGAGAGAGACATCCGTACGGAGGAAGCAATCTACCAATAGT
TGTGACCTGGACCCCCAAGCCCGCGTGGCCATCAAGTCTCTCACTGAGAGGCTTT
ATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAAAACTGCGGCTACCGCAGGT
GCCGCGCGAGCGGCGTACTGACTACTAGCTGTGGTAACACCCTCACTTGCTACAT
CAAGGCCCGAGCAGCCTGTCGAGCCGCAGGGCTCCGGGACTGCACCATGCTCGT
GTGTGGCGACGACCTAGTCGCTTATCTGCGAAAGTGCGGGGGTCCAGGAGGACG
CGGCGAGCCTGAGACTTCTCAACCGTCACTGAGAGAGACATCCGTACGGAGGAG
ACTCTTTACCAATGTTGTGACTTGGCCCCCGAAGCCCGCAAGGTCGTGAAGTCTC
TCACCGAGAGGCTTTACGTAGGGGGCCCCCTCACCAATTCCAAGGGACAAAACT
GCGGCTATCGCAGATGCCGCGCAAGCGGCGTCCTGACCACTAGCTGCGGTAACA
CCATTACTTGCTATCTCAAAGCTGCTGCTGCCTGTCGAGCTGCAAAGCTCCAGAG
CTGCACCATGCTGGTGTGCGGCGCACGCTCTCTAGTTATCTGTGAGAGCAGTGGG
GTTGAGGAGGACGCGGCGAGCCTGAGAGCATCTCAACCGTCACTGAGAGAGACA
TCCGTACGGAGGAAGCAATCTACCAATAGTTGTGACCTGGACCCCCAAGCCCGC
GTGGCCATCAAGTCTCTCACTGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATT
CAAGGGGGGAAAACTGCGGCTACCGCAGGTGCCGCGCGAGCGGCGTACTGACTA
CTAGCTGTGGTAACACCCTCACTTGCTACATCAAGGCCCGAGCAGCCTGTCGAGC
CGCAGGGCTCCGGGACTGCACCATGCTCGTGTGTGGCGACGACCTAGTCGCTTAT
CTGCGAAAGTGCGGGGGTCCAGGAGGACGCGGCGAGCCTGAGACTTCTCAACCG
TCACTGAGAGAGACATCCGTACGGAGGAGACTCTTTACCAATGTTGTGACTTGGC
CCCCGAAGCCCGCAAGGTCGTGAAGTCTCTCACCGAGAGGCTTTACGTAGGGGG
CCCCCTCACCAATTCCAAGGGACAAAACTGCGGCTATCGCAGATGCCGCGCAAG
CGGCGTCCTGACCACTAGCTGCGGTAACACCATTACTTGCTATCTCAAAGCTGCT
GCTGCCTGTCGAGCTGCAAAGCTCCAGAGCTGCACCATGCTGGTGTGCGGCGCA
CGCTCTCTAGTTATCTGTGAGAGCAGTGGGGTTGAGGAGGACGCGGCGAGCCTG
AGAGCA

>NGR-JS-LD15b
TTCTCAACCGTCACTGAGAGAGACATTATGACTGAAGAGTCCATCTACCAATACA
TGTGACTTACAACCTGAGGCGCGCACGGCAATACGGTCACTCACCCAACGCCTGT
ACTGTGGAGGCCCTATGTATAACAGCAAGGGGCAGCAATGTGGTTATCGTAGAT
GCCGCGCCAGCGGCGTCTTCACTACCAGTATGGGCAACACCATGACGTGTTACAT
TAAGGCTTTAGCCTCCTGTAGAGCTGCAAAGCTCCGGGATTGTACGCTCCTGGTG
TGTGGTGACGACCTTGTGGCCATCTGCGAGAGCCAAGGGACACACGAGGACGCG
GCGAGCCTGAGAGCACGCGCCCTCTCAACCGTCACTGAGAGAGACATCCGTACG
GAGGAGACTCTTTACCAATGTTGTGACTTGGCCCCCGAAGCCCGCAAGGTCGTGA
AGTCTCTCACCGAGAGGCTTTACGTAGGGGGCCCCCTCACCAATTCCAAGGGAC
AAAACTGCGGCTATCGCAGATGCCGCGCAAGCGGCGTCCTGACCACTAGCTGCG
GTAACACCATTACTTGCTATCTCAAAGCTGCTGCTGCCTGTCGAGCTGCAAAGCT
CCAGAGCTGCACCATGCTGGTGTGCGGCGCACCCCCTAGTCGTCATCTGTGAGAG
CAGTGGGGTTGAGGAGGACGCGGCGAGCCTGAGAGCATTCTCAACCGTCACTGA
GAGAGACATTATGACTGAAGAGTCCATCTACCAATACATGTGACTTACAACCTG
AGGCGCGCACGGCAATACGGTCACTCACCCAACGCCTGTACTGTGGAGGCCCTA

TGTATAACAGCAAGGGGCAGCAATGTGGTTATCGTAGATGCCGCGCCAGCGGCG
TCTTCACTACCAGTATGGGCAACACCATGACGTGTTACATTAAGGCTTTAGCCTC
CTGTAGAGCTGCAAAGCTCCGGGATTGTACGCTCCTGGTGTGTGGTGACGACCTT
GTGGCCATCTGCGAGAGCCAAGGGACACACGAGGACGCGGCGAGCCTGAGAGC
ACGCGCCCTCTCAACCGTCACTGAGAGAGACATCCGTACGGAGGAGACTCTTTA
CCAATGTTGTGACTTGGCCCCCGAAGCCCGCAAGGTCGTGAAGTCTCTCACCGAG
AGGCTTTACGTAGGGGGCCCCCTCACCAATTCCAAGGGACAAAACTGCGGCTAT
CGCAGATGCCGCGCAAGCGGCGTCCTGACCACTAGCTGCGGTAACACCATTACTT
GCTATCTCAAAGCTGCTGCTGCCTGTCGAGCTGCAAAGCTCCAGAGCTGCACCAT
GCTGGTGTGCGGCGCACCCCCTAGTCGTCATCTGTGAGAGCAGTGGGGTTGAGG
AGGACGCGGCGAGCCTGAGAGCA

>NGR-JS-LD16
TGGGGATCCCGTATGATACCCGCTGCTTTGACTCAACCGTTACCGAACATGACAT
TATGACTGAAGAGTCCATCTACCAATCATGTGACTTACAACCTGAGGCGCGCACG
GCAATACGGTCACTCACCCAACGCCTGTACTGTGGAGGCCCTATGTATAACAGCA
AGGGGCAGCAATGTGGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACTACCA
GTATGGGCAACACCATGACGTGTTACATTAAGGCTTTAGCCTCCTGTAGAGCTGC
AAAGCTCCGGGATTGTACGCTCCTGGTGTGTGGTGACGACCTTGTGGCCATCTGC
GAGAGCCAAGGGACACACGAGGATGAAGCGAGCCTGAGAGCCTTCACGGAGGC
TATGACCAGGAATTCCGCCATCTCAACCGTTCACTGAGAGAGACATAATGACTG
AAGAGTCTATTTACCAATGCACAGTAGACATGCAGCCTGAGGCGCGCGTGGCAA
TACGGTCACTCACCCAACGCCTGTACTGTGGAGGCCCTATGTATAACAGCAAGG
GGCAACAATGTGGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACCACTAGTAT
GGGCAACACTATGACGTGCTACATCAAGGCTTTAGCCTCCTGTAGAGCTGCAAA
GCTCCAGGATTGCACGCTCCTGGTGTGTGGTGATGACCTCGTGGCCATCTGCGAG
AGCCAGGGGACACACGAGGACGCGGCGAGCCCTGAGAGCATGGGGATCCCGTA
TGATACCCGCTGCTTTGACTCAACCGTTACCGAACATGACATTATGACTGAAGAG
TCCATCTACCAATCATGTGACTTACAACCTGAGGCGCGCACGGCAATACGGTCAC
TCACCCAACGCCTGTACTGTGGAGGCCCTATGTATAACAGCAAGGGGCAGCAAT
GTGGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACTACCAGTATGGGCAACAC
CATGACGTGTTACATTAAGGCTTTAGCCTCCTGTAGAGCTGCAAAGCTCCGGGAT
TGTACGCTCCTGGTGTGTGGTGACGACCTTGTGGCCATCTGCGAGAGCCAAGGGA
CACACGAGGATGAAGCGAGCCTGAGAGCCTTCACGGAGGCTATGACCAGGAATT
CCGCCATCTCAACCGTTCACTGAGAGAGACATAATGACTGAAGAGTCTATTTACC
AATGCACAGTAGACATGCAGCCTGAGGCGCGCGTGGCAATACGGTCACTCACCC
AACGCCTGTACTGTGGAGGCCCTATGTATAACAGCAAGGGGCAACAATGTGGTT
ATCGTAGATGCCGCGCCAGCGGCGTCTTCACCACTAGTATGGGCAACACTATGAC
GTGCTACATCAAGGCTTTAGCCTCCTGTAGAGCTGCAAAGCTCCAGGATTGCACG
CTCCTGGTGTGTGGTGATGACCTCGTGGCCATCTGCGAGAGCCAGGGGACACAC
GAGGACGCGGCGAGCCCTGAGAGCA

>NGR-JS-LD16b
TTCTCAACCGTCACTGAGAGAGACATTATACCTGAAGAGTCCATCTACCAATCAT
GTGACTTACAACCTGAGGCGCGCACGGCAATGCGGTCACTCACCCAACGCCTGT
ACTGTGGAGGCCCTATGTATAACAGCAAGGGGCAGCAATGTGGTTATCGTAGAT
GCCGCGCCAGCGGCGTCTTCACTACCAGTATGGGCAACACCATGACGTGTTACAT
TAAGGCTTTAGCCTCCTGTAGAGCTGCGAAGCTCCGGGATTGCACGCTCCTGGTG
TGTGGTGACGACCTTGTGGCCATCTGCGAGAGCCAAGGGACACACGAGGACGCG

GCGAGCCTGAGAGCATCTCAACCGTCACTGAGAGAGACATAATGACTGAAGAGT
CTATTTACCAATCATGTAGACTTGCAGCCTGAGGCGCGCGTGGCAATACGGTCAC
TCACCCAACGCCTGTACTGTGGAGGCCCTATGTATAACAGCAAGGGGCAACAAT
GTGGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACCACTAGTATGGGCAACAC
TATGACGTGCTACATCAAGGCTTTAGCCTCCTGTAGAGCTGCAAAGCTCCAGGAT
TGCACGCTCCTGGTGTGTGGTGATGACCTTGTGGCCATCTGCGAGAGCCAGGGGA
CACACGAGGACGCGGCGAGACTGNGAGCATTCTCAACCGTCACTGAGAGAGACA
TTATACCTGAAGAGTCCATCTACCAATCATGTGACTTACAACCTGAGGCGCGCAC
GGCAATGCGGTCACTCACCCAACGCCTGTACTGTGGAGGCCCTATGTATAACAGC
AAGGGGCAGCAATGTGGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACTACC
AGTATGGGCAACACCATGACGTGTTACATTAAGGCTTTAGCCTCCTGTAGAGCTG
CGAAGCTCCGGGATTGCACGCTCCTGGTGTGTGGTGACGACCTTGTGGCCATCTG
CGAGAGCCAAGGGACACACGAGGACGCGGCGAGCCTGAGAGCATCTCAACCGT
CACTGAGAGAGACATAATGACTGAAGAGTCTATTTACCAATCATGTAGACTTGC
AGCCTGAGGCGCGCGTGGCAATACGGTCACTCACCCAACGCCTGTACTGTGGAG
GCCCTATGTATAACAGCAAGGGGCAACAATGTGGTTATCGTAGATGCCGCGCCA
GCGGCGTCTTCACCACTAGTATGGGCAACACTATGACGTGCTACATCAAGGCTTT
AGCCTCCTGTAGAGCTGCAAAGCTCCAGGATTGCACGCTCCTGGTGTGTGGTGAT
GACCTTGTGGCCATCTGCGAGAGCCAGGGGACACACGAGGACGCGGCGAGACTG
NGAGCA

>NGR-JS-LD20
TCTCAACCGTCACTGAGAGAGACACCGGTACGGAGGAGGCAATCTACCAGTGTT
GTGACCTGGACCCCCAAGCCCGTGTGGCCATCAAGTCTCTCACCGAGAGGCTTTA
TGTCGGGGGCCCTCTTACCAATTCAAGGGGGGAAAACTGTGGCTATCGCAGGTG
CCGCGCGAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTCACTTGCTACATC
AAGGCCCAGGCAGCCTGTCGAGCCGCAGGGCTCCGGGACTGCACCATGCTCGTG
TGTGGCGACGACCTAGTCGTTATCAGCGAAAGTCAGGGAGTCCAGGAGGACGCG
GCGAGCCTGAGAGCATTGTTTTAAGGGCTCGCCGCGTCCTCTCAACCGTCACTGA
GAGAGACATAAGAACAGAAGAATCCATATATCAGGCTTGTTCCCTGCCCCAAGA
GGCCAGAACTGCCATACACTCGCTCACTGAGAGACTTTACGTAGGAGGGCCCAT
GACAAACAGCAAAGGGCAATCCTGCGGTTACAGGCGTTGCCGCGCAAGCGGCGT
CTTCACCACCAGCATGGGGAATACCATGACATGTTACATCAAAGCCCTTGCAGCG
TGCAAAGCTGCGGGGATTGTGGACCCTGTTATGCTGGTGTGTGGAGATGACCTGG
TCGTTATCTCAGAGAGCCAAGGTAACGAGGAGGACGCGGCGCCCTGATAGCTTG
TCTCTCTCAGTGACGGTTGAGGATCTCAACCGTCACTGAGAGAGACACCGGTACG
GAGGAGGCAATCTACCAGTGTTGTGACCTGGACCCCCAAGCCCGTGTGGCCATC
AAGTCTCTCACCGAGAGGCTTTATGTCGGGGGCCCTCTTACCAATTCAAGGGGGG
AAAACTGTGGCTATCGCAGGTGCCGCGCGAGCGGCGTACTGACAACTAGCTGTG
GTAACACCCTCACTTGCTACATCAAGGCCCAGGCAGCCTGTCGAGCCGCAGGGC
TCCGGGACTGCACCATGCTCGTGTGTGGCGACGACCTAGTCGTTATCAGCGAAAG
TCAGGGAGTCCAGGAGGACGCGGCGAGCCTGAGAGCATTGTTTTAAGGGCTCGC
CGCGTCCTCTCAACCGTCACTGAGAGAGACATAAGAACAGAAGAATCCATATAT
CAGGCTTGTTCCCTGCCCCAAGAGGCCAGAACTGCCATACACTCGCTCACTGAGA
GACTTTACGTAGGAGGGCCCATGACAAACAGCAAAGGGCAATCCTGCGGTTACA
GGCGTTGCCGCGCAAGCGGCGTCTTCACCACCAGCATGGGGAATACCATGACAT
GTTACATCAAAGCCCTTGCAGCGTGCAAAGCTGCGGGGATTGTGGACCCTGTTAT
GCTGGTGTGTGGAGATGACCTGGTCGTTATCTCAGAGAGCCAAGGTAACGAGGA
GGACGCGGCGCCCTGATAGCTTGTCTCTCTCAGTGACGGTTGAGGA

>NGR-JS-P02
TCTCAACCGTCACTGAGAGAGACATCCGTACGGAGGAGGCAATCTACCAATGTT
GTGACCTGGACCCCCAAGCCCGCGTGGCCATCAAGTCTCTCACTGAGAGGCTTTA
TGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAAAACTGCGGCTACCGCAGGTG
CTGCGCGAGCGGCGTACTGACTACTAGCTGTGGTAACACCCTCACTTGCTACATC
AAGGCCCGAGCAGCCTGTCGAGCCGCAGGGCTCCGGGACTGCACCATGCTCGTG
TGTGGCGACGACCTAGTCGTTATCTGCGAAAGTGCGGGGGTCCAGGAGGACGCG
GCGAGCCTGAGAGCACTCAACCGTCACTGAGAGAGACATAATGACCGAAGAGTC
CATTTACCAATCATGTGACTTGCAGCCTGAGGCGCGCGCGGCAATACGGTCACTC
ACCCAACGCCTGTACTGTGGAGGCCCTATGTATAACAGCAAAGGGCAACAATGT
GGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACCACTAGTATGGGCAACACCA
TAACGTGCTACGTTAAGGCTTTGGCCTCCTGTAGAGCTGCAAAGCTCCGGGACTG
CACGCTCTTGGTGTGTGGTGATGATCTTGGCGCGCTGCCTTCAGGATCTCAACCG
TCACTGAGAGAGACATCCGTACGGAGGAGGCAATCTACCAATGTTGTGACCTGG
ACCCCCAAGCCCGCGTGGCCATCAAGTCTCTCACTGAGAGGCTTTATGTTGGGGG
CCCTCTTACCAATTCAAGGGGGGAAAACTGCGGCTACCGCAGGTGCTGCGCGAG
CGGCGTACTGACTACTAGCTGTGGTAACACCCTCACTTGCTACATCAAGGCCCGA
GCAGCCTGTCGAGCCGCAGGGCTCCGGGACTGCACCATGCTCGTGTGTGGCGAC
GACCTAGTCGTTATCTGCGAAAGTGCGGGGGTCCAGGAGGACGCGGCGAGCCTG
AGAGCACTCAACCGTCACTGAGAGAGACATAATGACCGAAGAGTCCATTTACCA
ATCATGTGACTTGCAGCCTGAGGCGCGCGCGGCAATACGGTCACTCACCCAACG
CCTGTACTGTGGAGGCCCTATGTATAACAGCAAAGGGCAACAATGTGGTTATCGT
AGATGCCGCGCCAGCGGCGTCTTCACCACTAGTATGGGCAACACCATAACGTGC
TACGTTAAGGCTTTGGCCTCCTGTAGAGCTGCAAAGCTCCGGGACTGCACGCTCT
TGGTGTGTGGTGATGATCTTGGCGCGCTGCCTTCAGGA

>NGR-JS-P03
TCTCAACCGTTCACTGAGAGAGACATCCGTACGGAGGAGGCAATCTACCAATGT
TGTGACCTTGACCCCCAAGCCCGTGTGGCCATCAAGTCCCTCACCGAGAGACTTT
ACGTTGGGGGCCCTCTTACCAACTCAAGGGGGGAGAACTGCGGCTACCGCAGGT
GCCGCGCGAGCGGCGTACTGACAACCAGCTGTGGTAACACCCTCACTTGCTACAT
CAAGGCCCAGGCAGCCTGTCGAGCCGCAGGGCTCCGGGACTGCACCATGCTCGT
GTGTGGCGACGACTTGGTCGTTATCTGTGAAAGTGCGGGGATCCAGGAGGACGC
GGCGAGCCTGAGAGCACTCAACCGTCACTGAGAGAGACATCCGTACGGAAGAGG
CAATCTACCAGTGTTGTGACCTGGACCCCCAAGCCCGTGTGGCCATCAAGTCTCT
CACCGAGAGGCTTTATGTCGGGGGCCCCCTTACCAATTCAAGGGGGGAAAACTG
CGGCTATCGCAGGTGCCGCGCGAGCGGCGTACTGACAACTAGCTGTGGTAACAC
CCTCACTTGCTACATTAAGGCCCAAGCAGCCTGTCGAGCCGCAGGGCTCCGGGA
CTGCACCATGCTCGTGTGTGGCGACGACTTAGTCGTTATCTGTGAAAGTCAGGGA
GTCCAGGAGGACGCGGCGAGCCTGAGAGCTCTCAACCGTTCACTGAGAGAGACA
TCCGTACGGAGGAGGCAATCTACCAATGTTGTGACCTTGACCCCCAAGCCCGTGT
GGCCATCAAGTCCCTCACCGAGAGACTTTACGTTGGGGGCCCTCTTACCAACTCA
AGGGGGGAGAACTGCGGCTACCGCAGGTGCCGCGCGAGCGGCGTACTGACAACC
AGCTGTGGTAACACCCTCACTTGCTACATCAAGGCCCAGGCAGCCTGTCGAGCCG
CAGGGCTCCGGGACTGCACCATGCTCGTGTGTGGCGACGACTTGGTCGTTATCTG
TGAAAGTGCGGGGATCCAGGAGGACGCGGCGAGCCTGAGAGCACTCAACCGTCA
CTGAGAGAGACATCCGTACGGAAGAGGCAATCTACCAGTGTTGTGACCTGGACC
CCCAAGCCCGTGTGGCCATCAAGTCTCTCACCGAGAGGCTTTATGTCGGGGGCCC
CCTTACCAATTCAAGGGGGGAAAACTGCGGCTATCGCAGGTGCCGCGCGAGCGG

CGTACTGACAACTAGCTGTGGTAACACCCTCACTTGCTACATTAAGGCCCAAGCA
GCCTGTCGAGCCGCAGGGCTCCGGGACTGCACCATGCTCGTGTGTGGCGACGAC
TTAGTCGTTATCTGTGAAAGTCAGGGAGTCCAGGAGGACGCGGCGAGCCTGAGA
GC

>NGR-JS-P06
TTCTCAACCTTTCACTGAGAGAGCCATCCGTACGGAGGAGGCAATTTACCAATGT
TGTGACCTGGACCCCCAAGCCCGCGTGGCCATCAAGTCTCTCACTGAGAGGCTTT
ATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAAAACTGCGGCTATCGCAGGT
GCCGCGCGAGCGGCGTACTGACTACTAGCTGTGGTAACACCCTCACTTGCTACAT
CAAGGCCCGAGCAGCCTGTCGAGCCGCAGGGCTCCGGGACTGCACCATGCTCGT
GTGTGGCGACGACCTAGTCCTTATCTGCGAAAGTGCGGGGTCCAGGAGGACGCG
GCGGACCTGAAAGCAATCTCAACCGTCACTGAGAGAGACATAATGACCGAAGAG
TCCATTTACCAATCATGTGACTTGCAGCCTGAGGCGCGCGCGGCAATACGGTCAC
TCACCCAACGCCTGTACTGTGGAGGCCCTATGTATAACAGCAAAGGGCAACAAT
GTGGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACCACTAGTATGGGCAACAC
CATAACGTGCTACGTTAAGGCTTTGGCCTCCTGTAGAGCTGCAAAGCTCCGGGAC
TGCACGCTCTTGGTGTGTGGTGATGATCTTGTGGCCATCTGCGAAAGCCAGGGGA
CACACGAGGACGCGGCGAGCCTGAGAGCTTCTCAACCTTTCACTGAGAGAGCCA
TCCGTACGGAGGAGGCAATTTACCAATGTTGTGACCTGGACCCCCAAGCCCGCGT
GGCCATCAAGTCTCTCACTGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCA
AGGGGGGAAAACTGCGGCTATCGCAGGTGCCGCGCGAGCGGCGTACTGACTACT
AGCTGTGGTAACACCCTCACTTGCTACATCAAGGCCCGAGCAGCCTGTCGAGCCG
CAGGGCTCCGGGACTGCACCATGCTCGTGTGTGGCGACGACCTAGTCCTTATCTG
CGAAAGTGCGGGGTCCAGGAGGACGCGGCGGACCTGAAAGCAATCTCAACCGTC
ACTGAGAGAGACATAATGACCGAAGAGTCCATTTACCAATCATGTGACTTGCAG
CCTGAGGCGCGCGCGGCAATACGGTCACTCACCCAACGCCTGTACTGTGGAGGC
CCTATGTATAACAGCAAAGGGCAACAATGTGGTTATCGTAGATGCCGCGCCAGC
GGCGTCTTCACCACTAGTATGGGCAACACCATAACGTGCTACGTTAAGGCTTTGG
CCTCCTGTAGAGCTGCAAAGCTCCGGGACTGCACGCTCTTGGTGTGTGGTGATGA
TCTTGTGGCCATCTGCGAAAGCCAGGGGACACACGAGGACGCGGCGAGCCTGAG
AGC

>NGR-JS-P07
TCTGAACCGTTCACTGAGAGAGACATCCGTACGGAGGAGGCAATCTACCAATGT
TGTGACCTGGACCCCCAAGCCCGCGTGGCCATCAAGTCCCTCACTGAGAGGCTTT
ATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGCTACCGCAGGT
GCCGCGCGAGCGGCGTACTGACTACTAGCTGTGGTAACACCCTCACTTGCTACAT
CAAGGCCCGGGCAGCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGT
GTGTGGCGACGACTTAGTCGTTATCTGTGAAAGTGCGGGGGTCCAGGAGGACGC
GGCGAGCCTGAGAGCATCTCAACCGTCACTGAGAGAGACATAATGACCGAAGAG
TCCATTTACCAATCATGTGACTTGCAGCCTGAGGCGCGCGCGGCAATACGGTCAC
TCACCCAACGCCTGTACTGTGGAGGCCCTATGTATAACAGCAAGGGGCAACAAT
GTGGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACCACTAGTATGGGCAACAC
CATAACGTGCTACGTTAAGGCTTTGGCCTCCTGTTGAGCTGCAAAGCTCCGGGAC
TGCACGCTCTTGGTGTGTGGTGATGATCTTGTGGCCATCTGCGAGAGCCCGGGGA
CACACGAGGACGCGGCGACCCTGAGAGCATCTGAACCGTTCACTGAGAGAGACA
TCCGTACGGAGGAGGCAATCTACCAATGTTGTGACCTGGACCCCCAAGCCCGCG
TGGCCATCAAGTCCCTCACTGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTC

AAGGGGGGAGAACTGCGGCTACCGCAGGTGCCGCGCGAGCGGCGTACTGACTAC
TAGCTGTGGTAACACCCTCACTTGCTACATCAAGGCCCGGGCAGCCTGTCGAGCC
GCAGGGCTCCAGGACTGCACCATGCTCGTGTGTGGCGACGACTTAGTCGTTATCT
GTGAAAGTGCGGGGGTCCAGGAGGACGCGGCGAGCCTGAGAGCATCTCAACCGT
CACTGAGAGAGACATAATGACCGAAGAGTCCATTTACCAATCATGTGACTTGCA
GCCTGAGGCGCGCGCGGCAATACGGTCACTCACCCAACGCCTGTACTGTGGAGG
CCCTATGTATAACAGCAAGGGGCAACAATGTGGTTATCGTAGATGCCGCGCCAG
CGGCGTCTTCACCACTAGTATGGGCAACACCATAACGTGCTACGTTAAGGCTTTG
GCCTCCTGTTGAGCTGCAAAGCTCCGGGACTGCACGCTCTTGGTGTGTGGTGATG
ATCTTGTGGCCATCTGCGAGAGCCCGGGGACACACGAGGACGCGGCGACCCTGA
GAGCA

>NGR-JS-P09a
TCTCAACCGTCACTGAGAGAGACATACGTACCGAGCAGCCTATCTATCGATTGTG
CACTGTCTCGTGCCAAGCTCGCGATGGCCATCAAGTCCCTGACTGAGAGACTATT
ATGTGGGAGGCCCTATGACCAACACAAAGGGGGAATACTGCGGCTACAGCAGGT
GCCGCGCGAGCGGAGTGCTCACTACTAGCATGGGTAACACCCTCACGTGCTACG
TCAAGGCCCAGGCAGCCTGTAGAGCCGCGGGCATCCAGGCCCCCACCATGCTGG
TATGTGGCGACGACCTAGTCGTCATCTCAAAGAGTGCGGGGCTCCAGGAGGACG
CGGCGAGCCCTGAAAGCCTCTCAACCGTCACTGAGAGAGACATCCGTGTTGAGG
AGTTAATTTACCAATCATGTGACTTGCACCCTGAGGCGCGCGCGGCAATACGGTC
ACTCACCCAACGCCTGTATATTGGAGGCCCTATGTATAACAGCAAGGGGCAACA
ATGTGGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACCACTAGTATGGGCAAC
ACCATAACGTGCTACGTTAAGGCTTTGGCCTCCTGTGGAGCTGCAAAGCTCCGGG
ACTGCCACGCTCTTGGTGTGTGGTGATGATCTCGTGGCCATCTGCGAGAGCCAGG
GGACACACGAGGACGCGGCGACCGTGAGAGCTCTCAACCGTCACTGAGAGAGAC
ATACGTACCGAGCAGCCTATCTATCGATTGTGCACTGTCTCGTGCCAAGCTCGCG
ATGGCCATCAAGTCCCTGACTGAGAGACTATTATGTGGGAGGCCCTATGACCAA
CACAAAGGGGGAATACTGCGGCTACAGCAGGTGCCGCGCGAGCGGAGTGCTCAC
TACTAGCATGGGTAACACCCTCACGTGCTACGTCAAGGCCCAGGCAGCCTGTAG
AGCCGCGGGCATCCAGGCCCCCACCATGCTGGTATGTGGCGACGACCTAGTCGT
CATCTCAAAGAGTGCGGGGCTCCAGGAGGACGCGGCGAGCCCTGAAAGCCTCTC
AACCGTCACTGAGAGAGACATCCGTGTTGAGGAGTTAATTTACCAATCATGTGAC
TTGCACCCTGAGGCGCGCGCGGCAATACGGTCACTCACCCAACGCCTGTATATTG
GAGGCCCTATGTATAACAGCAAGGGGCAACAATGTGGTTATCGTAGATGCCGCG
CCAGCGGCGTCTTCACCACTAGTATGGGCAACACCATAACGTGCTACGTTAAGGC
TTTGGCCTCCTGTGGAGCTGCAAAGCTCCGGGACTGCCACGCTCTTGGTGTGTGG
TGATGATCTCGTGGCCATCTGCGAGAGCCAGGGGACACACGAGGACGCGGCGAC
CGTGAGAGC

>NGR-JS-P09b
TCTCAACCGTCACTGAGAGAGACATAATGACCGAAGAGTCCATTTACCAATCAT
GTGACTTGCAGCCTGAGGCGCGCGCGGCAATACGGTCACTCACCCAACGCCTGT
ACTGTGGAGGCCCTATGTATAACAGCAAAGGGCAACAATGTGGTTATCGTAGAT
GCCGCGCCAGCGGCGTCTTCACCACTAGTATGGGCAACACCATAACGTGCTACGT
TAAGGCTTTGGCCTCCTGTAGAGCTGCAAAGCTCCGGGACTGCACGCTCTTGGTG
TGTGGTGATGATCTTGTGGCCATCTGCGAAAGCCAGGGGACACACGAGGACGCG
GCGAGCCTGAGAGCACCAATCATGTGACTTGCAGCCTGAGGCGCGCGCGGCAAT
ACGGTCACTCACCCAACGCCTGTACTGTGGAGGCCCTATGTATAACAGCAAGGG
GCAACAATGTGGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACCACTAGTATG

GGCAACACCATAACGTGCTACGTTAAGGCTTTGGCCTCCTGTGGAGCTGCAAAGC
TCCGGGACTGCACGCTCTTGGTGTGTGGTGATGATCTCGTGGCCATCTGCGAGAG
CCAGGGGACATCTCAACCGTCACTGAGAGAGACATAATGACCGAAGAGTCCATT
TACCAATCATGTGACTTGCAGCCTGAGGCGCGCGCGGCAATACGGTCACTCACCC
AACGCCTGTACTGTGGAGGCCCTATGTATAACAGCAAAGGGCAACAATGTGGTT
ATCGTAGATGCCGCGCCAGCGGCGTCTTCACCACTAGTATGGGCAACACCATAA
CGTGCTACGTTAAGGCTTTGGCCTCCTGTAGAGCTGCAAAGCTCCGGGACTGCAC
GCTCTTGGTGTGTGGTGATGATCTTGTGGCCATCTGCGAAAGCCAGGGGACACAC
GAGGACGCGGCGAGCCTGAGAGCACCAATCATGTGACTTGCAGCCTGAGGCGCG
CGCGGCAATACGGTCACTCACCCAACGCCTGTACTGTGGAGGCCCTATGTATAAC
AGCAAGGGGCAACAATGTGGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACC
ACTAGTATGGGCAACACCATAACGTGCTACGTTAAGGCTTTGGCCTCCTGTGGAG
CTGCAAAGCTCCGGGACTGCACGCTCTTGGTGTGTGGTGATGATCTCGTGGCCAT
CTGCGAGAGCCAGGGGACA

>NGR-JS-P10
TCTCAACCGTCACTGAGAGAGACATCCGTACCGACGAGTCTATGTACCAATCATG
TGACCTGGACCCGCAGGCGCGCGCGGCCATTACAAGTCACCTCACCCAAAAGGC
TGTATGGTGGAGGCCCTCTTACCAATTCAAGGGGGGAGAAATGGGGCTTCCGCA
GATGCCGCGCCAGCGGCGTCTTCACCGACTAGCTGGGGTAACACCCTCACGTGCT
ACTTCAAGGCCTTGGCATCCTGTAGGCTGCAAAGCTCCAGGACTGCACCCTGTTG
GTGTGTGGTGACGATCTTGTGGCCATCTGCGAAAGTGAGGAAACACACGAGGAC
GCGGCGAGCCTGAGAGCTCTCAACCGTCACTGAGAGAGACATAATGACCGAAGA
GTCCATTTACCAATCATGTGACTTGCAGCCTGAGGCGCGCGCGGCAATACGGTCA
CTCACCCAGCGCCTGTACTGTGGAGGCCCTATGTACAACAGCAAGGGGCAACAA
TGTGGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACCACTAGTATGGGCAACA
CCATAACGTGCTACGTTAAGGCTTTGGCCTCCTGTAGAGCTGCAAAGCTCCGGGA
CTGCACGCTCTTGGTGTGTGGTGATGATCTTGTGGCCATCTGCGAGAGTCAGGAG
ACACACGAGGACGCGGCGAGCCTGAGAGTCTCAACCGTCACTGAGAGAGACATC
CGTACCGACGAGTCTATGTACCAATCATGTGACCTGGACCCGCAGGCGCGCGCG
GCCATTACAAGTCACCTCACCCAAAAGGCTGTATGGTGGAGGCCCTCTTACCAAT
TCAAGGGGGGAGAAATGGGGCTTCCGCAGATGCCGCGCCAGCGGCGTCTTCACC
GACTAGCTGGGGTAACACCCTCACGTGCTACTTCAAGGCCTTGGCATCCTGTAGG
CTGCAAAGCTCCAGGACTGCACCCTGTTGGTGTGTGGTGACGATCTTGTGGCCAT
CTGCGAAAGTGAGGAAACACACGAGGACGCGGCGAGCCTGAGAGCTCTCAACC
GTCACTGAGAGAGACATAATGACCGAAGAGTCCATTTACCAATCATGTGACTTG
CAGCCTGAGGCGCGCGCGGCAATACGGTCACTCACCCAGCGCCTGTACTGTGGA
GGCCCTATGTACAACAGCAAGGGGCAACAATGTGGTTATCGTAGATGCCGCGCC
AGCGGCGTCTTCACCACTAGTATGGGCAACACCATAACGTGCTACGTTAAGGCTT
TGGCCTCCTGTAGAGCTGCAAAGCTCCGGGACTGCACGCTCTTGGTGTGTGGTGA
TGATCTTGTGGCCATCTGCGAGAGTCAGGAGACACACGAGGACGCGGCGAGCCT
GAGAG

>NGR-JS-P13
TCTCAACCGTCACTGAGAGAGACATAATGACCGAAGGGTCCATTTACCAATCAT
GTGACTTGCAGCCTGAGGCGCGCGCGGCAATACGGTCACTCACCCAACGCCTGT
ACTGTGGAGGCCCTATGTATAACAGCAAAGGGCAACAATGTGGTTATCGTAGAT
GCCGCGCCAGCGGCGTCTTCACCACTAGTATGGGCAACACCATAACGTGCTACGT
TAAGGCTTTGGCCTCCTGTAGAGCTGCAAAGCTCCGGGACTGCACGCTCTTGGTG

TGTGGTGATGATCTTGTGGCCATCTGCGAGAGCCAGGGGACACACGAGGACGCG
GCGAGCCTGAGAGCTCTCAACCGTCACTGAGAGAGACATCCGTACGGAGGAGGC
GATCTACCAATGTTGTGACCTGGACCCCCAGGCCCGCGTGGCCATCAAGTCCCTC
ACTGAAAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGC
GGCTACCGCAGGTGCCGCGCAAGCGGCGTACTGACGACTAGCTGTGGTAACACC
CTCACTTGCTATATCAAGGCCAAGGCAGCCTGTCGAGCCGCAGGGCTCCAGGAC
TGCACCATGCTCGTGTGTGGCGACGACCTAGTCGTTATCTGTGAAAGTGTGGGAG
TCCAGGAGGACGCGGCGAGACTGAGAGCATCTCAACCGTCACTGAGAGAGACAT
AATGACCGAAGGGTCCATTTACCAATCATGTGACTTGCAGCCTGAGGCGCGCGC
GGCAATACGGTCACTCACCCAACGCCTGTACTGTGGAGGCCCTATGTATAACAGC
AAAGGGCAACAATGTGGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACCACT
AGTATGGGCAACACCATAACGTGCTACGTTAAGGCTTTGGCCTCCTGTAGAGCTG
CAAAGCTCCGGGACTGCACGCTCTTGGTGTGTGGTGATGATCTTGTGGCCATCTG
CGAGAGCCAGGGGACACACGAGGACGCGGCGAGCCTGAGAGCTCTCAACCGTC
ACTGAGAGAGACATCCGTACGGAGGAGGCGATCTACCAATGTTGTGACCTGGAC
CCCCAGGCCCGCGTGGCCATCAAGTCCCTCACTGAAAGGCTTTATGTTGGGGGCC
CTCTTACCAATTCAAGGGGGGAGAACTGCGGCTACCGCAGGTGCCGCGCAAGCG
GCGTACTGACGACTAGCTGTGGTAACACCCTCACTTGCTATATCAAGGCCAAGGC
AGCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGTGTGTGGCGACGA
CCTAGTCGTTATCTGTGAAAGTGTGGGAGTCCAGGAGGACGCGGCGAGACTGAG
AGCA

>NGR-JS-P15a
TCTCAACCGTCACTGAGAGAGACATCCGTACGGAGGAGGCAATCTACCAATGTT
GTGACCTGGACCCCCAAGCCCGTGTGGCCATCAGGTCTCTCACCGAGAGGCTTTA
TGTCGGGGGCCCTCTTACCAACTCAAGGGGGGAAAACTGCGGCTATCGCAGGTG
CCGCGCAAGCGGCGTGCTGACAACTAGCTGTGGTAACACCCTCACTTGCTACATC
AAGGCCCGAGCAGCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGTG
TGTGGCGACGACTTGGTCGTTATCTGTGAAAGTCAGGGGGTCCAGGAGGACGCG
GCGAGCCTGAGAGGCTCTCAACCGTCACTGAGAGAGACATAATGACCGAAGGGT
CCATTTACCAATCATGTGACTTGCAGCCTGAGGCGCGCGCGGCAATACGGTCACT
CACCCAACGCCTGTACTGTGGAGGCCCTATGTATAACAGCAAAGGGCAACAATG
TGGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACCACTAGTATGGGCAACACC
ATAACGTGCTACGTTAAGGCTTTGGCCTCCTGTAGAGCTGCAAAGCTCCGGGACT
GCACGCTCTTGGTGTGTGGTGATGATCTTGTGGCCATCTGCGAGAGCCAGGGGAC
ACACGAGGACGCGGCGAGCCTGAGAGCTCTCAACCGTCACTGAGAGAGACATCC
GTACGGAGGAGGCAATCTACCAATGTTGTGACCTGGACCCCCAAGCCCGTGTGG
CCATCAGGTCTCTCACCGAGAGGCTTTATGTCGGGGGCCCTCTTACCAACTCAAG
GGGGGAAAACTGCGGCTATCGCAGGTGCCGCGCAAGCGGCGTGCTGACAACTAG
CTGTGGTAACACCCTCACTTGCTACATCAAGGCCCGAGCAGCCTGTCGAGCCGCA
GGGCTCCAGGACTGCACCATGCTCGTGTGTGGCGACGACTTGGTCGTTATCTGTG
AAAGTCAGGGGGTCCAGGAGGACGCGGCGAGCCTGAGAGGCTCTCAACCGTCAC
TGAGAGAGACATAATGACCGAAGGGTCCATTTACCAATCATGTGACTTGCAGCC
TGAGGCGCGCGCGGCAATACGGTCACTCACCCAACGCCTGTACTGTGGAGGCCC
TATGTATAACAGCAAAGGGCAACAATGTGGTTATCGTAGATGCCGCGCCAGCGG
CGTCTTCACCACTAGTATGGGCAACACCATAACGTGCTACGTTAAGGCTTTGGCC
TCCTGTAGAGCTGCAAAGCTCCGGGACTGCACGCTCTTGGTGTGTGGTGATGATC
TTGTGGCCATCTGCGAGAGCCAGGGGACACACGAGGACGCGGCGAGCCTGAGAG
C

>NGR-JS-P15b
CTCAACCGTCACTGAGAGAGACATCCGTGTTGAGGAGTCAATTTACCAATGTTGT
GACTTAGCCCCCGAAGCCAGACAGGCCATACGGTCGCTCACAGAGCGGCTTTAT
ATCGGGGGCCCCCTGACTAATTCAAAAGGGCAGAACTGCGGTTATCGCCGGTGC
CGCGCAAGCGGCGTGCTGACGACTAGCTGTGGCAATACCCTCACATGCTACTTGA
AGGCCGCTGCGGCCTGTCGAGCTGCAAAGCTCCAGGACTGCACGATGCTCGTGA
ACGGAGACGACCTTGTCGTTATCTGTGAGAGCGCGGGAACCCAAGAGGACGCGG
CAAGGCTGAGAGCAACCATCATGTGACTTGCAGCCTGAGGCGCGCGCGGCAATA
CGGTCACTCACCCAACGCCTGTACTGTGGAGGCCCTATGTATAACAGCAAAGGG
CAACAATGTGGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACCACTAGTATGG
GCAACACCATAACGTGCTACGTTAAGGCTTTGGCCTCCTGTAGAGCTGCAAAGCT
CCGGGACTGCACGCTCTTGGTGTGTGGTGATGATCTTGTGGCCATCTGCGAGAGC
CAGGGGACACTCAACCGTCACTGAGAGAGACATCCGTGTTGAGGAGTCAATTTA
CCAATGTTGTGACTTAGCCCCCGAAGCCAGACAGGCCATACGGTCGCTCACAGA
GCGGCTTTATATCGGGGGCCCCCTGACTAATTCAAAAGGGCAGAACTGCGGTTAT
CGCCGGTGCCGCGCAAGCGGCGTGCTGACGACTAGCTGTGGCAATACCCTCACA
TGCTACTTGAAGGCCGCTGCGGCCTGTCGAGCTGCAAAGCTCCAGGACTGCACG
ATGCTCGTGAACGGAGACGACCTTGTCGTTATCTGTGAGAGCGCGGGAACCCAA
GAGGACGCGGCAAGGCTGAGAGCAACCATCATGTGACTTGCAGCCTGAGGCGCG
CGCGGCAATACGGTCACTCACCCAACGCCTGTACTGTGGAGGCCCTATGTATAAC
AGCAAAGGGCAACAATGTGGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACC
ACTAGTATGGGCAACACCATAACGTGCTACGTTAAGGCTTTGGCCTCCTGTAGAG
CTGCAAAGCTCCGGGACTGCACGCTCTTGGTGTGTGGTGATGATCTTGTGGCCAT
CTGCGAGAGCCAGGGGACA

>NGR-JS-P16
TCTCAACCGTTCACTGAGAGAGACATCCGTACGGAGGAGGCAATCTACCAATGT
TGTGACCTTGACCCCCAAGCCCGTGTGGCCATCAAGTCCCTCACCGAGAGACTTT
ACGTTGGGGGCCCTCTTACCAACTCAAGGGGGGAGAACTGCGGCTACCGCAGGT
GCCGCGCGAGCGGCGTACTGACAACCAGCTGTGGTAACACCCTCACTTGCTACAT
CAAGGCCCAGGCAGCCTGTCGAGCCGCAGGGCTCCGGGACTGCACCATGCTCGT
GTGTGGCGACGACTTGGTCGTTATCTGTGAAAGTGCGGGGATCCAGGAGGACGC
GGCGAGCCTGAGAGCACTCTCAACCGTCACTGAGAGAGACATAATGACCGAAGA
GTCCATTTACCAATCATGTGACTTGCAGCCTGAGGCGCGCGCGGCAATACGGTCA
CTCACCCAACGCCTGTACTGTGGAGGCCCTATGTATAACAGCAAAGGGCAACAA
TGTGGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACCACTAGTATGGGCAACA
CCATAACGTGCTACGTTAAGGCTTTGGCCTCCTGTAGAGCTGCAAAGCTCCGGGA
CTGCACGCTCTTGGTGTGTGGTGATGATCTTGTGGCCATCTGCGAGAGCCAGGGG
ACACACGAGGACGCGGCGAGCCTGAGAGCATCTCAACCGTTCACTGAGAGAGAC
ATCCGTACGGAGGAGGCAATCTACCAATGTTGTGACCTTGACCCCCAAGCCCGTG
TGGCCATCAAGTCCCTCACCGAGAGACTTTACGTTGGGGGCCCTCTTACCAACTC
AAGGGGGGAGAACTGCGGCTACCGCAGGTGCCGCGCGAGCGGCGTACTGACAA
CCAGCTGTGGTAACACCCTCACTTGCTACATCAAGGCCCAGGCAGCCTGTCGAGC
CGCAGGGCTCCGGGACTGCACCATGCTCGTGTGTGGCGACGACTTGGTCGTTATC
TGTGAAAGTGCGGGGATCCAGGAGGACGCGGCGAGCCTGAGAGCACTCTCAACC
GTCACTGAGAGAGACATAATGACCGAAGAGTCCATTTACCAATCATGTGACTTG
CAGCCTGAGGCGCGCGCGGCAATACGGTCACTCACCCAACGCCTGTACTGTGGA
GGCCCTATGTATAACAGCAAAGGGCAACAATGTGGTTATCGTAGATGCCGCGCC
AGCGGCGTCTTCACCACTAGTATGGGCAACACCATAACGTGCTACGTTAAGGCTT
TGGCCTCCTGTAGAGCTGCAAAGCTCCGGGACTGCACGCTCTTGGTGTGTGGTGA

TGATCTTGTGGCCATCTGCGAGAGCCAGGGGACACACGAGGACGCGGCGAGCCT
GAGAGCA

>NGR-JS-P18
CCTTCTCAACCGTCACTGAGAGAGACATAAGAACAGAAGGGCAAGAATGCCAAG
GCTAGAACCTGCCCCAAGAGGCCAGAACTGCCATACACTCGCTCACTGAGAGAC
TTTACGTAGGAGGGCCCATGACAAACAGCAAAGGGCAATCCTGCGGTTACAGGC
GTTGCCGCGCAAGCGGCGTCTTCACCACCAGCATGGGGAATACCCTGACATGTTA
CATCAAAGCCCTTGGAGCGTACAAAGCTGCGGGGCCTGCGGACCCTGTTATGCT
GGTGTGTGGCGGAATCCTGGTCATCGCCTCCGCAAGCCAAGGCAGCAGGGAGGA
CGCGGCGAGCATAAGGGCATTTCTCAACCGTCACTGAGAGAGACATCATTACGG
AAGAGTCAATCTACCAGTCGTGTGGACTTGCAACCCCAGGCGCGTGTGGCCATC
AAGTCTCTCACCCAGAGGCTTTACGTCGGAGGCCCTATGACCAATTCAAGGGGG
GAAAACTGTGGTTATCGCAGATGCCGCGCGAGCGGCGTCTTCACCACTAGTCATG
GCAACACCCTCACTTGCTACATTAAGGCCCAAGCATCCTGTCGAGCTGCAAAGCT
CCGGGACTGCACCCTCCTGGTGTGTGGAGACGACCTTGTCGCCATCTGCGAGAGC
CAGGGGACACACGAGGACGCGGCGAGCCTGAGAGCACCTTCTCAACCGTCACTG
AGAGAGACATAAGAACAGAAGGGCAAGAATGCCAAGGCTAGAACCTGCCCCAA
GAGGCCAGAACTGCCATACACTCGCTCACTGAGAGACTTTACGTAGGAGGGCCC
ATGACAAACAGCAAAGGGCAATCCTGCGGTTACAGGCGTTGCCGCGCAAGCGGC
GTCTTCACCACCAGCATGGGGAATACCCTGACATGTTACATCAAAGCCCTTGGAG
CGTACAAAGCTGCGGGGCCTGCGGACCCTGTTATGCTGGTGTGTGGCGGAATCCT
GGTCATCGCCTCCGCAAGCCAAGGCAGCAGGGAGGACGCGGCGAGCATAAGGG
CATTTCTCAACCGTCACTGAGAGAGACATCATTACGGAAGAGTCAATCTACCAGT
CGTGTGGACTTGCAACCCCAGGCGCGTGTGGCCATCAAGTCTCTCACCCAGAGGC
TTTACGTCGGAGGCCCTATGACCAATTCAAGGGGGGAAAACTGTGGTTATCGCA
GATGCCGCGCGAGCGGCGTCTTCACCACTAGTCATGGCAACACCCTCACTTGCTA
CATTAAGGCCCAAGCATCCTGTCGAGCTGCAAAGCTCCGGGACTGCACCCTCCTG
GTGTGTGGAGACGACCTTGTCGCCATCTGCGAGAGCCAGGGGACACACGAGGAC
GCGGCGAGCCTGAGAGCA

>NGR-JS-P21
CTCTCAACCGTCACTGAGAGAGACATCCGCATTGAGGAGTCAATCTACCAATGTT
GCGACTTGGCCCCCGAGGCCAGACAGGCTATAAGGTCGCTCACAGAGCGGCTTT
ATATCGGGGGTCCCCTGACTAATTCAAAAGGGCAGAACTGCGGTTATCGCCGGT
GCCGCGCAAGCGGCGTGCTGACGACTAGCTGCGGTAATACCCTCACATGTTACTT
GAAAGCCACTGCGGCCTGTCGAGCTGCAAAGCTCCAGGACTGCACAATGCTCGT
GAACGGGGACGACCTTGTCGTCATATGTGAGAGCGCGGGAACCCAAGAGGACGC
GGCGAGCCCTGAGACGGTCTCAACCGTCACTGAGAGAGACATAATTACGGAAGA
GACCATTTACCAATCATGTGACTTGCAGCCTGCAGGCGCGCGCGGCCATTACAGG
TCACTCACCCAACGCCTGTACTGTGGAGGCCCTATGTACAACAGCAAGGGGCAA
CAATGTGGTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACCACTAGTATGGGCA
ACACCATAACGTGCTACGTTAAGGCTTTGGCCTCCTGTAGAGCTGCAAAGCTCCG
GGACTGCACGCTCTTGGTGTGTGGTGATGATCTTGTGGTCATCTGCGGAAGTCAG
GAAACCCACGAGGACGCGGCGAGCCTGAGAGCGCCATCAGGCTCGCCGTGTCCT
CTCAATGGTCTGGTATTACAGGTTTGCATTCCTTCGCCCAATCTGCTTATCTTACT
ATATCACTGCGCGTCCTTGCTCAATGAGAGTCTTTGNTTACATCAAAAGCAAGAG
AGTTTTTTCGGTGATGCTTTATTACTTGCATATCATCCCACTTGGGAGAATATGCA
GATCATAACTAGCATCACAGAAGCTGAAAATGACTGTGATCAAAGACAATATTC

TTAACAAGCACGCTCACTGGCTTAGTAAGATAAGCCCTCTGGGCACAGTGGCTCA
AACCTGTAATCCCGGGCCTTTGGGAGGACACGGCGAGCCTGAGAGCACTCTCAA
CCGTCACTGAGAGAGACATCCGCATTGAGGAGTCAATCTACCAATGTTGCGACTT
GGCCCCCGAGGCCAGACAGGCTATAAGGTCGCTCACAGAGCGGCTTTATATCGG
GGGTCCCCTGACTAATTCAAAAGGGCAGAACTGCGGTTATCGCCGGTGCCGCGC
AAGCGGCGTGCTGACGACTAGCTGCGGTAATACCCTCACATGTTACTTGAAAGCC
ACTGCGGCCTGTCGAGCTGCAAAGCTCCAGGACTGCACAATGCTCGTGAACGGG
GACGACCTTGTCGTCATATGTGAGAGCGCGGGAACCCAAGAGGACGCGGCGAGC
CCTGAGACGGTCTCAACCGTCACTGAGAGAGACATAATTACGGAAGAGACCATT
TACCAATCATGTGACTTGCAGCCTGCAGGCGCGCGCGGCCATTACAGGTCACTCA
CCCAACGCCTGTACTGTGGAGGCCCTATGTACAACAGCAAGGGGCAACAATGTG
GTTATCGTAGATGCCGCGCCAGCGGCGTCTTCACCACTAGTATGGGCAACACCAT
AACGTGCTACGTTAAGGCTTTGGCCTCCTGTAGAGCTGCAAAGCTCCGGGACTGC
ACGCTCTTGGTGTGTGGTGATGATCTTGTGGTCATCTGCGGAAGTCAGGAAACCC
ACGAGGACGCGGCGAGCCTGAGAGCGCCATCAGGCTCGCCGTGTCCTCTCAATG
GTCTGGTATTACAGGTTTGCATTCCTTCGCCCAATCTGCTTATCTTACTATATCAC
TGCGCGTCCTTGCTCAATGAGAGTCTTTGNTTACATCAAAAGCAAGAGAGTTTTT
TCGGTGATGCTTTATTACTTGCATATCATCCCACTTGGGAGAATATGCAGATCAT
AACTAGCATCACAGAAGCTGAAAATGACTGTGATCAAAGACAATATTCTTAACA
AGCACGCTCACTGGCTTAGTAAGATAAGCCCTCTGGGCACAGTGGCTCAAACCT
GTAATCCCGGGCCTTTGGGAGGACACGGCGAGCCTGAGAGCA

Figure 4.8: HCV NS5B Gene Sequences of all study Isolates

Table 4.5: HCV NS5B Protein Sequences in MEGA7.0

```
1         10         20         30         40         50         60         70         80         90
   TNVVTWTPKPAWPSSPSLRGFMLGALLPIQGGKTAATAGAARAAY*QLAVVTPSLATSRPGQPVEPQGSRTAPCSCVATT*SLSVKVRG
BD2TNVVTWTRKPAWPSSPKLRGFMLGALLPIQGRKTAATAGAQRAAYRQLAVVTPSLITSRRGQPVEPQGSRTAPCSCVATT*SFMDKVRG
BD7TNHVTYNLRRARQYGHSPNACTVEALCITARGSNVVIVDAAPAASSLPVWATP*RVTLRL*PPVELQSSGIVRSWVVVTTLWPSARAKG
P02TNHVTCSLRRARQYGHSPNACTVEALCITAKGNNVVIVDAAPAASSPLVWATP*RATLRLWPPVELQSSGTARSWCVVMILARCLQ?KG
P09TNHVTCTLRRARQYGHSPNACILEALCITARGNNVVIVDAAPAASSPLVWATP*RATLRLWPPVELQSSGTARSWCVVMISWPSARAKG
P09TNHVTCSLRRARQYGHSPNACTVEALCITARGNNVVIVDAAPAASSPLVWATP*RATLRLWPPVELQSSGTARSWCVVMILWPSARAKG
P15TNHVTCSLRRARQYGHSPNACTVEALCITAKGNNVVIVDAAPAASSPLVWATP*RATLRLWPPVELQSSGTARSWCVVMILWPSARAKG
BD6TSVVTWTPKPAWPSSLSLRGFMLGALLPIQGGKTAATAGAARAAY*LLAVVTPSLTSRREGQPVEPQGSRTAPCSCVATT*SLSVKVR
BD9TNHVTYNLRRARQYGHSPNACTVEALCITARGSNVVIVDAAPAASSLPVWATP*RVTLRL*PPVELQSSGIVRSWCVVTTLWPSARAKG
D10?PRETYNLRRARQCGHSPNACTVEALCITARGSNVVIVDAAPAASSLPVWATP*RVTLRL*PPVELRSSGIARSWCVVTTLWPSARAKG
D12TNVVTMTVKPAWLSPPSLRSFMLGALLPIQVGITAATAGAARAAY*QLAVVTPSLATRRPGQPVEPGGSRTAPCSCVATT*ASLSE---
D13NA?ETWTPKPAWPSSLSLRGFMLGALLPIQGGKTAATAGAARAAY*LLAVVTPSLATSRPEQPVEPQGSGTAPCSCVATT*SLSAKVRG
D20TNVVTLTPKPVWPSSPSPRDETLGALLPTQGGRTAATAGAARAAY*QPAVVTPSLATSRPRQPVEPQGSGTAPCSCVATTWSLSVKVRG
D23TNVVTWTPKPAWPSSLSLRGFMLGALLPIQGGKTAAIAGAARAAY*LLAVVTPSLATSRPEQPVEPQGSGTAPCS?VATT*SLSAKVRG
D28TNDVTWTPKPAWPSSPSLRGFMLGALLPIQGGRTAATAGAARAAY*LLAVVTPSLATSRPGQPVEPQGSRTAPCSCVATT*SLSVKVRG
D30*MTL?RPR?PRLPYTP*LRDYMWEAP*LTPGGNPAATGGAARAAECQPPVGTPSLATS*PGQRVEKQRARMAPCSWYATT*SLSVQRRG
BD1TNHVTCSLRRARQYGHSPNACTVEALCTTARGNNVVIVDAAPAASSPLVWATP*RATLRR*PPVELQSSGTARSWCVVMILWPSARVRK
BD3TNVVTWTPKPVWPSGLSPRGFMSGALLPTQGGKTAAIAGAAQAAC*QLAVVTPSLATSRPEQPVEPQQSRTAPCSCVATTWSLSVKVRG
LD2THCLTWNPRRARHYSHSPKGCTVEAL?TTSKGDNVVIRNGAPAASSPTIWGNPLRATLRL*PPV?L?SSRTA?SWCVVATLWPSAKVRK
P03TSVVTWTPKPVWPSSLSPRGFMSGAPLPIQGGKTAAIAGAARAAY*QLAVVTPSLATLRPKQPVEPQGSGRTAPCSCVAT*SLSVKVRE
P07 TNHVTCTLRRARQYGHSPNACILEALCITARGNNVVIVDAAPAASSPLVWATPRATLRLWPPVELQSSGTARSWCVVMISWPSARARG
P10TNHVTCSLRRARQYGHSPNACTVEALCTTARGNNVVIVDAAPAASSPLVWATP*RATLRL*PPVELQSSGTARSWCVVMILWPSARVRR
P13TNVVTWTPRPAWPSSPSLKGFMLGALLPIQGGRTAATAGAAQAAY*RLAVVTPSLAISRPRQPVEPQGSRTAPCSCVATT*SLSVKVWE
P15T?HVTCSLRRARQYGHSPNACTVEALCITAKGNNVVIVDAAPAASSPLVWATP*RATLRLWPPVELQSSGTARSWCVVMILWPSARARG
P16TNHVTCSLRRARQYGHSPNACTVEALCITAKGNNVVIVDAAPAASSPLVWATP*RATLRLWPPVELQSSGTARSWCVVMILWPSARARG
LD5T?VVT*PPKPDRPYGRSQSGFISGAP*LIQKGRAAVIAGAAQAACQRLAVAIPSHAT*RPLRPVELQSSRTARCS*TETTLSLSVRARE
P18TSRVTCNPRPVWPGHPPLRCTAEGACTTPRGGKVVIVDGAP?ASSPLVCAVTPSLALSL*PPVELQSSGTARSWSCEATLWPLAVAVE-
P21CHQARRVLSMVGITGLHSFAQSAYLTISLRVLAQ*ESL?TSKAREFFR*CFITCISSHLGEYADHN*HHSS*K*L*SKTIFLTSLTGL
P06TNHVVCSLRRARQYGHPPNACTVEALCITAKGNNVVIVDAAPAASSPLVWATP*RATLRLWPPVELQSSGTARSWCVVMILWPSARARG
P09TNHVVCSLRRARQYGHPPNACTVEALCITARGNNVVIVDAAPAASSPLVWATP*RATLRLWPPVELQSSGTARSWCVVMISWPSARARG
LD1TNHVTCSLRRPRQYGHSPNACTVEALCITAKGNNVVIVDGAPRASSP*VWAVT*RSLLTLWPQPVLPSGGRTRSWSCVMILWPSARARG

LD3TNHVTCSLRRPRQYGHPPNACTVEALCTTARKNNVVIVDAAPAASSPLVWATP*RATLRL*PPVELQSSGTARSWCVVMILWPVARGR

LD4TNVVTWTPKPVWPSGLSPRGFMSGALLPTQGGKTAAIGAARQAAC*QLAVVTPSLATSRPGEPVEPQGSRWAPCSCVAT

LD6TNVVTWTPRPAWPSSPSLKGFMLGALLPIQGGRAATAGAARAQAY*QRAVVTPSLATIRPGRPVEPQGSRTAPCSCVATT*SLSVKVRWE

LD7A?-?ARTCPKRPE?AI?LAH*ETLRRRAHDKQQRAILRLQALPRKRRLH?PAWGIP*HVTSKPLERT?SCGACGPCY?-----------

L10TN-??ATWPPRPD?AI?VAHRAALYRGSPD*FKRAELRLSPVPRKRRAD?LAAV-IPSHVT*KPLRPV?SCKAPGLHNA----------

L0bTN???ATWPPRPD?AI?VAHRAALYRGSPD*FKRAELRLSPVPRKRRAD?LAAV-IPSHVT*KPLRPV?SCKAPGLHNA----------

LD12I?-?LVPCPKRPE?AI?LAY*ETLRRRAHDKQQRAILRLQALPRKRRLH?PAWG-IP*HVTSKPLQRA?SCGDRGPCYA?---------

LD13I?-?LVPCPKRPE?AI?LAH*ETLRRRAHDKQQRAILRLQALPRKRRLH?PAWG-IP*HVTSKPLQRA?SCG-----------------

LD15aTN-??VTWPPKPA?VV?VSHREALRRGPPHQFQGTKLRLSQMPRKRRPD?LAAV-TPLLAISKLLLPV?SCKAPELHHA---------
LD15bTN-??VTWPPKPA?VV?VSHREALRRGPPHQFQGTKLRLSQMPRKRRPD?LAAV-TPLLAISKLLLPV?SCKAPELH?----------
LD16TN??Q*TCSLRRA?AI?VTHPTPVLWRPYV*QQGATMWLS*MPRQRRLH?LVWA-TL*RATSRL*PPV?SCKAPGLHA?----------
16bTN-??*TCSLRRA?AI?VTHPTPVLWRPYV*QQGATMWLS*MPRQRRLH?LVWA-TL*RATSRL*PPV?SCKAPGLHAP?---------
L20I?-?LVPCPKRPE?AI?LAH*ETLRRRAHDKQQRAILRLQALPRKRRLH?PAWG-IP*HVTSKPLQRA?SC?-----------------
```

Key*: stop codons; -: deletions; ---: end of sequence; ?:gap in sequence

Table 4.6: Amino acid substitutions in all HCV sequences in comparison to HCV Prototype Reference strain H77 indicating conserved sites (with red dots) and variable sites (letters)

```
   1        10        20        30        40        50        60        70        80        89
   I        I         I         I         I         I         I         I         I         I
H77TNVVTWTPKPAWPSSPSLRGFMLGALLPIQGGKTAATAGAARAAY*QLAVVTPSLATSRPGQPVEPQGSRTAPCSCVATT*SLSVKVRG
BD1..H..CSLRR.RQYGH.PNACTVE..CTTAR.NNVVIVD..P..SSP.VWA..*R..L.L*P...L.S.G..RSW..VMILWP.AR..K
BD2.........R.........K.............R........Q....R.........I...R.......................FMD---
BD3...........V...GL.P....S.....T.......I....Q..C..............E...................W........
LD2.HCL..N.RR.RHY.H.PK.CTVE..?TTSK.DNVVIRN..P..SSPTIWGN.LR..L.L*P..?L?S....?SW..V..LWP.A...K
BD6.S.............L.............................L..............E........G..............A....
BD7..H..YNLRR.RQYGH.PNACTVE..CITAR.SNVVIVD..P..SSLPVWA..*RV.L.L*P...L.S.GIVRSW..V..LWP.ARAK.
BD9..H..YNLRR.RQYGH.PNACTVE..CITAR.SNVVIVD..P..SSLPVWA..*RV.L.L*P...L.S.GIVRSW..V..LWP.ARAK.
D10?PRE.YNLRR.RQCGH.PNACTVE..CITAR.SNVVIVD..P..SSLPVWA..*RV.L.L*P...LRS.GI.RSW..V..LWP.ARAK.
D12.....M.V....L.P....S.............I...  ...................R.........G..............A...E..
D13NA?E...........L.............................L..............E........G..............A....
D20.....L....V......P.D.T......T...R..............P.............R........G..........W........
D23.................L.....................I.........L..............E........G.....?........A....
D28..D..............................R.............L.....................................
D30*MTL.RPR?.RL.YT.*..DY.WE.P*LTP..NP...G.....EC.PP.G.......*...R..K.RA.M.P..WYA.......QR...
LD1..H..CSLRR.RQYGH.PNACTVE..CITAK.NNVVIVD..P..SSP.VWA..*R..L.LWP...L.S.G..RSW..VMILWP.ARA..

P02..H..CSLRR.RQYGH.PNACTVE..CITAK.NNVVIVD..P..SSP.VWA..*R..L.LWP...L.S.G..RSW..VMILARCLQ?-

P21CHQARRVLSMVGITGLHSFAQSAYLTISLRVLAQ*ESL?TSKAREFFR*CFITCISSHLGEYADHN*HH.S*K*L*SK.IFLT.LLTGL

LD3..H..CSLRR.RQYGH.PNACTVE..CTTAR.NNVVIVD..P..SSP.VWA..*R..L.L*P...L.S.G..RSW..VMILWP.AR..R

P06..H..CSLRR.RQYGH.PNACTVE..CITAK.NNVVIVD..P..SSP.VWA..*R..L.LWP...L.S.G..RSW..VMILWP.ARA..

P09..H..CSLRR.RQYGH.PNACTVE..CITAR.NNVVIVD..P..SSP.VWA..*R..L.LWP...L.S.G..RSW..VMISWP.ARA..
P07..H..C.LRR.RQYGH.PNACI.E..CITAR.NNVVIVD..P..SSP.VWA..*R..L.LWP...L.S.G..RSW..VMISWP.ARA..
P09..H..CSLRR.RQYGH.PNACTVE..CITAR.NNVVIVD..P..SSP.VWA..*R..L.LWP...L.S.G..RSW..VMILWP.ARA..
LD4...........V...GL.P....S.....T.......I....Q..C..............E...................W........
P15..H..CSLRR.RQYGH.PNACTVE..CITAK.NNVVIVD..P..SSP.VWA..*R..L.LWP...L.S.G..RSW..VMILWP.ARA..
P03.S.........V....L.P....S..P..........I.....................L..K........G.................E
P10..H..CSLRR.RQYGH.PNACTVE..CTTAR.NNVVIVD..P..SSP.VWA..*R..L.L*P...L.S.G..RSW..VMILWP.AR..R
P13.........R.........K.............R........Q....R.........I...R.........................WE
P15.?H..CSLRR.RQYGH.PNACTVE..CITAK.NNVVIVD..P..SSP.VWA..*R..L.LWP...L.S.G..RSW..VMILWP.ARA..
P16..H..CSLRR.RQYGH.PNACTVE..CITAK.NNVVIVD..P..SSP.VWA..*R..L.LWP...L.S.G..RSW..VMILWP.ARA..
LD5.?...*P...DR.YGR.QS..IS..P*L..K.R..VI....Q..C.R...AI..H..*..LR...L.S....R..*TE..L....RA.E
P18.SR..CN.R.V...GH.P..CTAE..CTTPR...VVIVD..P?.SSP.VCA......L.L*P...L.S.G..RSW..E..LWP.A.A.E
LD6.........R.........K.............R........Q....R.........I...R.........................WE
LD7A...AR.CPKRPE.AI.lAH*ETLRRRAHDKQQR.ILRLQ.LPRKRR.H?PAWG-IP*HVTSKPLERT.SCGA.GPCY?----------
L10...A..P.R.D?AI?VAH.AALYRGSPD*FKRAELRLSPVP.KRRAD?LAAV-IPSHVT*KPLRPV?SCK.PGLHNA?----------
L12I?-?LVPCPKRPE.AI?.AY*ETLRRRAHDKQQR.ILRLQ.LPRKRR.H?PAWG-IP*HVTSKPLQRA?SCGDRGPCYA?---------

L0b...A..P.R.D?AI?VAH.AALYRGSPD*FKRAELRLSPVP.KRRAD?LAAV-IPSHVT*KPLRPV?SCK.PGLHNA?----------

L6b.H..Q*.CSLRRA.AI.VTHPTPVLWR.YV*QQG.TMWLS*MPRQRR.H?L.WA-.L*.ATSRL*.PV.SCKAPGLHA?----------

L5b.?...PPKPA?VV.V.HREALRR.PPHQF..T.LRLSQMP.KRRPD.L.A.-TP.LAISKLLLPV?SCKAPELHHA----------

L13I?-?LVPCPKRPE.AI?.AH*ETLRRRAHDKQQR.ILRLQ.LPRKRR.H?PAWG-IP*HVTSKPLQRA?SCG----------------

L15.?...PPKPA?VV.V.HREALRR.PPHQF..T.LRLSQMP.KRRPD.L.A.-TP.LAISKLLLPV?SCKAPELHHA------------

L16.H..Q*.CSLRRA.AI.VTHPTPVLWR.YV*QQG.TMWLS*MPRQRR.H?L.WA-.L*.ATSRL*.PV.SCKAPGLHA?----------

L20I?-.LVPCPKRPE.AI..AH*ETLRRRAHDKQQR.ILRLQ.LPRKRR.H?PAWG-IP*HVTSKPLQRA.SC?----------------
```

Table 4.7: Amino acid alignment of study HCV sequences in comparison with reference sequences from other countries in MEGA 7.0.

```
                  10        20        30        40        50        60        70        80        89
                   I         I         I         I         I         I         I         I         I
AF009606.1   TNVVTWTPKPA-PSSPSLRGFMLGALLPIQGGKTAATAGAAR-A-*QLAVVTPSLATSRPGQP-EPQGSRTA-CSC--TT*S-SVKV-G
EU849145.1   CLLLSIQNQPKWR-TRWPLVLTTTVPMAVL?YEADDVIRTHPAVYLVFRTAINPRAGPQ*HL?VAVRYVGATPLRYVAM?DLLWARPRC
EF442256.1   MNG-AWVWGGALLRCPR-F*YKAGAHWGI-GAWPITPAGAT-SSW-C-QGSMPQQPPSVAEQPTMPEG--LLAFYCGRQAK-QLIT-M?
FJ791092|C   ??VVTW?L-?AR*Y?ASL?YS?W??LYIT?EETCAG?ADA????CT?QASGTH?--SSRPQRPPKPQA?RTA-CWC-VTT?S-SPKATV
GU088166.1   ISVVTWSPK-VRQS?A*QRDSMWA--CITARETCAGIADAAQ-ASSPPASGTH*--ISKPQPLSKR-G*ETA-CWSAV-TWSSSP-ATA
Y12083_6a_   TNVVTWTPK-A-PSSPSLRGFMLGALLPIQGGKTAATAGAAR-A-*QLAVVTPSL-TSRPGQP-EPQGSRTA-CSC--TT*S-SVKV-G
FN401108_4   -?VVT*SPK-ARLL?ASQRDSMWAALCITAEETCVGIGDAAQ-ASSRPASGTH*--ISKPQPLPKR-G*ETA-CWFAVTTWSSSL-VMA
JQ272969_A   TNVVTWTPK-A-PSSPS-RGFMLGALLPIQGGRTAATAGAA--A-*RLAVVTPSL-TSRPKQP-EPQGSRTA-CSC--TT*S-SVKV-G
EF470324_B   TNVVTWTPK-A-PSSPSLRGFMSGALLPIQGGRTAATAGAAR-A-*QPAVVTPSL-TSRPRQP-EPQGSRTA-CS---TT*S-SVKV-E
JF511093_A   -IV--CPRRLELLY-H*LRGY-WEA-*QTARGNPAGTGVAAR-ECSPPAW-THS--T*KPRRR-TPQALLPP-CW-A-TT*-SSQ-AKE
JQ272964|_   TNVVTWTPK-A-PSSPSLRGFMLGALLPIQEGRTAATAGAA--A-*QPAVVTPSL-ISRPKQP-EPQGSRTA-CSC--TT*S-SV-VQG
JQ318499|_   TNVVTWTPK-A-PSSPSLRGFMLGALLPIQGGRTAATAGAAR-A-*RLAAVTPSL-TSRPRQP-EPQGS?TA-CSC--TT*S-S--V-G
JQ679054|_   TNVVTWPPK-VGS*SLS-RGS-*GA-SPIPRDRTAATAGAAQ-AS*PLAA--P*PVISKLLRP-ELQSSRTA-CWCA-TIWS?------
JQ679065|N   TNVVTWPPR-A-P*SPSLRGFISG-H*PIPRGK-AAIADAVQ-A-*RLAA--PSP-TLKPVRP-EL-SSRTA-CWCA-TT*S?------
JQ679058|N   TNAVTWPLK-ARS*SPS-RGF-*GA-SPTPKDKTAAIAGAAQ-ES*LLAVTPSP-ISKLPRP-ELQSSRTA-CWCA-TTWS?-------
JQ679055|N   TNVVTWPPK-VRS*GLS-RGS-*GA-SPTPRDKTAAIAGAAQ-AS*PLAAV-P*PVISKLLRP-ELQSSRTA-CWCA-TIWS?------
JQ318533|C   TSVVTWTPK-V-PSSPS-RGFMSGA-LPIQGGKTAAIAGAAR-A-*QLAVVTPSL-TLRPKQP-EPQGS-TA-CSC--TT*S-SVKV-E
KC351707|I   TSVVTWTPK-A-PSSPS-RGFMSGALLPIQGGKTAAIAGAAR-A-*QLAVVTPSL-TLRPKQP-EPQGS-TA-CSC--TT*S-SVKV-E
KC351693|_   TNVVTWTPK-A-PSSPS-RGFMSGA-LPIQGGKTAATAGAAR-A-*QLAVVTPSL-TLRPKQP-EPQGS-TA-CSC--TT*S-SV-V-E
JX677480|C   TNHVTCSLRRARQYGHS-SAC-VEALCTTARGNNVVIVDAA---SSPLVW-TP*--ILRL*PP-ELQSS-TARSWC-VTTLW?------
DQ333626|B   TNHVTCSLRRARQYGHS-NAC-VEA-CTTARGNNVVIVDAA--ASSPLVW-TP*--TLRL*PP-ELQSS-IARSWC-V?----------
AB204665|_   TNHVTCSPRRARQYG-S-NAC-VEALCITARGNNVVIVDAA--ASSPLVW-TP*--TLRLWPP-ELQSS-TARSWC-V-TLWPS--A-G
DQ333681|S   TNHVTCSLRRARQYGHS-NAY-VEALCITARGNNVVIVDAA--ASSPLVW-TP*--TLRLWPP-ELQSS-IARSWC-V?----------
DQ333672|S   TNHVTCSLRRARQYGHS-NAC-AEALCITARASNVVIADAA--ASSPLVW-TP*--TLRL*PP-ELQSSRTARSWC-V?----------
AB204665|S   TNHVTCSPRRARQYG-S-NAC-VEALCITARGNNVVIVDAA--ASSPLVW-TP*--TLRLWPP-ELQSS-TARSWC-V-TLWPS--A-G
GQ490838|B   TNVVTWPPK-DRP*G-SQSGFISGA-*LIQKGK-AVIAGAAR-AC*RLAAV-PSHVT*RPLRP-ELQSSRTARCSCAETTLS-SV----
CS101285|G   ISAVTLTLRLAP*SPPSLRDC-WVA-CST-GARRLATGGAE-V--SPLAWGTP*--ISRPWQPARL?GLVGAISWC---TW-SS--AGE
GQ866993|_   TNVVTWPPK-DRP*G-SQSGFMWAAR*PTAKARTAAIVVAVR-AC*RLAAVTPSHVT*RPLRP-ELQSSRTA-CW?------------
GQ153856|G   -I?--YLRK-HSLRDC-WGA-C*TAR?NPA-TGDAA--ASSPQA-----WETPS--MLKPSRPARPLA?I-L-CWC-G-TW?SSQ-?-E
AB204674|_   TNHVTCSLRRAQQYGHS-NAC-AEALCITARGNNVVIVDAA--ASSPLVW-TP*P-TLRLWPP-ELQSS-TARSWC-V-ILWPS--A-G
FN401088.P   -?VVTWSPKLARLF?ASRKDSMWAALCITARETCVGTGDAAQ-ASSRPASGTH*--ISKPQ-LSKRQG*ETA-CWFAVTTWSSSL-VTA
NGR-JS-BD2   TNVVTWTPR-A-PSSPSLKGFMLGALLPIQGGRTAATAGAAQ-A-*RLAVVTPSL-ISRPRQP-EPQGSRTA-CSC--TT*S-FMD---
NGR-JS-BD7   TNHVTYNLRRARQYGHS-NAC-VEALCITARGSNVVIVDAA--ASSLPVW-TP*-VTLRL*PP-ELQSS-IVRSWC-VTTLWPS--AKG
NGR-JS-P09   TNHVTCTLRRARQYGHS-NACILEALCITARGNNVVIVDAA--ASSPLVW-TP*--TLRLWPP-ELQSS-TARSWC-V-ISWPS--A-G
NGR-JS-P02   TNHVTCSLRRARQYGHS-NAC-VEALCITAKGNNVVIVDAA--ASSPLVW-TP*--TLRLWPP-ELQSS-TARSWC-V-ILARCLQ?--
NGR-JS-P09   TNHVTCSLRRARQYGHS-NAC-VEALCITARGNNVVIVDAA--ASSPLVW-TP*--TLRLWPP-ELQSS-TARSWC-V-ILWPS--A-G
NGR-JS-P15   TNHVTCSLRRARQYGHS-NAC-VEALCITAKGNNVVIVDAA--ASSPLVW-TP*--TLRLWPP-ELQSS-TARSWC-V-ILWPS--A-G
NGR-JS-BD6   TSVVTWTPK-A-PSSLSLRGFMLGALLPIQGGKTAATAGAAR-A-*LLAVVTPSL-TSRPEQP-EPQGS-TA-CSC--TT*S-S-KV-G
NGR-JS-BD9   TNHVTYNLRRARQYGHS-NAC-VEALCITARGSNVVIVDAA--ASSLPVW-TP*-VTLRL*PP-ELQSS-IVRSWC-VTTLWPS--AKG
NGR-JS-D10   ?PRETYNLRRARQCGHS-NAC-VEALCITARGSNVVIVDAA--ASSLPVW-TP*-VTLRL*PP-EL-SS-IARSWC-VTTLWPS--AKG
NGR-JS-D12   SA-VTWTPM-V-PSSLS-RGFMSGALLPIQGGKTVAIAGAAR-A-*QLAVVTPSL-TSRPRQP-EPQGS-TA-CSC--TT*S-S-KV-E
NGR-JS-D13   NA?ETWTPK-A-PSSLSLRGFMLGALLPIQGGKTAATAGAAR-A-*LLAVVTPSL-TSRPEQP-EPQGS-TA-CSC--TT*S-S-KV-G
NGR-JS-D20   TNVVTLTPK-V-PSSPS-RDF-LGALLPTQGGRTAATAGAAR-A-*QPAVVTPSL-TSRPRQP-EPQGS-TA-CSC--TTWS-SVKV-G
NGR-JS-D23   TNVVTWTPK-A-PSSLSLRGFMLGALLPIQGGKTAAIAGAAR-A-*LLAVVTPSL-TSRPEQP-EPQGS-TA-CS?--TT*S-S-KV-G
NGR-JS-D28   TNDVTWTPK-A-PSSPSLRGFMLGALLPIQGGRTAATAGAAR-A-*LLAVVTPSL-TSRPGQP-EPQGSRTA-CSC--TT*S-SVKV-G
NGR-JS-D30   *MT-?RPR?-RLPY-P*LRDYMWEA-*LTPGGNPAATGGAAR-EC*PPAGVTPSL-T*RPGRP-KP-ASMTP-CW-A-TT*S-SQ-V-G
NGR-JS-BD1   TNHVTCSLRRARQYGHS-NAC-VEALCTTARGNNVVIVDAA--ASSPLVW-TP*--TLRL*PP-ELQSS-TARSWC-V-ILWPS--V-K
NGR-JS-BD3   TNVVTWTPK-V-PSGLS-RGFMSGALLPTQGGKTAAIAGAAQ-AC*QLAVVTPSL-TSRPEQP-EPQGSRTA-CSC--TTWS-SVKV-G
NGR-JS-LD2   THC-TWNPRRARHYSHS-KGC-VEAL?TTSKGDNVVIRNAA--ASSPTWGNPL--TLRL*PP-?L?SSRTA?SWC-VT TLWPS-KV-K
NGR-JS-P03   TSVVTWTPK-V-PSSLS-RGFMSGA-LPIQGGKTAAIAGAAR-A-*QLAVVTPSL-TLRPKQP-EPQGS-TA-CSC--TT*S-SVKV-E
NGR-JS-P10   TNHVTCSLRRARQYGHS-NAC-VEALCTTARGNNVVIVDAA--ASSPLVW-TP*--TLRL*PP-ELQSS-TARSWC-V-ILWPS--V-R
NGR-JS-P13   TNVVTWTPR-A-PSSPSLKGFMLGALLPIQGGRTAATAGAAQ-A-*RLAVVTPSL-ISRPRQP-EPQGSRTA-CSC--TT*S-SVKVWE
NGR-JS-P15   T?HVTCSLRRARQYGHS-NAC-VEALCITAKGNNVVIVDAA--ASSPLVW-TP*--TLRLWPP-ELQSS-TARSWC-V-ILWPS--A-G
NGR-JS-P16   TNHVTCSLRRARQYGHS-NAC-VEALCITAKGNNVVIVDAA--ASSPLVW-TP*--TLRLWPP-ELQSS-TARSWC-V-ILWPS--A-G
NGR-JS-LD5   T?VVT*PPK-DRPYG-SQSGFISGA-*LIQKGRTAVIAGAAQ-AC*RLAV--PSH-T*RPLRP-ELQSSRTARCS*TETTLS-SV-A-E
NGR-JS-P18   TSRVTCNPR-V-PSGHS-RGC-AEALCTTPRGKTVVIVDAA-?ASSPLVC-TPSL-TLRL*PP-ELQSS-TARSWC-ETTLWPS-KA-E
NGR-JS-P06   TNHVTCSLRRARQYGHS-NAC-VEALCITAKGNNVVIVDAA--ASSPLVW-TP*--TLRLWPP-ELQSS-TARSWC-V-ILWPS--A-G
NGR-JS-P09   TNHVTCSLRRARQYGHS-NAC-VEALCITARGNNVVIVDAA--ASSPLVW-TP*--TLRLWPP-ELQSS-TARSWC-V-ISWPS--A-G
NGR-JS-LD7   A?-?ARTCPKRPE?AI?LAH*ETLRRRAHDKQQRAILRLQALPRKRRLH?PAWG-IP*HVTSKPLERT?SCGACGPCY?----------

NGR-JS-LD10  TN-??ATWPPRPD?AI?VAHRAALYRGSPD*FKRAELRLSPVPRKRRAD?LAAV-IPSHVT*KPLRPV?SCKAPGLHNA?---------
NGR-JS-L10b  TN???ATWPPRPD?AI?VAHRAALYRGSPD*FKRAELRLSPVPRKRRAD?LAAV-IPSHVT*KPLRPV?SCKAPGLHNA----------
NGR-JS-LD12  I?-?LVPCPKRPE?AI?LAY*ETLRRRAHDKQQRAILRLQALPRKRRLH?PAWG-IP*HVTSKPLQRA?SCGDRGPCYA?---------
NGR-JS-LD13  I?-?LVPCPKRPE?AI?LAH*ETLRRRAHDKQQRAILRLQALPRKRRLH?PAWG-IP*HVTSKPLQRA?SCG----------------
NGR-JS-LD15  TN-??VTWPPKPA?VV?VSHREALRRGPPHQFQGTKLRLSQMPRKRRPD?LAAV-TPLLAISKLLLPV?SCKAPELHHA----------
NGR-JS-L15b  TN-??VTWPPKPA?VV?VSHREALRRGPPHQFQGTKLRLSQMPRKRRPD?LAAV-TPLLAISKLLLPV?SCKAPELH?----------
NGR-JS-LD16  TN??Q*TCSLRRA?AI?VTHPTPVLWRPYV*QQGATMWLS*MPRQRRLH?LVWA-TL*RATSRL*PPV?SCKAPGLHA?-----------
NGR-JS-L16b  TN-??*TCSLRRA?AI?VTHPTPVLWRPYV*QQGATMWLS*MPRQRRLH?LVWA-TL*RATSRL*PPV?SCKAPGLHAP?----------
NGR-JS-LD20  I?-?LVPCPKRPE?AI?LAH*ETLRRRAHDKQQRAILRLQALPRKRRLH?PAWG-IP*HVTSKPLQRA?SC?-----------------
```

Table 4.8: Amino Acid Composition and Frequencies of study HCV Sequences

Isolates	Ala	Cys	Asp	Glu	Phe	Gly	His	Ile	Lys	Leu	Met	Asn	Pro	Gln	Arg	Ser	Thr	Val	Trp	Tyr	Total
H77	14.94	2.30	0.00	1.15	1.15	9.20	0.00	1.15	3.45	8.05	1.15	1.15	11.49	4.60	5.75	10.34	11.49	9.20	2.30	1.15	87
NGR-JS-BD2	15.48	2.38	1.19	1.19	2.38	7.14	0.00	2.38	1.19	8.33	2.38	1.19	11.90	4.76	7.14	9.52	10.71	7.14	2.38	1.19	84
NGR-JS-BD7	12.64	3.45	1.15	2.30	0.00	4.60	2.30	3.45	1.15	8.05	0.00	4.60	8.05	2.30	9.20	9.20	9.20	12.64	3.45	2.30	87
NGR-JS-P02	14.12	5.88	1.18	2.35	0.00	3.53	2.35	3.53	1.18	9.41	1.18	4.71	7.06	3.53	8.24	8.24	8.24	10.59	3.53	1.18	85
NGR-JS-P09a	14.77	4.55	1.14	2.27	0.00	4.55	2.27	4.55	0.00	7.95	1.14	4.55	7.95	2.27	10.23	9.09	7.95	9.09	4.55	1.14	88
NGR-JS-P09b	14.77	4.55	1.14	2.27	0.00	4.55	2.27	3.41	0.00	7.95	1.14	4.55	7.95	2.27	10.23	9.09	7.95	10.23	4.55	1.14	88
NGR-JS-P15a	14.77	4.55	1.14	2.27	0.00	4.55	2.27	3.41	1.14	7.95	1.14	4.55	7.95	2.27	9.09	9.09	7.95	10.23	4.55	1.14	88
NGR-JS-BD6	16.09	2.30	0.00	2.30	1.15	9.20	0.00	1.15	3.45	10.34	1.15	0.00	10.34	3.45	4.60	11.49	11.49	8.05	2.30	1.15	87
NGR-JS-BD9	12.64	3.45	1.15	2.30	0.00	4.60	2.30	3.45	1.15	8.05	0.00	4.60	8.05	2.30	9.20	9.20	9.20	12.64	3.45	2.30	87
NGR-JS-BD10	13.95	4.65	1.16	3.49	0.00	4.65	1.16	3.49	1.16	8.14	0.00	3.49	9.30	1.16	11.63	9.30	8.14	10.47	3.49	1.16	86
NGR-JS-BD12	14.94	2.30	0.00	2.30	1.15	8.05	0.00	2.30	2.30	8.05	2.30	0.00	11.49	4.60	5.75	12.64	9.20	9.20	2.30	1.15	87
NGR-JS-BD13	17.44	2.33	0.00	3.49	1.16	9.30	0.00	1.16	3.49	10.47	1.16	1.16	10.47	3.49	4.65	10.47	10.47	5.81	2.33	1.16	86
NGR-JS-BD20	13.64	2.27	1.14	1.14	1.14	7.95	0.00	0.00	2.27	6.82	0.00	1.14	13.64	4.55	6.82	10.23	13.64	10.23	2.27	1.14	88
NGR-JS-BD23	16.28	1.16	0.00	2.33	1.16	9.30	0.00	2.33	3.49	10.47	1.16	1.16	10.47	3.49	4.65	10.47	10.47	8.14	2.33	1.16	86
NGR-JS-BD28	14.94	2.30	1.15	1.15	1.15	9.20	0.00	1.15	2.30	9.20	1.15	1.15	11.49	3.45	6.90	10.34	11.49	8.05	2.30	1.15	87
NGR-JS-BD30	13.58	2.47	1.23	2.47	0.00	8.64	0.00	0.00	1.23	7.41	3.70	1.23	18.52	1.23	12.35	4.94	11.11	3.70	2.47	3.70	81
NGR-JS-BD1	13.79	4.60	1.15	2.30	0.00	3.45	2.30	2.30	1.15	8.05	1.15	4.60	8.05	2.30	10.34	9.20	9.20	11.49	3.45	1.15	87
NGR-JS-BD3	13.64	3.41	0.00	2.27	1.14	9.09	0.00	1.14	3.41	6.82	1.14	1.14	11.36	5.68	4.55	10.23	11.36	10.23	3.41	0.00	88
NGR-JS-BD4	10.71	3.57	1.19	1.19	0.00	3.57	3.57	2.38	4.76	8.33	0.00	4.76	9.52	0.00	10.71	11.90	8.33	4.76	1.19	1.19	84
NGR-JS-P03	13.79	2.30	0.00	2.30	1.15	8.05	0.00	2.30	4.60	6.90	1.15	0.00	12.64	4.60	4.60	11.49	10.34	10.34	2.30	1.15	87
NGR-JS-P10	13.79	4.60	1.15	2.30	0.00	3.45	2.30	2.30	0.00	8.05	1.15	4.60	8.05	2.30	11.49	9.20	9.20	11.49	3.45	1.15	87
NGR-JS-P13	14.94	2.30	0.00	2.30	1.15	6.90	0.00	2.30	2.30	8.05	1.15	1.15	11.49	4.60	6.90	10.34	10.34	9.20	3.45	1.15	87
NGR-JS-P15	14.94	4.60	1.15	2.30	0.00	4.60	2.30	3.45	1.15	8.05	1.15	3.45	8.05	2.30	9.20	9.20	8.05	10.34	4.60	1.15	87
NGR-JS-P16	14.77	4.55	1.14	2.27	0.00	4.55	2.27	3.41	1.14	7.95	1.14	4.55	7.95	2.27	9.09	9.09	7.95	10.23	4.55	1.14	88
NGR-JS-P17	14.46	2.41	1.20	3.61	1.20	6.02	1.20	4.82	2.41	7.23	0.00	0.00	9.64	4.82	12.05	10.84	9.64	7.23	0.00	1.20	83
NGR-JS-P18	11.49	5.75	1.15	4.60	0.00	4.60	1.15	1.15	2.30	8.05	0.00	1.15	12.64	1.15	8.05	11.49	12.64	9.20	3.45	0.00	87
NGR-JS-P06	14.77	4.55	1.14	2.27	0.00	4.55	2.27	3.41	1.14	7.95	1.14	4.55	7.95	2.27	9.09	9.09	7.95	10.23	4.55	1.14	88
NGR-JS-P09	14.77	4.55	1.14	2.27	0.00	4.55	2.27	3.41	0.00	6.82	1.14	4.55	7.95	2.27	10.23	10.23	7.95	10.23	4.55	1.14	88
Avg	14.32	3.51	0.83	2.31	0.54	6.15	1.24	2.48	1.90	8.17	1.03	2.64	10.03	3.01	8.25	9.82	9.82	9.45	3.26	1.24	

4.2.7 Substitutionsin Amino acid of NS5B gene

Tables 4.9- 4.12 show amino acid substitutions observed and the distribution among study populations. Major mutations with clinical implications that have been reported in previous studies associated with NS5B gene observed include **S15G** where amino acid serine is substituted by glycine (serine-glycine). S15G mutation affects the replication capacity of HCV NS5B polymerase and this reduces the replication fitness of the RNA template.Ten S15G mutation (36%)were observed among patients with HIV; 12 (43%) out of which 3 had S15G and 9 had substitutions with either Alanine or Valine (S15A/V) were observed in patients with clinical hepatitis; and 6 (21%) S15G observed in blood donors as shown in Tables 4.9 and 4.12.Among patients with HIV, all subtype 5a isolates, as well as lone subtype 3a had serine (S) substituted by glycine (G); while in all subtype 1a serine was conserved.Among patients with clinical hepatitis, serine was conserved in all subtype 1a; substituted by glycine in one isolate (subtype 1b) while all other 1b isolates had valine substitution. For subtype 5a, serine was substituted by glycine in 2 isolates, substituted by alanine in 2 isolates while conserved in 1 isolate. Blood donors had serine conserved in 6 isolates and glycine in 3 isolates of subtype 1a; while lone subtype 2c had Thymine substitution. Subtype 5a isolates had serine conserved in 1 isolate and substituted by glycine in 2 other isolates.

Mutation **Q47H** was not detected but its absence is of clinical importance. Most isolates from this study have this mutation absent. Glutamine (Q) was substituted by Histamine (H) at amino acid position 47. Among patients with clinical hepatitis, 14 (40%) substitutions were found thus, Histidine (H) was substituted by Proline (Q47P) in 4 subtype 5a isolates, 8 Q47Rmutation in 1 subtype 1a, subtype 1b (1), 2b (4), 5a (2) while glutamine was conserved in 2 isolates (1b). There were 2 Q47D (Aspartate) in 2 subtype 1b isolates; and 2 Q47 conserved in 2 subtype 1b isolates.Patients with HIV had Histidine substituted by Proline (P), Phenylalanine (F) (subtype 3a) or Arginine (R) (subtype 1a); in 11 (31%) individuals infected with subtype 5a and 1 had Q47 conserved However among blood donors, there were 10 (29%) Q47R/P/ L) substitutions, 3 Q47Q were conserved.

At position 79 **(T79M)**, there was substitution of Thymine by Methionine in 8(67%) in all subtype 5a isolates in patients with HIV, 3(25%) in subtype 5a isolates and alanine in 1 isolate (1b) in patients with clinical hepatitis and 1(8%) isolate in blood donors.

Several other mutations were also observed including **T7N/S** where 9(38%) isolates including subtypes 5a had serine substitutions, Thymine conserved in subtype 1a , while 3a had valine in patients with HIV; 6(24%) in clinical hepatitis, with all the 2b having proline, T conserved or deleted in5a ; and 9 (38%) were observed among blood donors where T was conserved in all subtype 1a isolates, subtype 5a had asparagine (N) substitution and subtype 2c had proline..

G61R/Emutation was observed in isolates among blood donors, where glycine was substituted by Arginine or glutamate in subtypes 1a (4 isolates), deleted in subtype 5a (3 isolates) and conserved in 1a (4 isolates). However, in patients with clinical hepatitis, glycine was substituted by Thymine (T) in subtypes 2b and 5a; and by Lysine (K) and Leucine (L) in subtypes 1b. Among patients with HIV, 7 (35%) isolates (subtype 5a) had glycine substituted by Tryptophan (W)while two (5a) had deletions, one subtype 1a had lysine and others conserved at that position

Table 4.12 gives the summary of the distribution of these substitutions in all the isolates found in the study populations. Other polymorphisms found were S54L, T71I, S84F, Q67R and G89E etc. The polymorphisms observed differ among the various HCV subtypes found in this study. This implies that diversity of HCV strains obtained is a reflection of the mutations observed in NS5B gene. Mutations are highest in HCV subtypes circulating in patients with clinical hepatitis, followed by individuals with HIV and least among blood donors. However, there was no significant difference in amino acid substitutions among the study groups ($p=0.23, P< .05$).

Table 4.9: Observed mutationsat different positions in Patients with HIV

Isolates in Patients with HIV		**Positions of Mutations in NS5B gene**										
ISOLATE ID	SUBTYPE	15	47	7	54	61	71	67	79	84	89	80
H77 Prototype	1a	S	Q	T	S	G	T	Q	T	S	G	T
NGR-JS-P03	1a	-	-	-	-	K	-	-	-	-	E	-
NGR-JS-P13	1a	-	R	-	-	-	-	-	-	-	E	-
NGR-JS-P09	5a	**G**	P	S	*	W	-	-	M	-	-	S
NGR-JS-P09b	5a	**G**	P	S	*	W	-	-	M	-	-	L
NGR-JS-P10	5a	**G**	P	S	*	*	-	-	M	-	R	I
NGR-JS-P06	5a	G	P	S	*	W	-	-	M	-	-	I
NGR-JS-P15& P15b	5a	**G**	P	**S**	*	W	-	-	M	-	-	I
NGR-JS-P16	5a	**G**	P	S	*	W	-	-	M	-	-	I
NGR-JS-P21	3a	**G**	F	V	C	-	S	*	-	-	L	I
NGR-JS-P18	5a	**G**	P	N	-	*	-	-	-	-	E	-
NGR-JS-P02	5a	**G**	P	S	*	W	-	-	M	C	-	I
NGR-JS-P07	5a	**G**	P	-	*	W	-	-	M	-	-	**I**

Key: (*) indicates deletions and (–) conserved amino acid

Table4.10: Observed Mutations in Patients with Clinical Hepatitis

Isolates in Patients with Clinical Hepatitis		**Positions of Mutations in NS5B gene**									
ISOLATE ID	SUBTYPE	15	47	7	54	61	71	67	79	84	89
H77 prototype	1a	S	Q	T	S	G	T	Q	T	S	G
NGR-JS- LD4	1a	-	P	N	L	*	-	-	-	-	K
NGR-JS- LD6	1a	-	R	-	-	-	-	-	-	-	E
NGR-JS- LD 5	1b	**G**	R	P	-	L	-	-	-	-	E
NGR-JS-LD10	1b	**V**	D	R	I	K	-	-	-	-	-
NGR-JS-LD10b	1b	*	**D**	*	I	K	*	-	*	-	-
NGR-JS-LD15	1b	**V**	-	P	P	L	P	S	*	-	-
NGR-JS-LD15b	1b	**V**	-	P	P	-	-	S	*	*	*
NGR-JS- LD7	2b	**A**	R	-	G	T	C	-	-	-	-
NGR-JS- LD12	2b	**A**	R	P	G	T	C	-	A	-	-
NGR-JS- LD13	2b	**A**	R	P	G	T	C	R	*	-	-
NGR-JS- LD20	2b	**A**	R	P	G	T	C	R	?	-	-
NGR-JS- LD1	5a	**G**	**P**	-	*	W	-	-	M	-	-
NGR-JS-LD2	5a	-	**P**	-	-	*	-	*	-	-	K
NGR-JS- LD3	5a	**G**	**P**	S	*	*	-	-	M	-	R
NGR-JS- LD16	5a	**A**	R	-	A	T	C	P	*	-	-
NGR-JS-LD16b	5a	**A**	R	*	A	T	C	*	*	-	-

Key: (*) indicates deletions and (–) conserved amino acid

Table 4.11: Observed mutations found in Blood Donors

HCV Isolates in Blood Donors		**Positions of Mutations in NS5B gene**									
ISOLATE ID	SUBTYPE	**15**	**47**	7	54	61	**71**	67	79	84	89
H77 Prototype	1a	**S**	**Q**	T	S	G	**T**	Q	T	S	G
NGR-JS-BD2	1a	**-**	R	-	*	-	-	-	-	F	-
NGR-JS-BD3	1a	**G**	-	-	**-**	E	**-**	-	-	-	-
NGR-JS-BD1	1a	**G**	**P**	S	*	*	**-**	**-**	M	**-**	K
NGR-JS-BD6	1a	-	L	-	-	E	-	-	-	-	-
NGR-JS-BD12	1a	-	**-**	-	-	-	**-**	-	-	-	E
NGR-JS-BD13	1a	**G**	L	-	-	E	-	-	-	-	-
NGR-JS-BD20	1a	-	-	-	**-**	-	**-**	-	-	-	-
NGR-JS-BD23	1a	**-**	**L**	-	-	E	-	**-**	-	-	-
NGR-JS-BD28	1a	-	L	-	-	-	-	-	-	-	-
NGR-JS-BD30	2c	**T**	P	P	-	-	-	R	-	-	-
NGR-JS-BD7	5a	**G**	L	N	*	*	I	-	-	-	-
NGR-JS-BD9	5a	-	L	N	*	*	I	-	-	-	K
NGR-JS-BD10	5a	**G**	**L**	**N**	*	*	I	**R**	-	-	-

Key: (*) indicates deletions and (–) conserved amino acid

Summary Table 4.12: Distribution of mutations among study participants

Study Populations	Polymorphism										
	S15G/ A/V	**Q47H**	**T7N/ S**	**S54L**	**G61R /E**	**T71I**	**Q67R**	**T79 M**	**S84F**	**G89E**	**p<.05**
Patients with HIV N= 13	10 (36%)	11 (31%)	9 (38%)	6 (27%)del etions	8 (35%)	1 (9%)	0 (0%)	8 (67%)	1 (50%)	5 (36%)	p=0.20
Blood Donors N =13	6 (21%)	10 (29%)	6 (24%)	5 (23%)	4 (17%)	3 (27%)	2 (29%)	1 (8%)	1 (50%)	4 (28%)	p=0.23
Patients with clinical hepatitis N =16	12 (43%)	14 (40%)	9 (38%)	11 (50%)	11 (28%)	7 (64%)	5 (71%)	3 (25%)	0 (0%)	5 (36%)	p =0.21

OBJECTIVE 3: To identify post-translational modifications in NS5B protein

4.2.8 Post-translational modifications in HCV NS5B protein

Tables 4.13 - 4.15 show the phosphorylation pattern of NS5B protein in all the HCV isolates in the study participants. In HCV, phosphorylation (addition of phosphoryl group $P03^-$) occurs mainly on serine and threonine amino acid residues. Analysis of NS5B protein showed different serine and threonine phosphorylated residues. For any amino acid to be phosphorylated, it must have a score more than 0.500 which is the set limit on this software.

Table 4.16 shows amino acid residues with highest scores(>0.900), which is any score above 0.500threshold among all study participants.

1. Among patients with clinical hepatitis, 11 isolates (69%) had Serine phosphorylation at Serine amino acid positions:Ser 6, 17, 35(2), 51, 52, 60(2), 71, 82. While 5 (31%) had threonine (Thr) phosphorylation at positions: 7, 52(3), 56. However, only 4 (31%) isolates (of subtypes 5a and 2b) had either Ser /Thr 52 or 73 which are viral markers of disease progression.
2. Blood donors on the other hand, had 7(54%) isolates phosphorylated at Serine positions 17, 53, 56(2), 57(2), 58 and 6 (46%) isolates had threonine phosphorylation at positions 7(2), 51, 52(2), 56. Two (15%**)** isolates (subtype 5a) had either Ser /Thr 52 or 73.
3. Among HCV/ HIV co-infected persons, 4 (31%) isolates were phosphorylated at Serine 6 (2), 9, 73; while 9 (69.23%) had threonine phosphorylation at positions 7, 52(6), 56, with 7 (54%**)** isolatesall of subtype 5a having either Ser /Thr 52 or 73. Overall, Ser / Thr 52 and 73 viral disease markers were highest among individuals with HIV, followed by patients with clinical hepatitis (p=0.04). Ser / Thr 52 and 73 markers were found only in HCV subtypes 5a and 2b. This result implies that HCV subtypes with Serine or threonine phosphorylated at positions 52 or 73 may be implicated in severe disease outcome or progress faster with hepatic pathology in infected patients.

Table 4.13: Phosphorylation of HCV NS5B Proteins (P68) at the Serine (Ser), Threonine (Thr) and Tyrosine residues among patients with clinical hepatitis

Sequence Name	**Serine (S) Phosphorylation >0.500 threshold**	**Threonine(T) Phosphorylation >0.500 threshold**	**Genotypes/Subtpes**	**Highest Score above threshold/Site**	**Protein Length**	**Protein Kinase**	**Phosphorylated sites with highest output scores**
NGR-JS-LD4	-	**Thr 7**	1a	0.982	87	unsp	VVTWTPRPR
NGR-JS-LD6	**Ser 82**	-	1a	0.983	88	unsp	ARSWSCEAT
NGR-JS-LD5	**Ser 71**	**-p**	1b	0.994	84	unsp	TARCSTETT
NGR-JS-LD10 &10b	**Ser 60**	**-**	1b	0.953	71	unsp	IPHVTSKPL
NGR-JS-LD15	**Ser 35**	**-**	1b	0.992	71	unsp	KLRLSQMPR
NGR-J-S-LD15b	**Ser 35**	-	1b	0.992	69	unsp	KLRLSQMPR
NGR-JS-LD7	**-**	**Thr 52**	2b	0.985	71	unsp	LQRASCGDR
NGR-JS-LD12	**Ser 60**	**-**	2b	0.991	70	unsp	LRRVSCKAP
NGR-JS-LD13	**-**	**Thr 52**	2b	0.953	66	unsp	IPHVTSKPL
NGR-JS-LD20	**-**	**Thr 52**	2b	0.953	62	unsp	IPHVTSKPL
NGR-JS-LD1	-	**Thr 56**	5a	0.954	87	unsp	TPRATLRLP
NGR-JS-LD2	**Ser 17**	-	5a	0.988	88	unsp	HYSHSPKGC
NGR-JS-LD3	**Ser 6**	-	5a	0.957	87	unsp	HVTCSLRRA
NGR-JS-LD16	**Ser 52**	**-**	5a	0.985	67	unsp	LRATSRLPP
NGR-JS-LD16b	**Ser 51**	**-**	5a	0.985	66	unsp	LRATSRPP

Table 4.14: Phosphorylation of HCV NS5B Proteins at the Serine(S), Threonine (T) residues among blood donors

Sequence ID	Serine (S) Phosphorylation >0.500 threshold	Threonine(T)Phosphorylation >0.500 threshold	Subtypes	Highest Score above threshold/Site	Protein Length	ProteinKinase	Phosphorylated sites with highest output scores
NGR-JS-BD2	**Ser 58**	-	1a	0.939	88	unsp	SLIT**S**RRGQ
NGR-JS-BD3	**Ser 57**	-	1a	0.992	88	unsp	SLAT**S**RPEQ
NGR-JS-BD6	**Ser 56**	-	1a	0.996	86	unsp	PSLTSRREG
NGR-JS-BD12	-	**Thr 7**	1a	0.926	84	unsp	VVTMTVKPA
NGR-JS-BD13	**Ser 56**	-	1a	0990	86	unsp	SLATSRPEQ
NGR-JS-BD20	**Ser 17**	-	1a	0.997	88	unsp	PSSPSPRDE
NGR-JS-BD23	**Ser 57**	-	1a	0.992	86	unsp	SLATSRPEQ
NGR-JS-BD28	-	**Thr 7**	1a	0.984	87	unsp	DVTWTPKPA
NGR-JS-BD30	**Ser 53**	-	2c	0.965	82	unsp	SLATSPGQR
NGR-JS-BD1	-	**Thr 56**	5a	0.954	87	unsp	TPRA**T**LRLP
NGR-JS-BD7	-	**Thr 52**	5a	0.990	87	unsp	PVWATPRVT
NGR-JS-BD9	-	**Thr 52**	5a	0.990	87	unsp	PVWATPRVT
NGR-JS-BD10	-	**Thr 51**	5a	0.990	86	unsp	PVWATPRVT

Table 4.15: Phosphorylation of NS5B Proteins at the Serine(S) and Threonine (T) residues among individuals with HIV

Sequence ID	Serine (S) Phosphorylation>0.500 threshold	Threonine(T) Phosphorylation with >0.500 threshold	Subtypes	Highest Score above threshold/Site	Protein Length	ProteinKinase	Phosphorylated protein sites with highest output scores
NGR-JS-P03	-	**Thr 7**	1a	0.973	87	unsp	VVTWTPKPV
NGR-JS-P13	-	**Thr 7**	1a	0.982	87	unsp	VVTWTPRPA
NGR-JS-P21	**Ser 9**	-	3a	0.926	82	unsp	RRVLSMVGI
NGR-JS-P18	**Ser 73**	-	5a	0.993	86	unsp	ARSWSCEAT
NGR-JS-P09b	-	**Thr 52**	5a	0.940	87	unsp	LVWATRAT
NGR-JS-P07	-	**Thr 52**	5a	0.940	88	unsp	LVWATRAT
NGR-JS-P02	-	**Thr 52**	5a	0.940	87	unsp	LVWATPRAT
NGR-JS-P06	-	**Thr 52**	5a	0.940	88	unsp	LVWATPRAT
NGR-JS-P09	-	**Thr 52**	5a	0.940	87	unsp	LVWATPRAT
NGR-JS-P10	-	**Thr 56**	5a	0.954	87	unsp	TPRATLRLP
NGR-JS-P15 & P15b	**Ser 6** **Ser 6**	-	5a 5a	0.957	87	unsp	HVTCSLRRA HVTCSLRRA
NGR-JS-P16	-	**Thr 52**	5a	0.940	88	unsp	LVWATPRAT

Table 4.16: Summary of Distribution of Phosphorylated amino acid residues in study populations

Phosphorylated residues/ sites	**Patients with Clinical Hepatitis N =16**	**Blood Donors N =13**	**Individuals with HIV N= 13**	**P< .05**
Serine	11 (69%)	7 (54%)	4 (31%)	
Threonine	5 (31%)	6 (46%)	9 (69%)	
Total	**16(100%)**	**13 (100%)**	**13 (100%)**	
Serine / Threonine 52 or 73(markers)	4(31%) Subtype 5a= 1 Subtype 2b=3	2(15%) Subtype 5a	7(54%) Subtype 5a	0.04

```
Protein Sequence Length= 84 amino acids
# Sequence                 # x   Context     Score   Kinase    Answer
# ------------------------------------------------------------------
# Sequence                  1 T   ----TVVTP   0.454   CaM-II      .
# Sequence                  4 T   -TVVTPPKP   0.632   unsp       YES
# Sequence                 12 Y   PDRPYGRSQ   0.447   SRC         .
# Sequence                 15 S   PYGRSQSGF   0.565   DNAPK      YES
# Sequence                 17 S   GRSQSGFIS   0.636   PKA        YES
# Sequence                 21 S   SGFISGAPL   0.492   PKA         .
# Sequence                 51 S   VAIPSHATR   0.500   PKC        YES
# Sequence                 54 T   PSHATRPLR   0.447   GSK3        .
# Sequence                 64 S   VELQSSRTA   0.675   unsp       YES
# Sequence                 65 S   ELQSSRTAR   0.460   cdc2        .
# Sequence                 67 T   QSSRTARCS   0.786   PKC        YES
# Sequence                 71 S   TARCSTETT   0.994   unsp       YES
# Sequence                 72 T   ARCSTETTL   0.531   PKA        YES
# Sequence                 74 T   CSTETTLSL   0.522   PKC        YES
# Sequence                 75 T   STETTLSLS   0.634   unsp       YES
# Sequence                 77 S   ETTLSLSVR   0.509   DNAPK      YES
# Sequence                 79 S   TLSLSVRAR   0.987   unsp       YES
#

TVVTPPKPDRPYGRSQSGFISGAPLIQKGRAAVIAGAAQAACQRLAVAIP   #     50
    SHATRPLRPVELQSSRTARCSTETTLSLSVRARE                   #    100
%1  ...T.........S.S.................................   #     50
%1  S............S..T...ST.TT.S.S.....
```

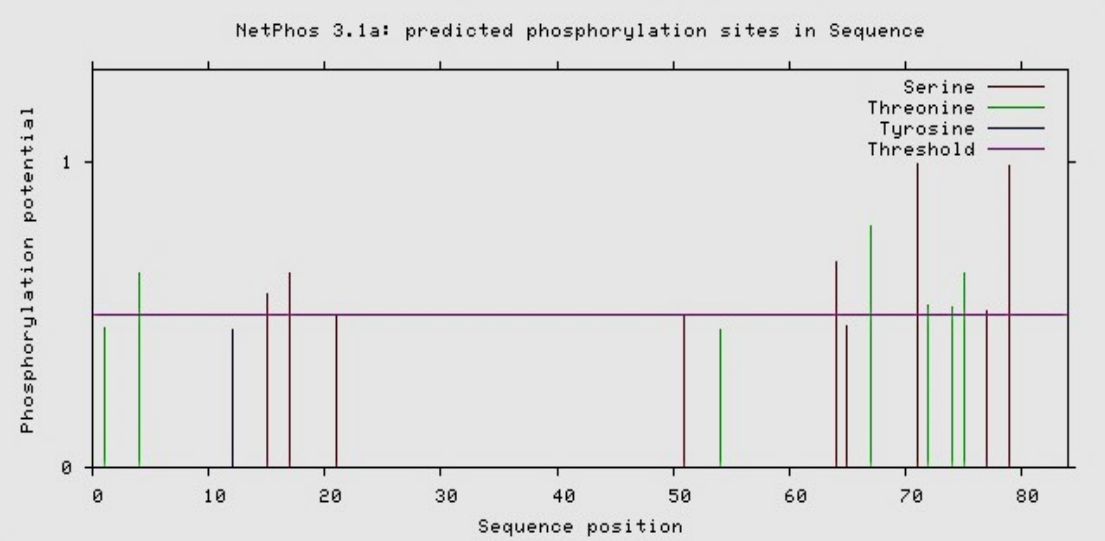

Fig 4. 9: Phosphorylation Graph of NS5B protein in a patient with clinical hepatitis (LD5) at Serine 71 position

```
Protein Sequence   Length = 86 amino acids
# Sequence                # x   Context     Score   Kinase    Answer
# ------------------------------------------------------------------
# Sequence                 4 T   -PRETYNLR   0.494   unsp        .
# Sequence                 5 Y   PRETYNLRR   0.371   INSR        .
# Sequence                16 S   QCGHSPNAC   0.840   unsp       YES
# Sequence                21 T   PNACTVEAL   0.442   GSK3        .
# Sequence                28 T   ALCITARGS   0.475   GSK3        .
# Sequence                32 S   TARGSNVVI   0.660   unsp       YES
# Sequence                44 S   APAASSLPV   0.463   cdc2        .
# Sequence                45 S   PAASSLPVW   0.479   DNAPK       .
# Sequence                51 T   PVWATPRVT   0.990   unsp       YES
# Sequence                55 T   TPRVTLRLP   0.932   unsp       YES
# Sequence                65 S   VELRSSGIA   0.467   GSK3        .
# Sequence                66 S   ELRSSGIAR   0.876   unsp       YES
# Sequence                71 S   GIARSWCVV   0.549   cdc2       YES
# Sequence                76 T   WCVVTTLWP   0.465   cdc2        .
# Sequence                77 T   CVVTTLWPS   0.455   GSK3        .
# Sequence                81 S   TLWPSARAK   0.815   PKC        YES
#


PRETYNLRRARQCGHSPNACTVEALCITARGSNVVIVDAAPAASSLPVWA   #     50
TPRVTLRLPPVELRSSGIARSWCVVTTLWPSARAKG                 #    100
%1  ..............S..............S.................   #     50
%1  T...T.........S....S.........S.....
```

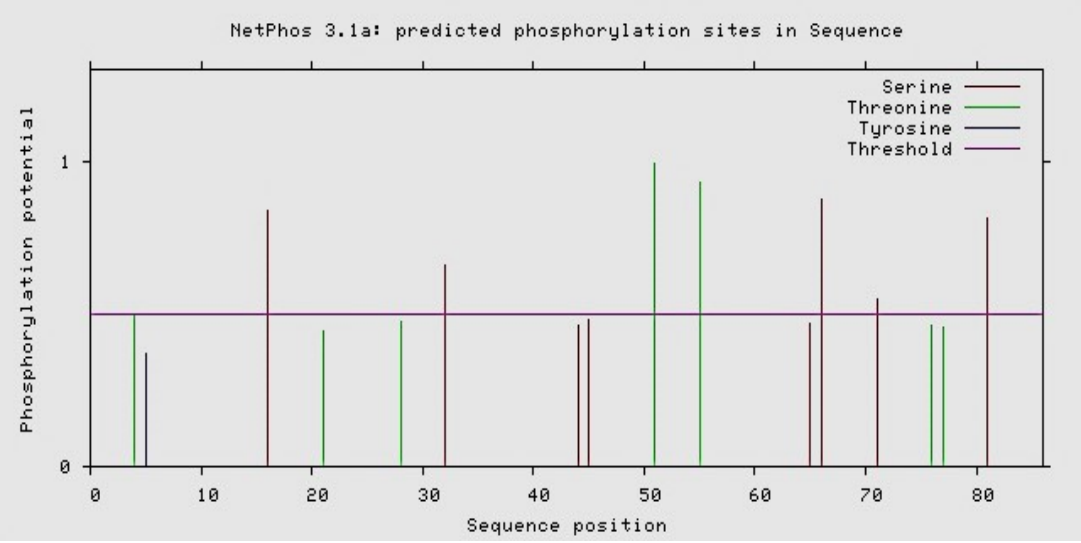

Fig.4.10: Phosphorylation graph of NS5B protein in a blood donor (NGR-JS-BD10)at Threonine 51 position

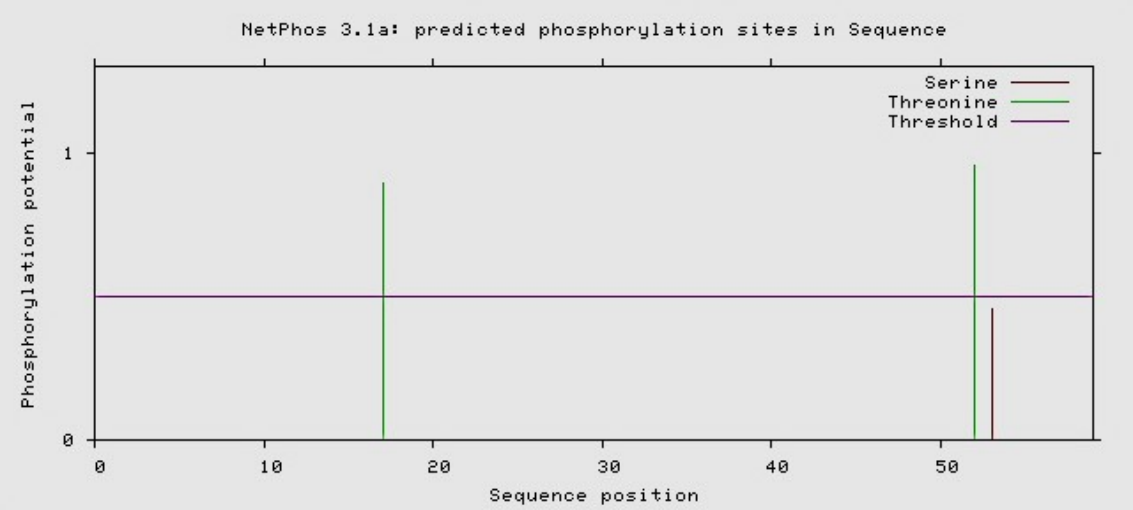

Fig. 4.11: Protein Phosphorylation graph of isolate LD 20 in a patient with clinical hepatitisat Threonine 52 position.

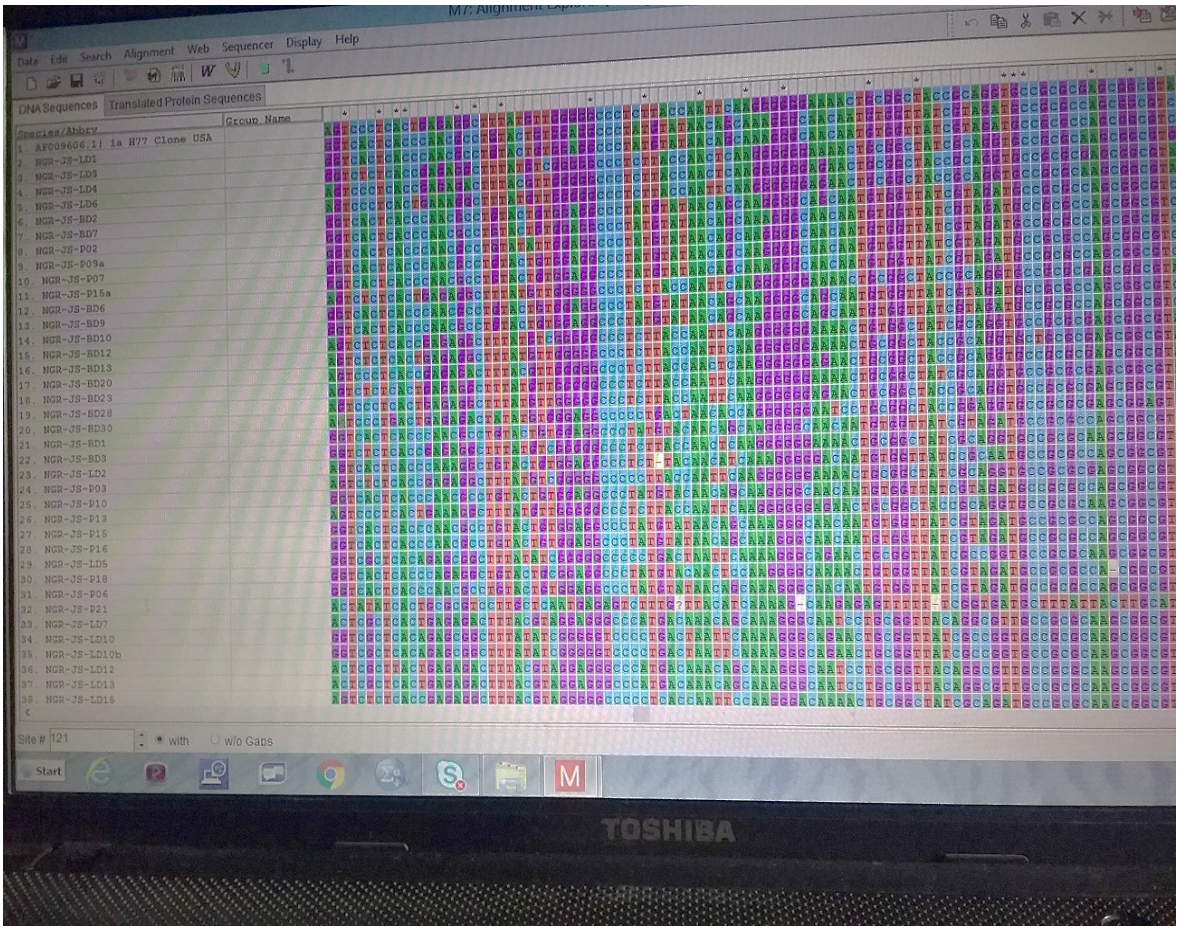

Fig. 4.12: Sequence alignment of study isolates in MEGA 7.0

4.3 SUMMARY OF MAJOR FINDINGS FROM THIS STUDY

1 Multiple genotypes of HCV exist in Ibadan with predominant subtypes as 1a and 5a.

2 Genotypes are distributed differently among the study participants, with subtypes 1b and 2b found only among patients with clinical hepatitis, 1a among donors, 5a in HIV and patients with clinical disease.

3 Mutations are highest in HCV subtypes circulating in patients with clinical hepatitis, followed by individuals with HIV and least among blood donors. There is no significant difference among the groups ($p=0.23$).

4 Ser / Thr 52 and 73 viral disease markers highest among individuals with HIV, followed by hepatitis patients, with a statistical significant difference ($p= 0.04$).

5 Ser / Thr 52 and 73 markers found only in Hepatitis CVirus subtypes 5a and 2b.

CHAPTER FIVE

DISCUSSION

In this study, hepatitis C virus (HCV) NS5B gene, a non-structural protein that codes for RNA-dependent-RNA polymerase (RdRp) enzyme, essentially involved in HCV RNA replicationhas been characterized. The Basic Local Alignment Search Tool (BLAST) in National Centre for Biotechnology Information (NCBI) was used to determine the evolutionary relatedness of the sequences to the HCV prototype strain H77 Accession number 009606 (Kolykhalov *et al*., 1997) and other reference sequences from different countries obtained from global data bank(GenBank). BLAST is a pop set evolutionary tool used to analyze and ascertain the evolutionary relatedness of isolates (NCBI). BLAST homology of between 97%- 100% shows very close identity. For instance, the closest HCV strains to Nigerian HCV isolates from this study are from South Africa with accession number KC 767832 and DQ333676 belonging to HCV subtype 5a with 99% identity and 100% respectively. NGR-JS-BD9 sequence representing a blood donor has 100% identity with DQ333676 from South Africa (Table 4.2 & 4.3). This shows that the virus found here has almost the same genetic identity with those from South Africa with about 1% difference across the genome in Blast identity of 99%. On the other hand, previously identified Nigeria isolate closest to our study isolates belonging to genotype 1 has accession number JQ679065 and belongs to genotype 1 with 89% identity, indicating a difference of about 10% between the two sequences. The previously found strain by Forbi *et al*., (2012), in North- Central Nigeria was notsubtyped; possibly they belonged to different subtypes accounting for the reason for the difference. The sequences vary most at the beginning with first three nucleotide codons seen as TCT in NGR-JS-P09, NGR-JS-B30, NGR-JS-P10; GGA in NGR-JS-BD2; TGG in NGR-JS-BD9 and TTG in NGR-JS-BD23 (Figure 4.5). This variation may have resulted from single/double/triple nucleotide insertion or deletion (substitution) across the genome and is responsible for the existence of HCV variants that have emerged as genotypes and subtypes in this work. Analysis of the NS5B genes revealed six subtypes of the virus as 1a, 1b, 2b, 2c, 3a and 5a belonging to four HCV genotypes: 1, 2, 3 and 5. In previous studies, multiple genotypes have been reported in Nigeria with genotype 1 as the most prevalent (Oni *et al*., 1996; Forbi et al. 2012). Subtype 3a in NGR-JS-P21 aligned with a mutant strain from Sri Lanka. This isolate had several substitutions and differed widely

from other isolates. The study sequences are available at the repository DOI: 10.6084/m9figshare.7471454.

Here, hepatitis C virus genotype 5 (subtype 5a) is reported for the first time in West Africa, responsible for 45.2% of HCV infections in the study participants. Previously, genotypes 1, 2 and 4 were known to be restricted to West Africa and Middle east respectively (Markov *et al*., 2009; Rao *et al*., 2015; Mora *et al*, 2016; Candotti *et al.*, 2003). But this study has shown that additional genotype 5 also circulate in West Africa.

Previously and epidemiologically, this genotype was known to be restricted to South Africa. In a study conducted by Naamani *et al* (2013), they showed that subtype 5a (the only subtype found so far in that group) is responsible for about 40% of infections caused by HCV in South Africa. In 1995, Smuts and Kannemeyer genotyped HCV in South Africa and found genotype 5a as the predominant strain (39.2%) in northern part of that country (Smuts and Kannemeyer, 1995). Other studies that have reported genotype 5 in their work include Payan *et al*.,(2005) who reported prevalence of 14.2% in France, Jover *et al*., (2001) got 10.3% prevalence in Southeast Spain, Antaki *et a*l., (2009) with 10% in Syria, D'Heygere *et al.,* (2011) 4.2% in Belgium, Antaki *et al* (2010) 5% in Montreal, Quebec Canada and Shobokshi *et al* (1999) who got 1% prevalence among the studied populations. This finding has therefore advanced knowledge on the diversity of HCV genotype in Western Africa, which has impacts on antiviral therapy, vaccine formulation and management policies on HCV infections.

There is paucity of data on the pathological and clinical characteristics of HCV genotype 5 infections. Previous studies that involved patients were mainly women aged fifty years and older, with some already having cirrhosis at presentation time; in a small sample sized study, with major means of HCV transmission as through blood transfusion and intravenous drug use (Naamani *et al.,* 2013).

Genotype 5 found in this study is closely related to genotype 5 from South Africa. However, not much is known about its point of introduction into Nigeria, just like other HCV genotypes. Obviously, globalization contributes majorly to transmission of viruses including hepatitis Cvirus; yet, several other factors may facilitate species adaptation in new areas. Based on this, future research that can answer pertinent questions regarding HCV genotype 5a epidemic history and evolution in West African region may be necessary.

This study also identified HCV subtype 1a as the second predominant strain after genotype 5a among the study populations. Previous studies have confirmed the presence of divergent strains of genotype 1 in West Africa, where it has been endemic with its lineage traced as far back as 17^{th} -20^{th} centuries (Markov et al., 2009; Forbi et al., 2012; Pouillot et al., 2008b).

HCV subtype 1a is distributed to all regions of the world. Its global distribution has been established in most studies (Forbi *et al*., 2012; Smith *et al*., 1997) within and outside the continent of Africa. The presence of subtype 1a among blood donors than in patients with hepatitis is a direct opposite of what we found with subtype 5a, where more subjects with HIV and clinical hepatitis were infected than blood donors. These findings appear very interesting and may be a pointer to why some people clear the virus after infection or progress to chronic stage, or respond faster during antiviral therapy, with regard to genotype and virulence. This prompts another important question about the pathogenesis of various HCV genotypes in chronic and acute HCV infections. Subtype 1a from this work formed a long cluster with strains from Argentina, Brazil and Canada but were seen further away from the previous Nigerian genotype 1 strains. The only previous Nigerian genotype1 strain seen close to our subtype 1a isolates on the phylogenetic tree is strain with accession number JQ679065.

Subtype 1b was found only in patients with hepatitis, as well as subtype 2b (Table 2). These subtypes were absent among individuals with HIV and blood donors. Subtype 1b is closely related to a strain from Brazil but seen closer to previous Nigerian genotype 1 strains than subtype 1a, while subtype 2b clustered also with strains from Brazil and USA but not those from Africa. Prior to this time, not many studies from Nigeria have shown the prevalence of 1b and 2b subtypes among this population. This study has contributed and advanced knowledge on HCV strains circulatingamong patients with clinical hepatitis, since these subtypes were strictly found among this subpopulation in this study.

HCV genotype 2 (subtype 2c) caused infection in one blood donor subject (Table 2). On the phylogenetic tree, this genotype formed a cluster with genotype 2 strains from Guinea Bissau (Figure 4.2). It is closely related to a recently isolated genotype 7 strain with accession number CS101285 in Canada (Murphy *et al.,* 2015) in an immigrant from Democratic Republic of Congo, and genotype 4 with accession number FJ791092 from Central African Republic (Njouom *et al.,* 2009). Genotype 2 of HCV is a common genotype found in West Africa (Markov *et al*., 2009; Candotti *et al*., 2003; Forbi *et al*., 2012).

Statistically, a significant difference (p=0.0004) existsbetween the HCV genotypes responsible for infection in patients with clinical disease, HIV–infected patients and blood

donors) (Table 4.3). The reason for this disparity is not known but may likely be due to route of transmission of this virus or some other host's and viral factors yet to be determined. This therefore, may necessitate further research into HCV genotype-specific infections in symptomatic and asymptomatic HCV-infected individuals.

Hepatitis C virus like other RNA genomes has great genetic variability due to error in replication or inability to proofread incorporated sequences by RNA-dependent-RNA polymerase, which then results in high mutation in the virus. The accumulation of these genetic polymorphisms results in numerous genetic variants in clades, genotypes, subtypes, isolates and quasispecies (Le Guillou-Guillemette *etal*., 2007). This study is consistent with previous findings on genetic variability both at the nucleotide and amino acid levels, as seen in multiple strains obtained in this study (Figure 4.4; Tables 4.5- 4.7). This is in agreement with the work done by Okwuraiwe *et al,* (2012) where multiple genotypes were also found in Lagos State, Nigeria.

On the phylogenetic trees (Figures 2 to 6) some of the sequences were seen branched further away from the rest. For instance, NGR-JS-BD30, an isolate in found among blood donors, subject has a longer branch than the others. This actually indicates how distantly related these sequences are compared to those in their groups. NGR-JS- P21 belongs to genotype 3 but seen branched further away from other groups (Figure 4.5). The longer the branch, the higher the nucleotide variation and the more distantly related the isolate is to others.

HCV response to therapy is dependent on the infecting genotype, whether with interferon and ribavirin or with direct-acting antivirals (DAAs), which are known to achieve sustained virologic response of up to 50% - 90% (Lawitz et al., 2013). Studies have shown that genotypes 1 &4 are the most problematic HCV types to treat (WHO, 2012; Tucker, 2013; Dhawan, 2016). In the past, other studies (Forbi *et al.*, 2012; Okwuraiwe, 2012) have revealed higher prevalence of genotypes 1 similar to high prevalence of genotype 1 reported in the present study, because they clustered together on phylogenetic tree though with little differences.

There were 56 amino acid substitutionsobserved in NS5B gene as shown in Table 4.6. Major clinically- important mutations associated with NS5B gene observed in the study include S15G where amino acid serine is substituted by glycine, Q47H, T7N/S, S54L, A56V, G61R/E, T71I, Q67R, T79M, S84F, and G89E. These mutations were distributed differently among the study population (Tables 4.9 -4.12). Mutations S282T typically known for

itsconferment of resistance to Sofosbuvir as reported by Tong *et al.,*(2014),was not detected probably due to length of NS5B gene sequenced.

NS5B polymorphisms may explain differences in treatment outcomes among patients and mutations at the NS5B region could be associated with poor prognosis of the disease in HCV-infected patients as such mutations alter the polymerase activity of NS5B. Hence, identifying drug resistance mutations (DRMs) is important as new direct -acting antiviral (DAA) drugs become available.

According to Castilho*et al* (2015), changes in the amino acid position 109 of NS5B is notably associated with resistance to antiviral therapy and infectivity of the virus. Resistance to antiviral drugs may also be as a result of uncharacterized mutations and interactions among mutations.

L159/L320F mutations in NS5B (polymerase) confer a low resistance to most HCV polymerase inhibitors especially Mericitabine (a prodrug of HCV NS5B polymerase inhibitor PSI-6130); and to Sofosbuvir (Tong *et al,* 2014). In the study sequences, this mutation was not observed but at other positions along the NS5B gene, L (Leucine) was substituted by different amino acids other than phenylalanine (F). These polymorphisms may be linked to more dysfunctionality in the gene and not only resistance.

S15G mutation affects the replication capacity of HCV NS5B polymerase which reduces the replication fitness of the RNA template (Tong *et al.,* 2014; Iam *et al.,* 2014). S15G was detected in this work and this major mutation at amino acid position 15 in which Serine (S) was substituted by Glycine (G), Alanine or Valine (Tables 4.11 and 4.12). This mutation occurs more in HCV genotype 5a infected persons and might be under positive selection in this group, hence may determine the outcome of infection in the groups. Another factor that affects replication fitness is co-infection. HCV co-infection with HIV/HBV has been linked with a reduced treatment response. HIV can affect NS5B variability, suggesting that an already compromised immune structure(by HIV) can actually affect genetic diversity of HCV by pathogenically influencing the viral replication fitness(Marascio *et al.*,2014, Blackard *et al.,* 2010).

Another major mutation is **Q47H**(absence in the gene has great implication for response to Daprevir and disease prognosis). In this study, glutamine is substituted by histidine Table 4.12). Glutamine is substituted mostly by Leucine, Proline (P), and Arginine (R) in the isolatesTables 4.11 and 4.12. According to a study by Tong et al., HCV-infected persons who possess Q47H mutation in their viral NS5B gene had improved prognosis and response to

Daprevir and Mericitabine, while those who did not have the mutation did not improve during therapy(Tong et al., 2014). The implication in the present study is that none of the study subjects possess this mutation (Q47H) and may likely not respond to treatment with Daprevir and Mericitabine (a prodrug of NS5B polymerase inhibitor – PSI-6130).

Substitution of Asparagine (N) with Thymine (T) at position 142 of the finger domain is selected for resistance to NS5B nucleotide inhibitor-Sofosbuvir (SOF). Sofosbuvir has demonstrated high efficacy in HCV-infected patients in combination therapy (Hedskog *et al.,* 2015.

Other sites include A56V, T71I Q67R in which all the subjects had no substitution at those amino acid sites 56, 67 and 71. A56V mutation was not observed in all the groups and most reference sequences around the world except in EF442256 strain from Ireland where Valine (V) was substituted by Proline (P) and JQ679058 from a previous study in Nigeria where valine was substituted by leucine (L)(Table 4.7).

With reference to H77 and other HCV strains in other parts of the world, Q67R, was present in EU849145 from Sri Lanka. But glutamine (Q) was substituted by Glutamate (E) in EU442256 from Ireland, Leucine in GQ153856 from Guinea Bissau, and Serine in JQ 679058 from Nigeria. There were numerous other substitutions that are different from the rest in other patients as well as deletions at amino acid positions 6, 27, 46, 58 and 76. At position 61, Glycine (G) is substituted by Leucine (L), while in the rest of the patients; it is replaced by Tryptophan (W) in G61R/E. In T7N/S, Proline substituted Thymine while the rest had Serine substitution.

Viral infectivity, virulence, disease progression and prognosis in infected individuals as well as antiviral resistance are seen to be affected by several substitutions (amino acids) in the NS5B polymerase. The mechanism of antiviral resistance is mainly achieved through a change in the NS5B conformation that affects the hydrophobic binding of the residues (Lam *et al.,* 2014). This confirms the regulatory role of NS5B phosphorylation in HCV RNA replication, suggesting that the post-translation modification of NS5B protein as observed in this study, regulates both RNA synthesis and viral resistance to NS5B polymerase inhibitors. Hence, amino acid substitutions at the finger loop (domain) of NS5B polymerase in the different genotypes have deleterious effects not only to antivirals but also to pathogenesis of these viral variants in infected individuals, which may affect disease outcome. This further implies that whether an individual is able to clear HCV infection or progress to liver disease depends partly on the type of mutation and position of this mutation on the finger domain of

NS5B polymerase actively involved in the replication HCV RNA and the subtypes the individual is infected with.Phosphorylation of amino acid residues is a post-translational modification that affects the structure of NS5B, which eventually affects the expression and function of that protein in HCV RNA replication and general life cycle. Phosphorylation of proteins plays important role in regulating the structures and functions of proteins.

The circulating HCV subtypesin patients with clinical hepatitis, individuals with HIV and blood donors showed different amino acid substitutions (Table 4.6). These substitutions are mostly at the Finger domain of the NS5B polymerase close to the active site.

The Finger domain is the most variable region of the NS5B polymerase, followed by the Thumb domain while the Palm is the most conserved. The amino acid positions in this finger subdomain range from numbers 1-187 and 228-286 (Chinnaswamy et al., 2008). Amino acid substitutions result from substitutions with one, two or more nucleotide bases.Such a substitution could change a codon to one that encodes a different amino acid and causes a change in the resulting protein. This can seriously affect the protein structure which may completely change the function of the protein (Han et al., 2014).

Molecular characterization of NS5B protein in HCV-infected populations in this study has revealed the presence of multiple Serine/Threonine phosphorylated amino acid residues in HCV NS5B gene. Phosphorylation is the addition of a phosphoryl group $(PO_3)^-$ to a molecule (Lee et al, 2000). Studies have shown that phosphorylation of amino acid Serine at some sites has a profound effect on the severity or outcome of disease (Cohen, 2002). NS5B phosphorylation is an important cellular process in HCV RNA replication that may alter the closed conformational nature of NS5B structure thereby affecting the function and expression of that protein in HCV life cycle. Phosphorylation of amino acid residues in protein is important in regulating the configuration and functions of proteins (Tarrant & Cole, 2009). Owing to multiple phosphate acceptor sites identified in the Serine and Threonine finger loop domain of the NS5B protein, it implies therefore, that hepatitis C virus isolates in the study populations are differently phosphorylated resulting in:

(1) The viral genetic diversity observed in boththe nucleotide and amino acid levels in the study population due to altered nature or conformation of the RdRp enzyme during post-translational modification or phosphorylation, resulting in generation of well adapted HCV variants seen as multiple subtypes (Table 4.1).

(2) Clinical disease manifestation in those showing symptoms of infection and delayed or lack of clinical signs of infection in asymptomatic subjects (blood donors). This could mean that symptomatic HCV-infected individuals with viral proteins phosphorylated at Serine or threonine may or may not proceed to liver disease; or that asymptomatic subjects (blood donors) may clear the virus, not progress to chronic infection or remain asymptomatic all through the period of infection and vice versa. These phosphorylated sites may be good targets for antiviral therapy that will inhibit RNA replication or vaccines that will stimulate immune response against the viral strains.

The pattern of amino acid residue phosphorylation in the various genotypes circulating among study participants suggests that these viral strains with Serine and Threonine phosphorylated sites have similar NS5B conformation (closed or open) due to addition of phosphate groups to the finger domain of the enzyme, leading to similar interactions through hydrogen bond with finger loop or other subdomains of the NS5B (Palm and Thumb).

Although multiple Serine and Threonine phosphorylated sites were observed among study participants, Ser /Thr 52 or 73 sites have been linked to disease progression in a previous study by Toivolaet al., (2004) where Ser 52 and Thr 73 were evidently phosphorylated at very high levels in cirrhoticpatients as compared with non-cirrhotic patients where S33 showed higher phosphorylation.

In the present study, 4 (31%) isolates including one isolate of subtype 5a and three of 2b had either Ser /Thr 52 or 73, out of the 16 isolates found in this group (Table 4.13). Two (15%**)** isolates (subtype 5a) had either Ser /Thr 52 or 73 among the blood donors (Table 4.14), while 7 (54%**)** isolatesall of subtype 5a with either Ser /Thr 52 or 73 were found among individuals with HIV (Tables 4.13- 4. 15) as summarized in table 4.16. However, it is not clear whether the co-infection status is responsible for high phosphorylation in this group. An important deduction though, is that individuals infected with HIV with high Ser /Thr 52 or 73 phosphorylation may have severe disease outcome or progress faster with hepatic pathology.

In general, it could be that lack of these modifications in NS5B protein of each genotype is responsible for delayed symptom, in spite of active RNA replication in blood donors or their presence responsible for clinical disease outcome in some patients with clinical hepatitis. It could also mean that among blood donors with infecting HCV isolates phosphorylated at Ser /Thr 52 or 73 sites, they may eventually progress to liver disease in the course of their infection while clinical patients without phosphorylated residues may or may not recover.

The only limitation here is that this study did not pre-assess the patients to determine the level of liver pathology using either biopsy or Fibro scan but depended on clinical manifestation of disease and indications on request forms hence cannot conclude that these deductions are final. However, this study provides a vital clue into the pathogenesis of HCV among donors, co-infected persons and patients with clinical symptoms depicting HCV infection. Although host's immune response to HCV infection leads to clinical disease symptoms such as jaundice, ascites, ankle edema etc, and influences progression to liver diseases, cirrhosis and hepatocellular carcinoma, supporting the notion that cell - mediated immune responses induce liver injury(Hajarizadeh *et al*., 2013), obviously, viral-related factors and host factors together play a part in acceleration of liver disease (Einav *et al*, 2002).

Sequel to these findings, these post-translational modifications (PTM) such as phosphorylation by protein kinase may have induced structural changes in HCV NS5B proteins which may affect viral RNA synthesis and may influence outcome of HCV infections, resistance to antiviral drugs and response to antiviral therapy in HCV-infected individuals. Hence, there is increased need to identify additional protein characteristics that may answer pertinent questions regarding HCV pathogenesis. In the work done by Han and others,two protein kinase (PRK2) phosphorylation sites; Ser29 and Ser42 of the delta 1 finger loop required for the formation of closed conformation of HCV NS5B and important in synthesis of new HCV RNA were found affecting HCV replication in 460 HCV sequences obtained from HCV database. They demonstrated the critical role of NS5B phosphorylation in HCV RNA replication (Han *et al.*, 2014).

The correlation between increased phosphorylation and disease occurrencehas been reported and this has actually positioned protein kinases that catalyze these processes as important targets for drug designs. Thishas led to the use of kinase enzyme inhibitors for therapeutic purposes (Cohen, 2002; Lee et al., 2000).

The importance of Serine and threonine phosphorylation in NS5B finger loop domain interactions with thumb and palm domains in maintaining the structural integrity of NS5B has been reported in previous study by Chinnaswamy et al (2008). They found that phosphorylation of the Serine residues within the finger domain could result in disruption of its electrical charge, which may cause the site to assume a varied configuration and when this happens, HCV NS5B phosphorylation when regulated would then have

significantconsequence for modifications in the RNA polymerase of the virus. Consequently, that can affect processes like initiation and extension during RNA synthesisand consequently viral infectivity and disease outcome (Chinnaswamy *et a*l., 2010).

HCV proteins are very important in the virus life cycle because viral proteins inhibit antigen presentation through induced interferon negative immune regulation and other processes that ensure its survival and evasion of immune responses, leading to persistent infection of the host. NS5B enzyme lacks the proof-reading capability; hence, it is prone to errors during replication which result in high variation in the gene. Mutations in the finger domain are associated with disease prognosis and lack of response to treatment with interferon and ribavirin, as well as to newly approved antivirals including direct acting antiviral agents (DAAs) -NS5B inhibitors (Nucleoside and Non-nucleoside inhibitors) (Gouklani *et al.,* 2012).In general, disease evolution is believed to be as a result of interactions that exist between HCV and the host cells (Codes *et al*., 2003).

Through this study, the followingare the contributions made to advance knowledge:

- Multiple genotypes of Hepatitis CVirus exist in Ibadan, with predominant subtypes as 1a and 5a; and subtype 5a reported for the first time in West Africa.
- HCV genotypes are distributed differently among the study population, with subtypes 1b and 2b found only among patients with clinical hepatitis, 1a predominant among donors, and 5a predominant among HIV and patients with clinical symptoms.
- Phosphorylated Ser / Thr at amino acid positions 52 and 73, which are viral markers of disease progression, were highest among individuals co-infected with HIV, followed by hepatitis patients, with a statistical significant difference ($p= 0.04$) between the groups.
- Only HCV subtypes 5a and 2b had these Ser / Thr 52 and 73 markers.

CHAPTER SIX

CONCLUSION AND RECOMMENDATIONS

In this study, molecular characterization of HCV NS5B gene has revealed varying degree of HCV diversity among different subpopulations infected with HCV and the differences among these diverse forms in relation to disease outcome. The genotype of hepatitis C virus (HCV) is an important indicator for antiviral and treatment response and may also determine disease course and outcome in infected individuals. Proper identification of HCV genotype will enhance optimal treatment routine in patients. During therapy, monitoring of HCV RNA is used to ascertain treatment duration in HCV infection. Molecular assays are used to assess treatment response and the end of treatment or sustained virological response (SVR)i.e. undetectable HCV RNA several months after end of therapy. Accurate subtype determination coupled with drug resistance profiling of patients and donors' viral population (diversity) may considerably guide the optimal choice of drug, especially in difficult - to - treat cases. Although, there has been a lot of improvement in the treatment of infection since inception of direct acting antivirals (DAAs) known as IFN-free therapy, however, high cost associated with this class of antiviral agents and difficulty in procurement have made IFN-based regimen to remain standard therapy in many parts of the world especially in developing nations. IFN-based therapy may remain suboptimal due to factors as host immune response, IL28B phenotype, liver disease such as cirrhosis, HCC and viral factors as HCV RNA (viral load) or HCV genotypes. As a result, HCV mediated-IFN resistance may then play an important role in treatment failure, emphasizing the need for HCV genotyping irrespective of the setting.

In this study, the diversity of HCV among infected populations has been shown. Hitherto, a limited number of molecular studies have been carried out to determine HCV population in Nigeria. Some of the studiescarried out in the past involved small populations from North Central and Lagos in Southwest Nigeria. The results from each location revealed genotypes 1 and 2 as predominant strains. However, these study populations were neither stratified in terms of prevalent strains nor strains linked to disease outcome in the study groups, resulting in a wide research gap on specific population–based genotypes and the involvement of these genotypes in disease progression or outcome in infected individuals. These gaps were addressed in this study in which valuable data was gathered and knowledge advanced on HCV diversity among different subgroups in Nigeria.

Detection of circulating HCV strains in this study has a tremendous impact on HCV epidemiology in Nigeria and also a global significance with regards to therapy and design of a broadly protective vaccine against hepatitis c virus. Based on presence of multiple HCV genotypes and viral markers of disease present as indicated by phosphorylated serine and threonine residues, the following recommendations are therefore suggested:

1. Every infected person should undergo genotyping test for proper target treatment and resistance monitoring, in addition to HCV RNA quantification (viral load).
2. Individuals infected with subtypes 1b, 2b and 5a should be treated early without delay, with or without symptoms.
3. Equipped blood banks and screening centers should be sited in urban and rural areas to facilitate proper blood screening before transfusion.
4. Due to presence of multiple HCV genotypes circulating in some parts of the world, global effort towards the development of broadly- protective vaccine against hepatitis c virus should be intensified.
5. There should be unhindered access to direct acting antivirals in developing nations where HCV burden is on the increase yet receiving little or no attention.
6. Countries should come out with policies that will enable infected individuals obtain treatment with new drugs at subsidized costs.
7. There should be increased public awareness about the risk factors and transmission of HCV infection.
8. Increased community-based studies are needed in Nigeria to ascertain theburden of HCV infection as well asemerging strains, since its distribution varies.
9. All strategies involved at local, national and international levels in HIV care should be deployed in HCV care to ensure large treatment coverage.
10. As a rule, every family member of infected person should be screened as well as intending couples for marriage, as it is the case with HIV.
11. All healthcare facilities should be strengthened to facilitate medical access to care and antiviral drugs.

REFERENCES

American Association for the study of the Liver/Infectious Disease Society of America/International Antiviral Society –USA (2015). Recommendations for testing, managing, and treating hepatitis C. HCV Guidance. 2015.

Acheson N.H (2007). Fundamentals of Molecular Virology, *John Wiley and Sons Inc.*ISBN-10 0-471-335151-2.

Adewole OO, Anteyi E, Ajuwon Z, Wada I, Elegba F, Ahmed P, Betiku Y, Okpe A, Eze S, *et al* (2009). Hepatitis B and C co-infection in Nigerian patients with HIV infection. *J Infect Dev Ctries* 3, 369-375.

Afolabi AY, Abraham A, Oladipo EK, Fagbami AH (2012).Hepatitis C Virus in Potential Blood Donors in Ibadan, Nigeria.*Global Advanced Research Journal of Microbiology* (ISSN: 2315-5116) Vol. 1(9) pp. 155-159.

Agnello V, De Rossa FG (2004). Extrahepatic disease manifestations of HCV infection: Some current issues. *J Hepatol* 40(2) : 34-52.

Alter MJ (1997). Epidemiology of HCV.*Hepatology.*1997, Vol 26 (pg 62S-65S).

Alter MJ, Kruszon-Moran D, Nainan OV *et al.*(1999). The prevalence of hepatitis C virus infection in the United States, 1988 through 1994. *New England Journal of Medicine* 341(8):556-62.

Antaki N, Haddad M, Kebbewar K *et al.* (2009). The unexpected discovery of a focus of hepatitis c virus genotype 5 in a Syrian province. *Epidemiology of Infection* 2009; 137:79-84.

Antaki N, Craxi A, Kamal S *et al*. (2010). The neglected hepatitis c virus genotypes 4, 5 and 6: An International consensus report. *Liver International,*: 30: 342-55.

Arai M, Tokunaga Y, Takagi A, Tobita Y, Hirata Y, *et al.* (2013) Isolation and Characterization of Highly Replicable Hepatitis C Virus Genotype 1a Strain HCV-RMT. PLoS ONE 8(12): e82527.

Argentini C, Genovese D, Dettori S, Rapicetta M (2009). HCV genetic variability: from quasispecies evolution to genotype classification. *Future Microbiol* 4: 359-373

Ashfaq UA, Javed T, Rehman S, Nawaz Z, Riazuddin S (2011).An overview of HCV molecular biology, replication and immune responses.*VirologyJournal;* 8: 161

Ayolabi CI, Taiwo MA, Omilabu SA, Abebisi AO, Fatoba OM (2006). Sero-prevalence of Hepatitis C Virus among Blood Donors in Lagos, Nigeria. *African Journal of Biotechnology.*, 5(20): 1944-1946

Ayesh BM, Zourob SS, Abu-Jadallah SY, Shemer- Avna Y (2009). Most Common Genotypes and Risk Factors for HCV in Gaza Strip: A Cross Sectional Study. *Virology Journal*; 6.1- 7

Azocar J *et al* (2003). MHC class II genes in HCV viral clearance of hepatitis C infected Hispanic patients. *Human Immunology*2003; 64(1):99-102. 112.

Balogun TM, Emmanuel S, Wright KO (2010). Hepatitis C virus co-infection in HIV positive patients.*Nig QJ Hosp Med* 20, 117-120.

Bartenshchlager R, Lohmann V (2000). Replication of hepatitis C. *Liver International*, 29(Suppl 1): S74-S81.

Bartenschlager R, Lohmann V, Penin F (2013). The molecular and structural basis of advanced antiviral therapy for hepatitis c virus infection. *Nature Review Microbiology.*2013; 11: 482-96.

Bedossa P, Poynard T(1996). METAVIR Coorporative Study Group. An algorithm for the grading of activity in chronic hepatitis C. *Hepatology*.1996; 24: 289-293.

Blackard JT, Ma G, Limketkai BN, Welge JA, Dryer PD, Martin CM, Hiasa Y, Taylor LE, Mayer KH, Jamieson DJ, Sherman KE: Variability of the polymerase gene (NS5B) in hepatitis C virus-infected women. *Journal of Clinical Microbiology* 2010, 48(11):4256-4259.

Blom N, Gammeltoft S, Brunaks S(1999). Sequence and Structural-based prediction of eukaryotic protein phosphorylation sites. For generic predictions. *Journal of Molecular Biology*: 294 (1): 1351-1362.

Blom N, Sicheritz-Ponten T, Gupta R, Gammeltoft S, Brunak S (2004). Prediction of post-translational glycosylation and phosphorylation of proteins from amino acid sequence. *Proteomics*: 4(6): 1633-49 review.

Borgia, S. M. *et al* (2018).Identification of a Novel Hepatitis C Virus Genotype From Punjab, India: 314 Expanding Classification of Hepatitis C Virus into 8 Genotypes. *The Journal of infectious diseases*218: 1722-1729, doi:10.1093/infdis/jiy401.

Boyer JL, Chang EB, Collyar DE, *et al* (2002). NIH Consensus Statement on Management of Hepatitis C: 2002. *NIH Consensus State Science Statements* 19(3):1-46.

Brimacombe CL, Wilson GK, Hübscher SG, McKeating JA, Farquhar MJ(2014).A role for CD81 and Hepatitis C Virus in Hepatoma Mobility *Viruses*, *6*, 1454-1472

Busch MP, Shafer KA. Acute-phase hepatitis C virus infection: implications for research, diagnosis, and treatment. *Clin Infect Dis*. 2005;40: 959-61.

Busch MP, Murthy KK, Kleinman SH, et al. Infectivity in chimpanzees (Pan troglodytes) of plasma collected before HCV RNA detectability by FDA-licensed assays: implications for transfusion safety and HCV infection outcomes. *Blood*.2012; 119: 6326-34.

Business Wire. Hepatiq receives FDA clearance. Hepatiq.com. 2014 Dec 19; Accessed: January 11, 2015.

Cacoub P, Poynard T, Ghillani P, et al. Extrahepatic manifestations of chronic hepatitis C. MULTIVIRC Group. Multidepartment Virus C. *Arthritis Rheum*. 1999 Oct. 42(10):2204-12.

Cacoub P, Renou C, Rosenthal E (2000). The GERMIVIC Group. Extrahepatic manifestations associated with hepatitis C virus infection. A prospective multicenter study of 321 patients. *Medicine (Baltimore)*. 2000 Jan. 79(1):47-56.

Campollo RO. (2002). Hepatitis C virus infection and alcohol. *Review Gastroenterology Mex*2002; 67(Suppl 2):S80-3.

Candotti D, Temple J, Sarkodie F, Allain J (2003). Frequent recovery and braod genotype 2 diversity characterize hepatitis c virus infection in Ghana West Africa. *Journal of Virology* 77, 7914-7923

Castilho MC, Martins AN, Horbach IS, Perez R deM, Figueiredo FAF, Pinto P de T, Nabuco LC, Lima DB, Tanuri A, Porto LC, Ferreira Junior Oda C (2015).Association of hepatitis C virus NS5B variants with resistance to new antiviral drugs among

untreated patients.*Mem Inst Oswaldo Cruz*, Rio de Janeiro, Vol. *106*(8): 968-975.

Ceci O, Margiotta M, Marello F et al (2001). Vertical transmission of hepatitis c in a cohort of 2447 HCV seronegative pregnant women: a 24 -month prospective study. *Journal of Pediatric Gastroenterol ogy Nutrition*, 2001; 33: 570-575.

Center for Disease Control and Prevention(CDC) (2012). Hepatitis C. Online review.

Chen SL (2006). Int J Med Sci. 2006; 47-52

Chinnaswamy S, Yarbrough I, Palaninathan S, Kumar CT, Vijayaraghavan V, Demeler B, Lemon SM, Sacchettini JC, Kao CC (2008). A locking mechanism regulates RNA synthesis and host protein interaction by the hepatitis C virus polymerase. *J. Biol .Chem*. 283:20535–20546.

Chinnaswamy S, Murali A, Li P, Fujisaki K, Kao CC (2010). Regulation of de novo-initiated RNA synthesis in hepatitis C virus RNA-dependent RNA polymerase by intermolecular interactions. *Journal of Virology.* 84:5923–5935.

Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, et al. (1989) Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* 244: 359–362.

Chung RT (2005). Acute Hepatitis C infection.*Clinical Infectious Disease,* 2005.

Codes L, de Freitas L.A.R, Santos- Jesus R, Vitvitski R L, Silva L.K, Trepo C, Reis M.G, Paraná R (2003). Comparative Study of Hepatitis C Virus Genotypes 1 and 3 in Salvador, Bahia. Brazilian Journal Infectious Disease, 2003; 7

Cohen (2002). Protein kinases-the major drug targets of the twenty- first century? *Nat Rev DrugDiscov* 2002; 1: 309-315.

Cooper BW, Krusell A, Tilton RC (1992).Seroprevalence of antibodies to hepatitis c virus in high-risk hospital personnel. *Infect Control Hosp Epidemiol* 13: 82-85.

Crane M. FDA approves first genotyping test for patients with HCV. Medscape Medical News from WebMD. 2013 Jun 20; Accessed: June 23, 2013.

Centers for Disease Control and Prevention (CDC).National Notifiable Diseases Surveillance System (NNDSS). Hepatitis C, acute: 2012 case definition.

Chukwurah EF, Ogbodo SO, Obi GO (2005). Seroprevalence of Hepatitis C Virus (HCV) among blood donors in South Eastern States of Nigeria. Biomedical research; 16,133-135.

Chung RT. Acute hepatitis C virus infection. Clin Infect Dis. 2005; 41 Suppl 1:S14-7.

Dhawan VK (2016). Hepatitis C. *Medscape*, www. emedicine.medscape.com

Davis GL, Balart LA, Schiff ER *et al.,* (1989).Treatment of chronic hepatitis c with recombinant interferon alfa. A multicenter randomized, controlled trial. Hepatitis Interventional Therapy Group.*New England Journal of Medicine* .1989, 321(22): 1501-6.

De Francesco R, Migliaccio G. Challenges and successes in developing new therapies for hepatitis C. *Nature* 2005; 436: 953-960

Della Rossa A, Tavoni A, Baldini C, Bombardieri S (2003). Mixed cryoglobulinemia and hepatitis C virus association: ten years later. *Israel Medical Association Journal*. 3(6):430-4.

D'Heygere F, George C, Van Vlierberghe, *et al.* (2011). Efficacy of interferon-based antiviral therapy in patients with chronic hepatitis C infected with genotype: A meta-analysis of two large prospective clinical trials. *Journal of Medical Virology* .201; 83: 815-9.

Egah, DZ, Mandong BM, Iya D, Gomwalk NE, Audu ES, Banwat EB, Onile BA (2004). Hepatitis C Virus Antibodies among Blood Donors in Jos, Nigeria. Annals of Afr. Med., 3(1): 35-37.

Einav S, Koziel M (2002). Immunopathogenesis of Hepatitis C Virus in the Immunosuppressed Host. Transpl Infect Dis 4(2): 85-92.

El-Zayadi AR (2006). Heavy smoking and liver. *World Journal of Gastroenterology*2006; 12 (38):6098-101.

Everson GT, Weinberg H. *Living With Hepatitis C: A Survivor's Guide*. 3rd ed. Long Island City, New York: Hatherleigh Health

Farci P, Shimoda A, Coiana A *et al*. (2000). The Outcome of acute hepatitis c predicted by the evolution of the viral quasispecies. *Science*, 2000; 288:339-394.

Farci *et al* (2012). NIAID/ NIH NEWS, July 23, 2012.

Fasola F, Kotila T, Akinyemi JO (2008).Trends in Transfusion transmitted viral infections from 2001-2006 in Ibadan Nigeria. *Intervirology* 51:427- 431

Forbi JC, Purdi MA, Campo DS, Vaughan G, Dimitrova GE, Ganova-Raeva LM, Xia GL, Khudyakov YE (2012).Epidemic history of hepatitis C virus infection in two remote communities in Nigeria, West Africa.9o- 2012 Jul;93(Pt 7):1410-21.

Gao X, Cui Q, Shi X, Su J, Peng Z, Chen X, Lei N, Ding K, Wang L, Yu R, Wang N, (2011). Prevalence and trend of hepatitis C virus infection among blood donors in Chinese mainland: a systematic review and meta-analysis.BMC Infectious Diseases, 11:88

Gededzha MP, Selabe SG, Kyaw T, Rakgole JN, Blackard JT, Mphahlele MJ (2012).Introduction of new subtypes and variants of hepatitis C virus genotype 4 in South Africa.J Med Virol.84(4):601-7.

Gerlach JT, Diepolder HM, Jung MC *et al.,*(1999). Recurrence of hepatitis C virus after loss of virus-specific $CD4^{+}$ response in acute hepatitis C. Gastroenterol 1999; 117: 93-941

Gerlach JT, Diepolder HM, Zachoval R *et al.*, (2003). Acute hepatitis C: high rate of both spontaneous and treatment-induced viral clearance. *Gastrenterology,* 2003; 125: 80-88

Goldberge D, Cameron S, McMenamin J (1998). Hepatitis C virus antibody prevalence among injecting drug users in Glasgow has fallen but remains high. *Commun Dis Public Health* 1:95-97.

Goh PY, Tan YJ, Lim SP, Tan YH, Lim SG, Fuller-Pace F, Hong W (2004). Cellular RNA helicase p68 relocalization and interaction with the hepatitis C virus (HCV)

NS5B protein and the potential role of p68 in HCV RNA replication. *Journal of Virol*ogy, 78:5288–5298.

Gopalsamy A, Lim K, Ciszewski G, Park K, Ellingboe JW, Bloom J, Insaf S, Upeslacis J, Mansour TS, Krishnamurthy G, *et al*.(2004).Discovery of pyrano[3,4-b]indoles as potent and selective HCV NS5B polymerase inhibitors. *J Med Chem*47:6603-6608.

Gordon SC, Bayati N, Silverman AL (1998). Clinical outcome of hepatitis C as a function of mode of transmission. *Hepatology*.28(2):562-7.

Graham CS, Wells A, Edwards EM, Herren T, Tumilty S, Stuver SO, Samet JH, Nunes D, Horsburgh CR Jr, Koziel MJ (2007). Effect of exposure to injection drugs or alcohol on antigen-specific immune responses in HIV and Hepatitis C Virus coinfection. J Infect Dis 15; 195 (16) : 847-56.

Graham CS (2015). Hepatitis C and HIV co-infection: Closing the gaps. JAMA 313: 217-8.

Grakoui A, Shoukry NH, Woolard DJ *et al*(2003). HCV persistence and immune evasion in the absence of memory T cell help. *Science*, 2003; 302: 659-662.

Gremion C, Cerny A. Hepatitis C virus and the immune system: a concise review. *Review of Medical Virology* 2005; 15(4):235-68

Gretton SN, Taylor AI, McLauchlan J. Mobility of the hepatitis C virus NS4B protein on the endoplasmic reticulum membrane and membraneassociated foci. *J Gen Virol*.2005;86: 1415–1421.

Hajarizadeh B, Grebely J, & Dore GJ (2013).Epidemiology and natural history of HCV infection.*Nature Reviews. Gastroenterology & Hepatology*, *10*(9), 553–62.

HALT-C (Hepatitis C Antiviral Long-term Treatment against Cirrhosis) trial website. National Institute of Diabetes and Digestive and Kidney Diseases. January 11, 2015.

Han SH, Kim SJ, Kim EJ, Kim TE, Moon JS, Kim GW, Lee SM, Cho K, Yoo JS, Woo SS, Rhee JK, Han SH, Oh JW(2014). Phosphorylation of Hepatitis C Virus RNA Polymerases Ser29 and Ser42 by Protein Kinase C-Related Kinase 2 Regulates Viral RNA Replication.*J Virol.* 88(19): 11240–11252.

Hanafiah KM, Groeger J, Flaxman AD, Wiersma ST (2012). Global epidemiology of hepatitis C virus infection: New estimates of age-specific antibody to hepatitis C virus seroprevalence. *Hepatology*.doi: 10.1002.

Harris, H.J.; Clerte, C.; Farquhar, M.J.; Goodall, M.; Hu, K.; Rassam, P.; Dosset, P.; Wilson, G.K.; Balfe, P.; IJzendoorn, S.C.; *et al.*(2013). Hepatoma polarization limits CD81 and hepatitis C virus dynamics. *Cell. Microbiol.15*, 430–445

Hedskog C, Dvory Sobol H,Gontcharova V,Martin R, Ouyang W, Han B, Gane EJ, Brainard D Hyland RH,Miller MD,MOH, Svawukaia E (2015). Evolution of the HCV viral population from patients with S282T detected at relapse after Sofosbuvir monotherapy. *J Viral Hepat*.

Heim, M. H. (2013). Innate immunity and HCV. *Journal of Hepatology*, 2012.10.005.

Houghton M (1996). Hepatitis C virus. In: Fields BN, Knipe DM, Howley PM, eds. *Fields Virology, 3rd ed.* Philadelphia, Lippincott - Raven, 1996:1035-1058.

Iles JC, Raghwani J, AbbyHarrison GL , Pepin JC, Djoko CF, Tamoufe U, LeBreton M, Schneider BS, Fair J.N, Tshala FM, Kayembe P.K, Muyembe J.J, Edidi-Basepeo S, Wolfe N.D, Kennard J (2014). Hepatitis Fact, Dec19, 2014. *Understanding Hepatitis-About.com*

Innis MA and Gelfand DH (1990). PCR Protocols: A guide to Methods and Applications. *Academic Press Inc*

Irshad M, Dhar I (2006). Hepatitis C virus core protein: an update on its molecular biology, cellular functions and clinical implications. *Medical Princ Pract* 15: 405–416.

Irshad M, Mankotia SD, Irshad K (2013).An insight into the diagnosis and pathogenesis of hepatitis C virus infection.*World Journal of Gastroenterology*. 2013 Nov 28; 19(44): 7896–7909.

Isa AH, Hassan A, Mamman AI Bababdoko AA, Muktar HM, Ahmed AJ (2010).Seroprevalence of hepatitis C virus antibodies among blood donors in Ahmadu Bello University Teaching Hospital, Kaduna.*African Journal of Clinical Experimental Microbioliogy*; 11: No. 2; 75-78.

Jacobson IM, Gordon SC, Kowdley KV, et al.(2013). Sofosbuvir for hepatitis C genotype 2 or 3 in patients without treatment options. *N Engl J Med* 368(20):1867-77.

Jakubiec A, Tournier V, Drugeon G, Pflieger S, Camborde L, Vinh J, Hericourt F, Redeker V, Jupin I(2006). Phosphorylation of viral RNA-dependent RNA polymerase and its role in replication of a plus-strand RNA virus. *J. Biol. Chem*. 281:21236–21249.

Jaspe RC, Sulbarán YF, Sulbarán MZ, Loureiro CL, Rangel HR, Pujol FH(2012). Prevalence of amino acid mutations in hepatitis C virus core and NS5B regions among Venezuelan viral isolates and comparison with worldwide isolates. *Virology Journal*, 9:214.

Jezequel C, Bardou-Jacquet E, Desille Y, et al(2015). Survival of patients infected by chronic hepatitis C and F0F1 fibrosis at baseline after a 15 year follow-up. Presented at: 50th Annual Meeting of the European Association for the Study of the Liver (EASL); Vienna, Austria.

Jover R, Perez-sSerr J, de Vera F *et al*.(2001).Infection by genotype 5a of HCV in a district of southeast Spain.*American Journal of Gastroenterol*96: 3042-3.

Kang SM, Choi JK, Kim SJ, Kim JH, Ahn DG, Oh JW (2009). Regulation of hepatitis C virus replication by the core protein through its interaction with viral RNA polymerase. *Biochemical and Biophysics Research Communications* 386:55–59.

Kapoor A, Simmonds P, GeroldG,Qaisar N, Jain K, Henriquez JA, Firth C, Hirschberg DL, Rice CM, Shields S, Lipkin WI (2011). Characterization of a canine homolog of hepatitis C viru*s. PNAS* 108 no. 28 11608-11613

Kato N (2000). Genome of Human hepatitis virus (HCV): gene organization, sequence diversity, and variation. *Microb. Comp. Genomics* 5 (3): 129-51. PMID 11252351

Kazmierczak J, Pawelczyk A, Cortes KC, Radkowski M (2013). Seronegative hepatitis C virus Infection.*Arch.Immunol. Ther.Exp.*

Kenny-Walsh (1999). Clinical outcomes after hepatitis c infection from contaminated anti-D immune globin. Irish Hepatology Research Group. *New England Journal of Medine,* 340: 1228-1233.

Khakoo SI, Thio CL, Martin MP *et al*., (2004). HLA and NK cell inhibitory receptor genes in resolving hepatitis C virus infection. *Science,* 305: 872-874.

Kim WR. The burden of hepatitis C in the United States. *Hepatology*. 2002 Nov. 36(5 Suppl 1):S30-4.

Kim SJ, Kim JH, Kim YG, Lim HS, Oh JW (2004). Protein kinase C-related kinase 2 regulates hepatitis C virus RNA polymerase function by phosphorylation. *J. Biol. Chem.*279:50031–50041.

Kim SJ, Kim JH, Sun JM, Kim MG, Oh JW (2009). Suppression of hepatitis C virus replication by protein kinase C-related kinase 2 inhibitors that block phosphorylation of viral RNA polymerase. J. Viral Hepat. 16:697–704.

Kim MG, Moon JS, Kim EJ, Lee SH, Oh JW (2012). Destabilization of PDK1 by Hsp90 inactivation suppresses hepatitis C virus replication through inhibition of PRK2-mediated viral RNA polymerase phosphorylation. *Biochemical and Biophysics Research Communications.*421:112–118.

Kimura Y, Hayashida K, Ishibashi H, Niho Y and YanagiY (2000). Anti-body free virion titer greatly differs between hepatitis C virus genotypes. *J Med Virol* 61, 37-43.

Kishihara Y *et al.,* (2001). Human T lymphotropic virus type I infection influences hepatitis C virus clearance. *Journal of Infectious Disease*2001; 184:1114-9.

Kiyosawa K, Tanaka E, Sodeyama T *et al* (1994). Transmission of hepatitis C in an isolated area in Japan: community-acquired infection. The South Kiso Hepatitis Study Group. *Gastroenterology* 106(6):1596-602.

Kohla M, Bonacini M (2006). Pathogenesis of hepatitis c virus infection. *Minerva Gastroenterol Dietol* 52(2): 107-23.

Kolykhalov AA, Agapov E V, Blight K J, Mihalik K, Feinstone SM, Rice C M (1997). Transmission of hepatitis C by intrahepatic inoculation with transcribed RNA. *Science* 277 (5325), 570-574

Kolykhalov A. A., Mihalik K, Feinstone S. M, RiceC. M (2000). Hepatitis C virus-encoded enzymatic activities and conserved RNA elements in the 3′ non-translated region are essential for virus replication*in vivo*. *Journal of Virology.* 74**:**2046-2051.

Kuiken C, Yusim K, Boykin L, Richardson R (2005).The Los AlamosHCV sequence database.*Bioinformatics* 21:379-384.

Lawitz E, Sulkowski MS, Ghalib R, *et al* (2004). Simeprevir plus sofosbuvir, with or without ribavirin, to treat chronic infection with hepatitis C virus genotype 1 in non-responders to pegylated interferon and ribavirin and treatment-naive patients: the COSMOS randomised study. *Lancet*384(9956):1756-65

Lawitz E, Mangia A, Wyles D, *et al* (2013). Sofosbuvir for previously untreated chronic hepatitis C infection. *N Engl J Med*. 368(20):1878-87.

Lavanchy D (2009). The global burden of hepatitis C. *Liver International,* 29(Suppl 1):74-81.

Lauer, G. M. (2013). Immune responses to hepatitis C virus (HCV) infection and the prospects for an effective HCV vaccine or immunotherapies. *Journal of Infectious Diseases*, 207(SUPPL.1).

Lee JC, Kumar S, Griswold DE, Underwood DC, Votta BJ, Adams JL (2000).Inhibition of P38 MAP kinase as a therapeutic strategy. ImmunPharma-col 47: 85-201.

Lee PY, Costumbrado J, Hsu CY, Kim YH (2012). Agarose gel electrophoresis for the separation of DNA fragments.*Journal of Visualized Experiment*.JoVE, 2012. 20 ;(62): 3922

Legrand-Abravanel F, Sandres-Sauné K, Barange K, Alric L, Moreau J, Desmorat P, Vinel JP, Izopet J (2004).Hepatitis C virus genotype 5: epidemiological characteristics and sensitivity to combination therapy with interferon-alpha plus ribavirin. *J Infect Dis*189(8)**:**1397-1400.

Le Guillou-Guillemette H, Vallet S, Gaudy-Graffin C, Payan C, Pivert A, Goudeau A, Lunel- Fabiani F (2007). Genetic diversity of Hepatitis C Virus: Impact and Issues in antiviral therapy. *World Journal of Gastroenterology*13(17): 2416-2426.

Levrero M (2006).Viral hepatitis and liver cancer: the case of hepatitis C. *Oncogene* (2006) 25, 3834–3847. doi:10.1038/sj.onc.1209562

Liang TJ, Rehermann B, Seeff LB, Hoofnagle JH (2000). Pathogenesis, Natural History, Treatment and Prevention of Hepatitis C. *Ann Intern Med.* 2000; 132: 296-305. Doi:10.7326/0003-4819.

Li YP, Ramirez S, Mikkelsen L, Bukh J (2015). Efficient infectious cell culture systems of the hepatitis C virus (HCV) prototype strains HCV-1 and H77.*Journal of Virology*.

Lin W, Kim SS, Yeung E, Kamegaya Y, Blackard JT, Kim KA, Holtzman MJ, Chung RT (2006).Hepatitis C virus core protein blocks interferon signaling by interaction with the STAT1 SH2 domain. *J Virol*80:9226-9235.

Lindenbach BD and Rice CM (2005). Unravelling hepatitis C virus replication from genome to function. *Nature*, 436: 18, 2005.

Lok AS, Gardiner DF, Lawitz E, *et al.,*(2012). Preliminary study of two antiviral agents for hepatitis c genotype 1. *New England Journal of Medicine* 366: 216-224.

Ly KN, Xing J, Klevens RM, Jiles RB, Ward JW, Holmberg SD (2012). The increasing burden of mortality from viral hepatitis in the United States between 1999 and 2007. *Ann Intern Med* 156(4):271-8.

Mandal A (2015).Hepatitis C History.*News Medical.* Life Science and Medicine, 9: 2015.

Marascio N, Torti C, Liberto CM, Forca A (2014). Update on different aspects of HCV variability: focus on NS5B polymerase. *BMC Infectious Diseases* 14(Suppl 5).

Markov PV, Pepin J, Frost E, Deslandes S, Labbe A, Pybus OG (2009). Phytogeography and Molecular Epidemiology of Hepatitis C Virus genotype 2 in Africa. Journal of General Virology 90 :2086-2096. DOI 10, 1099/vir.0.011569-0.

Martell M, Esteban J, Quer J, Genesca J, Weiner A *et al.*(1992). Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. *Journal of Virology* 66:3225-3229. 9.

Maticic M. Lichen planus in hepatitis C virus infection: an early marker that may save lives. *Acta Dermatovenerol Alp Panonica Adriat*. 2007 Mar. 16(1):3-6.

McCombs J, Matsuda T, Tonnu-Mihara I, *et al.*(2014). The risk of long-term morbidity and mortality in patients with chronic hepatitis C: results from an analysis of data from a Department of Veterans Affairs Clinical Registry. *JAMA Internal Medicine* 174(2):204-12.

Mohd Hanafiah, K, Groeger, J, Flaxman, AD, Wiersma, ST (2013).Global epidemiology of hepatitis C virus infection: new estimates of age-specific antibody to HCV Seroprevalence. *Hepatology* 57, 1333–1342.

Moradpour D, Brass V, Gosert R, Wölk B, Blum H E (2002). Hepatitis C: molecular virology and antiviral targets.Trends in Molecular Medicine 8:476-482.3.

Moradpour D, Penin F, Rice C.M. (2007). Replication of hepatitis C virus. *Nature Reviews Microbiology* 5:453-463.2.

Mostafa A, Tailor SM, el-Daly M, el-Hoseiny M, Bakr I, Arafa N, Thiers V, Rimlinger F, Abdel-Hamid M, Fontanet A, Mohamed MK (2010). 'Is the hepatitis C virus epidemic over in Egypt? Incidence and risk factors of new hepatitis C virus infections .*Liver International* 13(4): 560-566.

Mukherjee S (2012). Hepatitis C.*Medscape.* 2012.

Nakashima K, Ikematsu H, Hayashi J, Kishihara Y, Mutsutake A, Kashiwagi S (1995). Intrafamilial transmission of hepatitis-C virus among the population of an endemic area of Japan. *JAMA* 274(18):1459-61.

Nasu A, Marusawa H, Ueda Y, Nishijima N, Takahashi K *et al*(2011). Genetic Heterogeneity of Hepatitis C Virus in Association with Antiviral Therapy Determined by Ultra-Dee Sequencing. Plos One 6(9):e24907.

Niederau C, Lange S, Heintges T (1998). Prognosis of chronic hepatitis C: results of a large, prospective cohort study. *Hepatology* 28(6):1687-1695.

Njouom R, Frost E, Deslandes S, Mamadou-Yaya F, Labbé AC, Pouillot R, Mbélesso P, Mbadingai S, Rousset D, Pépin J. (2009). Predominance of hepatitis C virus

genotype 4 infection and rapid transmission between 1935 and 1965 in the Central African Republic. *Journal General Virology.* 90 (Pt 10):2452-6.

Nwankiti OO, Ndako JA, Echeonu GO, Olabode AO, Nwosuh CI, Onovoh EM, Okeke LA, Akinola JO, Duru BN *et al* (2009). Hepatitis C virus infection in apparently healthy individuals with family history of diabetes in Vom, Plateau state Nigeria. *Virol J* 6, 110.

Ohno O, Mizokami M, Wu RR, SalehMG, Ohba K, Orito E, Mukaide M, Williams R, Lau JY *et al.* (2007). New hepatitis C virus (HCV) genotyping system that allows for identification of HCV genotypes Ia, 1b, 2a, 2b,3a, 3b 4,5 and 6a.*J ClinMicrobiol*35(1): 201-207.

Okwuraiwe, AP, Salu OB, Anomneze E, Audu, RA, Ujah, IAO(2014).Hepatitis C virus genotypes and viral ribonucleic acid titers in Nigeria. *J Gastroenterology, 2014.*

Onakewhor JU & Okonofua FE (2009).Seroprevalence of Hepatitis C viral antibodies in pregnancy in a tertiary health facility in Nigeria.*Niger J Clin Pract* 12, 65-73.

Oni AO and Harrison TJ (1996).Genotype of hepatitis C Virus in Nigeria.*J Med Virol*49: 178-186.

Op De Beeck A, Dubuisson J (2003).Topology of hepatitis C virus envelope glycoproteins.*Rev. Med.Virol*13 (4): 233-41. doi: 101002/rmv.391.

Osella AR, Misciagna G, Leone A, Di Leo A, Fiore G (1997).Epidemiology of hepatitis C virus infection in an area of Southern Italy. *Journal of Hepatology* 27(1):30-5.

Otegbayo JA, Taiwo BO, AkingbolaTS,Odaibo GN, Adedapo KS, Penugonda S, Adewole FI, Olaleye DO, Murphy R, Kanki P (2008). Prevalence of hepatitis Band C Seropositivity in Nigerian Cohort of HIV-infected Patients.*Annals of Hepatology* 7(2): 152-156.

Pawlotsky J (2006). Hepatitis C Virus population dynamics during infection.*Curr Top Microbiol Immunol* 299: 261-284. Pawlotsky JM. Therapy of hepatitis C: from empiricism to cure. *Hepatology* 2006; 43 (Suppl 1):S207–S220.

Pawlotsky JM. Therapy of hepatitis C: from empiricism to cure. *Hepatology* 43 (Suppl 1):S207–S220.

Payan C, Roudot-Thoraval F, Marcellin P *et al* (2005). Changing of hepatitis c virus patterns in France at the beginning of te third millennium: The GEMHEP GenoCII study. *Journal of Viral Hepatology* 12: 405-13.

Penin F, Dubuisson J, Rey FA, Moradpour D, Pawlotsky JM (2004).Structural biology of hepatitis C virus.*Hepatology*, 39(1)**:**5-19.

Pileri P, Uematsu Y, Campagnoli S, Galli G, Falugi F, Petracca R *et al* (1998). Binding of hepatitis C virus to CD81. *Science* 282: 938-941

Pineda JA, Caruz A, Rivero A, *et al.*(2010). Prediction of response to pegylated interferon plus ribavirin by IL28B gene variation in patients coinfected with HIV and hepatitis C virus. *Clinical Infectious Disease* 51(7):788-95.

Pouillot R, Lchenal G, Pybus OG, Rousset D, Njouom R(2008b).Variable epidemic histories of hepatits C virus genotype 2 infection in West Africa and Cameroun. *Infection Genetic Evolution* 8: 676-681

Purcell RH. Hepatitis C virus. In: Webster RG, Granoff A, eds. *Encyclopedia of Virology*. London, Academic Press Ltd, 1994:569-574.

Ramalho F (2003). Hepatitis C virus infection and liver steatosis.*Antiviral Res*60(2)**:**125-127.

Ramia S, Eid-Fares J (2006). Distribution of hepatitis C virus genotypes in the Middle East.*Int J Infect Dis*10(4)**:**272-277.

Shapiro CN, Tokars JI, Chamberland ME (1996). The American Academy of Orthopedic Surgeons Serosurvey Study Committee. Use of hepatitis B vaccine and infection with hepatitis B and C among orthopaedic surgeons.*Journal of Bone Joint Surgery* 78-A: 1791-1800.

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullins KB, Erlich HA (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase.*Science* 239: 487-491.

Saitou N, Nei M (1987).The Neighbor-Joining Method-a new method for reconstructing phylogenetic trees. *Mol Biol Evol*. 4: 406-425

Salami RO, Von Meding JK, Giggins H (2016).Assessing Habitats of Vulnerability in African Cities: A case of poverty housing in Ibadan metropolis. http:// PMsymposium.umd.edu. Doi: 10.13140/RG.2.1.3185

Samsouk M, Lauer GM, Casson D *et al* (2003). Spontaneous resolution of chronic hepatitis c disease after withdrawal of immunosuppression. *Gastroenterology* 124: 1946-1949.

Sanaullah K, Sobia A, Sultan A, Shahid NK, Sumaira S, Ijaz A, Muhammad B, Sami S(2011).Molecular epidemiology of HCV among health care workers of Khyber Pakhtunkhwa.*Hepatology Mon* 10(3): 205-214.

Sanger F; Nicklen S, Coulson AR(1977).DNA sequencing with chain-terminating inhibitors".*Proc. Natl. Acad. Sci.* U.S.A. 74 (12): 5463–7.

Schiff E. (1999). The alcoholic patient with hepatitis C virus infection. *American Journal of Medicine*107(6B):95S-9.

Seeff, LB (2000). Hepatitis C. *In Natural History of Hepatitis C*. Edited by T.J. Liang &J.H. Hoofnagle. San Diego, CA: Academic Press. In: Markov *et al* (2009). Phylogeography and molecular epidemiology of hepatitis C virus genotype 2 in Africa.

Shivkumar S, Peeling R, Jafari Y, Joseph L, Pant Pai N (2012). Accuracy of Rapid and Point-of-Care Screening Tests for Hepatitis C: A Systematic Review and Meta-analysis. *Ann Intern Med.* 157(8):558-566.

Shobokshi OA, Serebour FE, Skakni L, Al-Saffy YH, Ahdal MN (1999). Hepatitis C genotypes and subtypes in Saudi Arabia. *Journal of Medical Virology* 58: 44-8.

Simmonds P (2001). The origin and evolution of hepatitis viruses in humans. 2000 Fleming Lecture. *Journal of General Virology* 82: 693-712.

Simmonds P, Bukh J, Combet C, Deleage G, Enomoto N, Feinstone S, Halfon P, Inchauspe G, Kuiken C, Maertens G, Mizokami M, Murphy DG, Okamoto H, Pawlotsky J M, Penin F, Sablon E, Shin I, Stuyver LJ, Thiel HJ, Viazov S, Weiner AJ, Widell A (2005). Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. *Hepatology* 42:962–973.

Simmonds P, Klenerman P, Pybus O.G (2014). Phylogeography and epidemic history of hepatitis C virus genotype 4 in Africa. *Virology* 464-465(2014)233–243.

Sir D, Kuo C, Tian Y, Liu HM, Huang EJ, Jung JU, Machida K, Ou JJ (2012).Replication of Hepatitis C Virus RNA on AutophagosomalMembranes.JBC online.

Smith DB, Bukh J, Kuiken C, Muerhoff AS, Rice CM, Stapleton JT, Simmonds, P (2014). Expanded classification of hepatitisCvirusinto7genotypes and 67 subtypes: updated criteria and genotype assignment web resource. *Hepatology* 59, 318–327.

Smith DB *et al.*(2017).International Committee on Taxonomy of Viruses (ICTV). HCV Classification. A web resource to manage the classification and genotype and subtype assignments of hepatitis C virus. https://talk.ictvonline.org/ictv_wikis/flaviviridae/w/sg_flavi/56/hcv-classification

Smuts HE, Kannemeyer J (1995). Genotyping of hepatitis C virus in South Africa. *Journal of Clinical Microbiology* 33:1679-81.

Sugimoto K *et al.*(2003). Influence of ethnicity in the outcome of hepatitis C virus infection and cellular immune response.*Hepatology* 37:590-9.

Tamura K and Nei M (1993).Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpazees.*Molecular Biology and Evolution* 10: 512-526.

Tamura K., Nei M., and Kumar S. (2004).Prospects for inferring very large phylogenies by using the neighbor-joining method. Proceedings of the National Academy of Sciences (USA) 101:11030-1103

Tarrant MK, Cole PA (2009). The chemical biology of protein phosphorylation *Annu.Rev.Biochem.* 78:797-825.

Taylor DR *et al.*(2000). Hepatitis C virus and interferon resistance. Microbes and Infection 2: 1743-56.

Thimme R, Oldach D, Chang KM, Steiger C, Ray SC, Chisari FV(2001). Determinants of viral clearance and persistence during acute hepatitis C virus infection. *Journal of Experimental Medicine* 194, 1395-1406.

Thomas, DL, Astemborski, J, Rai RM, Anania FA, Schaeffer M, Galai N (2000). The natural history of hepatitis C virus infection: host, viral, and environmental factors. JAMA 284: 450-456

Thursz MR *et al* (1997). Host factors in chronic viral hepatitis. *Seminar on Liver Disease* 17:345-50.

Toivola DN, Ku NO, Resurreccion EZ, Nelson DR, Wright TL, Bishr Omary M (2004). Keratin 8 and 18 Hyperphosphorylation Is a Marker of Progression of Human Liver Disease. *Hepatology* 40:459–466.

Tong X, Pogam SL, Li L, Haines K, Piso K, Baronas V, Yan JM, So S, Klumpp K and Najera I (2014). *In vivo* Emergence of a Novel Mutant L159F/L320F in the NS5B Polymerase Inhibitors Mericitabine and Sofosbuvir. *Hoffman-La Roche Inc, Nutley, New Jersey.*

Tucker ME (2013). Hepatitis C viral suppression reduces liver morbidity, death. *Medscape Medical News.* Accessed: November 18, 2013.

Tucker ME (2013). FDA approves 'game changer' hepatitis C drug sofosbuvir. Medscape Medical News from WebMD. 2013 Dec 6; Accessed: December 21, 2013.

US Food and Drug Administration. FDA approves rapid test for antibodies to hepatitis C virus. *FDA News Release*. June 25, 2010.

US Food and Drug Administration. FDA approves first genotyping test for patients with hepatitis C virus. *FDA News Release*. 2013, Jun 20.

US Food and Drug Administration. FDA approves Sovaldi for chronic hepatitis C. FDA News Release. 2013 Dec 6; Accessed: December 23, 2013.

US Food and Drug Administration. 510(k) Premarket notification: Hepatiq. Accessed: January 11, 2015.

US Food and Drug Administration. FDA approves Viekira Pak to treat hepatitis C. FDA News Release. 2014 Dec 19; Accessed: April 6, 2015.

Vento S, Cainelli F (2002). Does hepatitis C virus cause severe liver disease only in people who drink alcohol? *Lancet Infectious Dis* 2(5):303-9.

Viso ATR (2007). Pathogenesis of hepatitis C-HCV Consensus 2007. *The Brazilian Journal of Infectious Disease* 11 Supplement 1.

World Health Organization (2014). Viral Hepatitis Factsheet

World Health Organisation(2017).Global hepatitis report, 2017, ISBN 978-992-4156545-4156545

World Health Organization (2018). Hepatitis C. Key Facts. https//www.who.int

Worobey M, Gemmel M, Teuwen DE, Haselkorn T, Kunstman K, Bunce M, Muyembe J J, Kabongo RM and other authors (2008). Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960.*Nature*455, 661-664.

Saitou N, Nei M (1987).The Neighbor-Joining Method-a new method for reconstructing phylogenetic trees. *Mol Biol Evol*. 4: 406-425.

Wyles D (2013). Antiviral resistance and the future landscape of hepatitis C virus infection therapy. *Journal of Infectious Disease* 207 (suppl 1):S33-39.

Yeung LT, King SM, Roberts EA (2001). Mother-to-infant transmission of hepatitis C virus. *Hepatology* 34(2):223-9.

Yun Z, Lara C, Johansson B, Lorenzana DR, Sonnerborg A (1996). Discrepancy of hepatitis c virus genotypes as determined by phylogenetic analysis of partial NS5B and core sequences. *Journal of Medical Virology* 49: 155-160. 35.

Zein NN (2000). Clinical significance of hepatitis C virus genotypes. *Clin Microbiol Rev*13, 223-235.

.

APPENDIX I

Pairwise Distance between some study HCV isolates and Reference Sequences

AF009606.1|_1a_H77_Clone_USA		0.546	0.532	0.235	0.079	0.185	0.239	0.606	0.600	0.594	0.555	0.249	0.735	0.594	0.016	0.653	0.606	0.578	0.602	0.606	0.012	0.653	0.664	0.022	0.016	0.023	0.012	0.006	0.350	0.599	0.021	0.680	0.024	0.599	0.016	0.606	0.606	0.146	0.433	0.606	0.602
FJ791092|Central_AFrican_Rep_4	0.471		0.129	0.566	0.531	0.540	0.549	0.681	0.649	0.689	0.615	0.570	0.707	0.697	0.554	0.709	0.677	0.651	0.681	0.677	0.556	0.709	0.720	0.621	0.584	0.578	0.556	0.546	0.731	0.685	0.559	0.820	0.606	0.685	0.554	0.677	0.677	0.677	0.849	0.677	0.680
GU088166.1|Congo.4c	0.487	0.207		0.446	0.499	0.466	0.429	0.557	0.513	0.529	0.549	0.489	0.619	0.568	0.520	0.573	0.538	0.512	0.536	0.538	0.538	0.573	0.579	0.568	0.567	0.507	0.538	0.532	0.672	0.539	0.510	0.635	0.526	0.539	0.520	0.538	0.538	0.588	0.639	0.538	0.534
JQ679054|_Nigeria_1	0.238	0.561	0.459		0.097	0.035	0.015	0.510	0.481	0.495	0.564	0.159	0.616	0.473	0.216	0.514	0.505	0.478	0.502	0.505	0.227	0.514	0.521	0.231	0.250	0.250	0.227	0.233	0.521	0.496	0.219	0.697	0.207	0.496	0.216	0.505	0.505	0.140	0.432	0.505	0.502
JQ679065|Nigeria_1	0.158	0.497	0.455	0.166		0.106	0.124	0.599	0.563	0.588	0.586	0.149	0.669	0.545	0.090	0.643	0.599	0.568	0.595	0.599	0.089	0.643	0.647	0.126	0.110	0.141	0.089	0.082	0.374	0.588	0.113	0.651	0.091	0.588	0.090	0.599	0.599	0.112	0.417	0.599	0.595
JQ679058|Nigeria_1	0.200	0.487	0.456	0.086	0.180		0.029	0.502	0.481	0.505	0.524	0.130	0.546	0.466	0.171	0.543	0.495	0.474	0.498	0.495	0.195	0.543	0.550	0.221	0.217	0.203	0.195	0.188	0.409	0.491	0.184	0.658	0.198	0.491	0.171	0.495	0.495	0.164	0.419	0.495	0.498
JQ679055|Nigeria_1	0.238	0.537	0.439	0.027	0.181	0.068		0.504	0.482	0.489	0.554	0.127	0.600	0.467	0.225	0.515	0.499	0.472	0.496	0.499	0.230	0.515	0.523	0.234	0.253	0.233	0.230	0.241	0.493	0.490	0.191	0.675	0.210	0.490	0.225	0.499	0.499	0.165	0.427	0.499	0.496
DQ333681|South_Africa_5a	0.509	0.582	0.512	0.487	0.530	0.457	0.477		0.021	0.017	0.569	0.526	0.579	0.014	0.628	0.019	0.017	0.020	0.016	0.017	0.606	0.019	0.023	0.622	0.635	0.634	0.606	0.611	0.760	0.018	0.610	0.173	0.621	0.018	0.628	0.017	0.017	0.596	0.102	0.017	0.016
DQ333672|South_Africa_5a	0.520	0.546	0.475	0.455	0.504	0.445	0.465	0.044		0.021	0.578	0.533	0.553	0.019	0.612	0.023	0.021	0.025	0.020	0.021	0.599	0.023	0.026	0.593	0.628	0.622	0.599	0.595	0.765	0.018	0.570	0.158	0.586	0.018	0.612	0.021	0.021	0.550	0.112	0.021	0.020
AB204665|South_Africa_5a	0.496	0.608	0.484	0.465	0.517	0.469	0.455	0.033	0.044		0.588	0.512	0.553	0.018	0.616	0.025	0.013	0.017	0.012	0.013	0.594	0.025	0.028	0.610	0.622	0.621	0.594	0.599	0.758	0.015	0.598	0.135	0.609	0.015	0.616	0.013	0.013	0.581	0.096	0.013	0.012
CS101285|Gt_7_DRC	0.520	0.604	0.550	0.633	0.640	0.571	0.630	0.500	0.512	0.525		0.709	0.614	0.599	0.537	0.598	0.586	0.558	0.583	0.586	0.566	0.598	0.608	0.591	0.584	0.542	0.566	0.550	0.674	0.578	0.581	0.715	0.593	0.578	0.537	0.586	0.586	0.738	0.626	0.586	0.582
GQ866993|_1b	0.244	0.517	0.472	0.207	0.213	0.186	0.185	0.501	0.493	0.478	0.721		0.565	0.532	0.265	0.531	0.518	0.497	0.521	0.518	0.251	0.531	0.540	0.278	0.275	0.262	0.251	0.246	0.407	0.513	0.230	0.631	0.251	0.513	0.265	0.518	0.518	0.067	0.467	0.518	0.521
GQ153856|Guinea_Bissau_2	0.673	0.653	0.571	0.646	0.657	0.564	0.632	0.575	0.538	0.548	0.564	0.617		0.586	0.737	0.581	0.553	0.522	0.550	0.553	0.737	0.581	0.597	0.809	0.758	0.772	0.737	0.740	0.365	0.549	0.749	0.681	0.738	0.549	0.737	0.553	0.553	0.674	0.638	0.553	0.549
AB204674|_South_Africa_5a	0.517	0.597	0.525	0.470	0.499	0.441	0.460	0.027	0.038	0.038	0.515	0.503	0.578		0.616	0.023	0.018	0.022	0.018	0.018	0.595	0.023	0.026	0.583	0.623	0.620	0.595	0.599	0.737	0.020	0.571	0.187	0.582	0.020	0.616	0.018	0.018	0.592	0.078	0.018	0.017
NGR-JS-BD2	0.027	0.479	0.487	0.229	0.164	0.205	0.243	0.525	0.514	0.512	0.489	0.253	0.701	0.533		0.665	0.628	0.600	0.624	0.628	0.022	0.665	0.676	0.026	0.025	0.024	0.022	0.014	0.381	0.621	0.031	0.703	0.030	0.621	0.000	0.628	0.628	0.159	0.452	0.628	0.624
NGR-JS-BD7	0.581	0.612	0.525	0.499	0.594	0.514	0.509	0.038	0.050	0.061	0.527	0.514	0.564	0.056	0.576		0.025	0.028	0.024	0.025	0.648	0.000	0.014	0.670	0.676	0.647	0.648	0.648	0.790	0.023	0.651	0.184	0.662	0.023	0.665	0.025	0.025	0.625	0.144	0.025	0.024
NGR-JS-P02	0.530	0.582	0.507	0.487	0.541	0.447	0.477	0.033	0.044	0.022	0.532	0.491	0.538	0.038	0.546	0.061		0.014	0.009	0.000	0.606	0.025	0.028	0.622	0.634	0.634	0.606	0.611	0.761	0.011	0.610	0.131	0.621	0.011	0.628	0.000	0.000	0.588	0.106	0.000	0.009
NGR-JS-P09a	0.499	0.581	0.488	0.449	0.499	0.439	0.439	0.044	0.055	0.033	0.512	0.481	0.517	0.050	0.514	0.073	0.022		0.012	0.014	0.579	0.028	0.031	0.595	0.607	0.607	0.579	0.583	0.727	0.015	0.583	0.164	0.594	0.015	0.600	0.014	0.014	0.564	0.123	0.014	0.010
NGR-JS-P09b	0.517	0.612	0.512	0.475	0.527	0.455	0.465	0.033	0.044	0.022	0.537	0.499	0.543	0.038	0.533	0.061	0.011	0.016		0.009	0.603	0.024	0.027	0.618	0.631	0.630	0.603	0.607	0.756	0.011	0.606	0.138	0.617	0.011	0.624	0.009	0.009	0.591	0.101	0.009	0.006
NGR-JS-P15a	0.530	0.582	0.507	0.487	0.541	0.447	0.477	0.033	0.044	0.022	0.532	0.491	0.538	0.038	0.546	0.061	0.000	0.022	0.011		0.606	0.025	0.028	0.622	0.634	0.634	0.606	0.611	0.761	0.011	0.610	0.131	0.621	0.011	0.628	0.000	0.000	0.588	0.106	0.000	0.009
NGR-JS-BD6	0.016	0.483	0.479	0.237	0.172	0.214	0.237	0.509	0.520	0.496	0.532	0.251	0.685	0.517	0.044	0.560	0.530	0.499	0.517	0.530		0.648	0.658	0.020	0.009	0.024	0.000	0.014	0.365	0.593	0.019	0.677	0.018	0.593	0.022	0.606	0.606	0.149	0.429	0.606	0.603
NGR-JS-BD9	0.581	0.612	0.525	0.499	0.594	0.514	0.509	0.038	0.050	0.061	0.527	0.514	0.564	0.056	0.576	0.000	0.061	0.073	0.061	0.061	0.560		0.014	0.670	0.676	0.647	0.648	0.648	0.790	0.023	0.651	0.184	0.662	0.023	0.665	0.025	0.025	0.625	0.144	0.025	0.024
NGR-JS-BD10	0.595	0.626	0.517	0.512	0.586	0.527	0.522	0.049	0.061	0.073	0.540	0.527	0.581	0.067	0.591	0.022	0.072	0.085	0.073	0.072	0.573	0.022		0.680	0.651	0.657	0.658	0.658	0.789	0.025	0.661	0.192	0.672	0.025	0.676	0.028	0.028	0.633	0.121	0.028	0.026
NGR-JS-BD12	0.044	0.549	0.504	0.238	0.194	0.231	0.238	0.504	0.509	0.491	0.541	0.269	0.756	0.486	0.056	0.576	0.525	0.494	0.512	0.525	0.038	0.576	0.591		0.024	0.022	0.020	0.023	0.387	0.615	0.022	0.699	0.016	0.615	0.026	0.622	0.622	0.165	0.398	0.622	0.618
NGR-JS-BD13	0.027	0.511	0.507	0.255	0.187	0.231	0.255	0.538	0.549	0.525	0.548	0.269	0.705	0.546	0.055	0.591	0.560	0.527	0.546	0.560	0.011	0.591	0.570	0.050		0.028	0.009	0.017	0.380	0.621	0.023	0.707	0.022	0.621	0.025	0.634	0.634	0.175	0.427	0.634	0.631
NGR-JS-BD20	0.050	0.498	0.445	0.255	0.222	0.229	0.253	0.525	0.535	0.512	0.499	0.261	0.709	0.533	0.056	0.538	0.546	0.514	0.533	0.546	0.056	0.538	0.552	0.050	0.067		0.024	0.021	0.339	0.627	0.026	0.699	0.024	0.627	0.024	0.634	0.634	0.181	0.429	0.634	0.630
NGR-JS-BD23	0.016	0.483	0.479	0.237	0.172	0.214	0.237	0.509	0.520	0.496	0.532	0.251	0.685	0.517	0.044	0.560	0.530	0.499	0.517	0.530	0.000	0.560	0.573	0.038	0.011	0.056		0.014	0.365	0.593	0.019	0.677	0.018	0.593	0.022	0.606	0.606	0.149	0.429	0.606	0.603
NGR-JS-BD28	0.005	0.471	0.487	0.231	0.165	0.207	0.246	0.520	0.509	0.507	0.509	0.236	0.685	0.527	0.022	0.570	0.540	0.509	0.527	0.540	0.022	0.570	0.584	0.050	0.033	0.044	0.022		0.353	0.604	0.023	0.686	0.025	0.604	0.014	0.611	0.611	0.141	0.437	0.611	0.607
NGR-JS-BD30	0.331	0.633	0.615	0.517	0.399	0.422	0.493	0.714	0.739	0.726	0.662	0.429	0.409	0.717	0.369	0.768	0.726	0.697	0.735	0.726	0.357	0.768	0.756	0.376	0.369	0.336	0.357	0.339		0.768	0.360	0.845	0.368	0.768	0.381	0.761	0.761	0.496	0.751	0.761	0.755
NGR-JS-BD1	0.517	0.597	0.499	0.465	0.517	0.445	0.455	0.038	0.038	0.027	0.520	0.488	0.527	0.044	0.533	0.055	0.016	0.027	0.016	0.016	0.496	0.055	0.067	0.512	0.525	0.533	0.496	0.527	0.743		0.597	0.125	0.608	0.000	0.621	0.011	0.011	0.599	0.083	0.011	0.011
NGR-JS-BD3	0.044	0.476	0.465	0.229	0.193	0.207	0.200	0.520	0.484	0.507	0.566	0.221	0.701	0.502	0.073	0.570	0.540	0.509	0.527	0.540	0.038	0.570	0.584	0.044	0.050	0.061	0.038	0.050	0.349	0.507		0.678	0.020	0.597	0.031	0.610	0.610	0.138	0.382	0.610	0.606
NGR-JS-BD4	0.606	0.789	0.650	0.627	0.598	0.590	0.602	0.235	0.220	0.211	0.727	0.565	0.749	0.244	0.624	0.244	0.205	0.229	0.211	0.205	0.595	0.244	0.252	0.613	0.627	0.617	0.595	0.617	0.912	0.197	0.595		0.691	0.125	0.703	0.131	0.131	0.691	0.257	0.131	0.138
NGR-JS-P03	0.050	0.535	0.467	0.221	0.178	0.214	0.221	0.522	0.507	0.509	0.559	0.251	0.685	0.504	0.067	0.573	0.543	0.512	0.530	0.543	0.033	0.573	0.587	0.027	0.044	0.056	0.033	0.055	0.366	0.509	0.038	0.609		0.608	0.030	0.621	0.621	0.125	0.391	0.621	0.617
NGR-JS-P10	0.517	0.597	0.499	0.465	0.517	0.445	0.455	0.038	0.038	0.027	0.520	0.488	0.527	0.044	0.533	0.055	0.016	0.027	0.016	0.016	0.496	0.055	0.067	0.512	0.525	0.533	0.496	0.527	0.743	0.000	0.507	0.197	0.509		0.621	0.011	0.011	0.599	0.083	0.011	0.011
NGR-JS-P13	0.027	0.479	0.487	0.229	0.164	0.205	0.243	0.525	0.514	0.512	0.489	0.253	0.701	0.533	0.000	0.576	0.546	0.514	0.533	0.546	0.044	0.576	0.591	0.056	0.055	0.056	0.044	0.022	0.369	0.533	0.073	0.624	0.067	0.533		0.628	0.628	0.159	0.452	0.628	0.624
NGR-JS-P15	0.530	0.582	0.507	0.487	0.541	0.447	0.477	0.033	0.044	0.022	0.532	0.491	0.538	0.038	0.546	0.061	0.000	0.022	0.011	0.000	0.530	0.061	0.072	0.525	0.560	0.546	0.530	0.540	0.726	0.016	0.540	0.205	0.543	0.016	0.546		0.000	0.588	0.106	0.000	0.009
NGR-JS-P16	0.530	0.582	0.507	0.487	0.541	0.447	0.477	0.033	0.044	0.022	0.532	0.491	0.538	0.038	0.546	0.061	0.000	0.022	0.011	0.000	0.530	0.061	0.072	0.525	0.560	0.546	0.530	0.540	0.726	0.016	0.540	0.205	0.543	0.016	0.546	0.000		0.588	0.106	0.000	0.009
NGR-JS-P17	0.206	0.594	0.551	0.177	0.196	0.199	0.185	0.554	0.506	0.530	0.766	0.113	0.689	0.543	0.212	0.574	0.543	0.522	0.551	0.543	0.213	0.574	0.589	0.221	0.229	0.230	0.213	0.199	0.466	0.568	0.192	0.616	0.197	0.568	0.212	0.543	0.543		0.447	0.588	0.592
NGR-JS-P18	0.395	0.706	0.549	0.409	0.391	0.409	0.400	0.149	0.149	0.142	0.540	0.435	0.595	0.135	0.407	0.171	0.156	0.171	0.157	0.156	0.377	0.171	0.162	0.347	0.384	0.395	0.377	0.404	0.693	0.136	0.345	0.257	0.346	0.136	0.407	0.156	0.156	0.439		0.106	0.100
NGR-JS-P06	0.530	0.582	0.507	0.487	0.541	0.447	0.477	0.033	0.044	0.022	0.532	0.491	0.538	0.038	0.546	0.061	0.000	0.022	0.011	0.000	0.530	0.061	0.072	0.525	0.560	0.546	0.530	0.540	0.726	0.016	0.540	0.205	0.543	0.016	0.546	0.000	0.000	0.543	0.156		0.009
NGR-JS-P09	0.520	0.605	0.507	0.477	0.530	0.457	0.467	0.033	0.044	0.022	0.532	0.501	0.538	0.038	0.535	0.061	0.011	0.011	0.005	0.011	0.520	0.061	0.072	0.514	0.549	0.535	0.520	0.530	0.726	0.016	0.530	0.212	0.533	0.016	0.535	0.011	0.011	0.554	0.156	0.011	

APPENDIX2

Arch. Bas. App. Med. 6 (2018) 79 – 85
www.archivesbamui.com

ARCHIVES OF BASIC & APPLIED MEDICINE

Research Article

Genetic Diversity of Hepatitis C Virus Among Blood Donors and Patients with Clinical Hepatitis in Ibadan, Nigeria

Shenge J.A., Odaibo G.N., *Olaleye D.O.
Department of Virology, College of Medicine, University of Ibadan, Nigeria.

Received: June, 2017; Revised version Accepted: November, 2017

Abstract
Hepatitis C virus (HCV) infection is responsible for liver diseases and hepatocellular carcinoma in chronically-infected patients. Owing to high sequence variability in HCV genome, numerous subtypes have emerged. This study determined HCV strains among patients with clinical hepatitis and blood donors in Ibadan. Blood samples were collected from consented 176 subjects who tested positive to HCV IgM antibodies, including 99 patients with clinical hepatitis and 77 apparently healthy blood donors. Viral RNA was extracted from blood samples, while presence of HCV was tested by amplifying the NS5B gene using polymerase chain reaction (PCR). The amplified NS5B gene was sequenced and sequences were aligned on MEGA 7.0. Phylogenetic tree was constructed with Neighbor-Joining method. Data were analyzed using descriptive statistics at P<0 .05. The NS5B gene was amplified in 38 samples, of which 29 were successfully sequenced. Phylogenetic analysis revealed three of seven known genotypes of HCV including genotypes / subtypes 1a (34.5%), 1b (17.2%), 2b (13.8%), 2c (3.6%) and 5a (31.3%). Subtypes 1b and 2b were found among patients with clinical hepatitis, while the single 2c was found among donors. Although subtype 1a was detected among both populations, its rate was higher among blood donors (P = 0 .003). Subtype 5a was found among the two groups (P= 1. 00). HCV subtypes 1a and 5a are the predominant strains in Ibadan. The diversity of HCV observed has implications for treatment of patients and design of a broadly protective vaccine against the virus.

Key Words: *Hepatitis C Virus, Patients, Blood donors, Genotypes, Nigeria*

INTRODUCTION

Over 185 million people worldwide are infected with hepatitis C virus (HCV) (Mora *et al*., 2016). An estimated 3-4 million people become infected every year globally (Ashfaq *et al*., 2011). This accounts for about 3% of the world's population that are infected with HCV, with most of these cases occurring in Africa (Kapoor *et al*., 2011). HCV infection is responsible for most deaths emanating from liver failure and liver cancer each year (Simmonds, 2004).

Studies have shown that almost 75% of HCV related deaths occur among adults between the ages of 45 and 64 as a result of long-term infection with HCV that leads to chronic liver disease including liver cirrhosis and hepatocellular carcinoma (Mukherjee, 2012; Dhawan, 2016; Ly *et al*., 2012; Davis *et al*., 1989). Chronic infection with HCV has been reported as the main cause of liver disease, and this might be the reason for carrying out most of the Orthotopic Liver Transplantation (OLT) procedures in the USA (Dhawan, 2016; Davis *et al*., 1989). In sub-Saharan Africa, HCV infection is a major health challenge and has been implicated in liver disease and its complications in chronically-infected-patients (Mora *et al*., 2016; WHO, 2012; Rao *et al*., 2015).

Hepatitis C virus is a member of the Flaviviridae family and the only member of the genus Hepacivirus (Simmonds, 2004; WHO, 2012). The virus is a small enveloped, spherical virus with a positive sense, single-stranded RNA genome (Simmonds, 2004). The HCV genome consists of a single, open reading frame (ORF) that is 9600 nucleotide bases long and 2 untranslated, but highly conserved regions namely 5'-UTR and 3'-UTR located at both ends of the genome (Kato, 2000). The genome encodes a single polyprotein starting with the core proteins (structural) and ending with the NS5B protein, a non-structural protein that codes for the RNA polymerase (Lindenbach *et al*., 2005). The NS5B gene codes for RNA-dependent RNA polymerase (RdRp), an enzyme that is essential for viral maturation and plays an important catalytic role during replication of HCV (Penin *et al*., 2004). According to Ashfaq *et al* (2011), this gene represents an ideal target for the development of antiviral drugs.

Genotypes and subtypes of HCV can be differentiated based on the sequences of the NS5B gene, a relatively variable region of HCV genome (Gededzha *et al*., 2012). This variability results in substitutions as the virus mutates. These nucleotide substitutions during HCV replication has resulted in the emergence of seven major HCV genotypes (1–7), each further divided into subtypes based on their genetic diversity (Simmonds *et al*., 2005; Ohno, 2007). To date 67 well defined and 20 unconfirmed subtypes have been identified (Messina *et al*., 2014).

HCV genotypes and subtypes are distributed differently throughout the world (Ramia and Eid-Fares 2006). Divergent strains of genotypes 1 and 2 have been shown to be endemic in West African countries including Burkina Faso, Ghana, Guinea Bissau, Benin Republic and Nigeria (Forb *et al*., 2012; Markov *et al*., 2009), genotypes 3 is found in South Asia, genotype 4, 5 and 6 are more predominant in central Africa and Middle East, South-east Asia, Northern region of South Africa and Belgium respectively (Markov *et al*., 2009). Genotype 7 has been reported only in central African

*Author for Correspondence: +2348126286260 E-mail: davidoolaleye@gmail.com

APPENDIX 3

PLOS ONE

RESEARCH ARTICLE

Phylogenetic analysis of hepatitis C virus among HIV/ HCV co-infected patients in Nigeria

Juliet A. Shenge*, Georgina N. Odaibo, David O. Olaleye

Department of Virology, College of Medicine, University of Ibadan, Ibadan, Nigeria

* jadamma@yahoo.com

OPEN ACCESS

Citation: Shenge JA, Odaibo GN, Olaleye DO (2019) Phylogenetic analysis of hepatitis C virus among HIV/ HCV co-infected patients in Nigeria. PLoS ONE 14(2): e0210724. https://doi.org/10.1371/journal.pone.0210724

Editor: Jason Blackard, University of Cincinnati College of Medicine, UNITED STATES

Received: July 4, 2018

Accepted: December 31, 2018

Published: February 6, 2019

Data Availability Statement: All relevant data are within the manuscript and its supporting files. NS5B sequences available at Dol: 10.6084/m9.figshare.7471454.

Funding: This study was supported by Medical Education Partnership Initiative Nigeria (MEPIN) through National Institute of Health (NIH) USA grant funded by Fogarty International Centre, the office of AIDS Research and National Human Genome Research Institute of NIH, the Health Resources and Services Administration (HRSA) and the Office of the U.S. Global AIDS Coordinator

Abstract

Hepatitis C virus (HCV) infection has been associated with liver disease including liver cirrhosis and hepatocellular carcinoma (HCC) in chronically-infected persons. However, in HIV/HCV co-infected patients, increased rate of progression to cirrhosis and HCC has been reported. Limited information exists regarding genetic variants of HCV circulating among co-infected patients, which could be important in the design of broadly protective vaccine and management of the disease. Here, we determined the genotypes of HCV isolates circulating among HIV/HCV co-infected patients in Ibadan, southwestern Nigeria. One hundred and twenty-five HIV/HCV IgM positive samples obtained from HIV laboratory, University of Ibadan were used for this study. HCV NS5B gene was amplified using polymerase chain reaction (PCR). The amplified NS5B gene was sequenced using gene specific primers. Twenty isolates were amplified, out of which 13 were successfully sequenced. Phylogenetic analysis of the 13 sequenced isolates showed three HCV subtypes 1a, 3a and 5a belonging to genotypes 1, 3 and 5 respectively. Ten isolates (77%) belong to subtype 5a, followed by 2 isolates (15%) subtype 1a and 1 isolate (8%) was subtype 3a. The predominant HCV genotype was 5, followed by genotype 1 (subtype 1a). The findings, as well as the observed mutations in NS5B gene, indicate the need for screening and monitoring of HIV/HCV co-infected patients. Further study to determine the phylogeny of isolates circulating in other parts of Nigeria will be carried out.

Introduction

Hepatitis C virus (HCV) still affects more than 185 million people worldwide despite availability of highly effective antiviral therapy such as direct acting antiviral agents (DAA) [1]. According to [2], an estimated 3–4 million people become infected every year, representing more than three percent of the world's population that are chronically infected, with most of these cases occurring in Africa [2, 3]. It is also estimated that about 350, 000 people die from liver failure and liver cancer caused by hepatitis C disease each year. Furthermore, about 2.3 million people infected with HIV are actually co-infected with hepatitis C virus globally [4].

It has been reported that about one third of HIV-infected patients are also infected with HCV with an undesirable impact on HCV pathogenesis [5]. HIV destroys the immune system by depleting CD4-bearing T cells while HCV causes necrosis of the hepatocytes. Liver disease