Journal of Bio-Molecular Sciences (JBMS) (2015) 3(2): 66-73.

Journal of Bio-Molecular Sciences
(JBMS)
ISSN: 2311-4630
www.directsciences.com/jbms

Determination of Hepatitis C Virus Genotypes by Direct Sequencing Analysis of 5' UTR of HCV Genome

Nasar Khan[1], Atta Ullah[1], Shaista Zafar[2], Iqbal Ahmad[1], Kashif Rahim[1], Farman Ullah[3], Shamim Saleha[3]*

[1]Department of Microbiology, Kohat University of Science and Technology (KUST), Kohat 26000, Pakistan.
[2] Pakistan Institute of Medical Sciences, Islamabad 44000, Pakistan
[3]Department of Biotechnology and Genetic Engineering, Kohat University of Science and Technology (KUST) Kohat 26000, Pakistan

Received: 15 May.2015; Accepted: 24 May 2015; Published 31 June 2015

Abstract: High degree of sequence variations are present throughout in the coding regions of Hepatitis C virus (HCV) genome. However, a high degree of sequence conservation within the 5'untranslated region (UTR) has made this region a target of choice for most of detection assays based on nucleic acid amplification. The current study was designed to determine the HCV genotypes in samples of HCV chronically infected patients in Pakistan. Specific primers targeting 5'UTR region were designed, which were then used for both amplification and sequencing of all isolates. All HCV isolates were sequenced and genotyped based upon phylogenetically informative regions within the 5'UTR of HCV genome. Out of 15 samples, 9 (60 %) samples 5'UTR of HCV genome was successfully sequenced. However, only 6 (66.7 %) samples were assigned genotypes based upon sequence comparison with reference sequences of database. Results of this study revealed that five isolates that were assigned genotype 3a, showed 97 to 99 % sequence conservation to isolates of genotype 3a previously reported from Pakistan as well as from rest of world, grouped together and coincided with their positions in the phylogenetic tree. Moreover, one isolate that was assigned genotype 4a, showed 97 % sequence conservation to isolates of genotype 4a previously reported from rest of world, coincided with its position in the phylogenetic tree. Our findings suggest that direct sequencing analysis of the 5'UTR of HCV genome is a sensitive and efficient approach of HCV genotyping; it may be adopted as a routine diagnostic procedure for HCV genotyping in clinical settings.

Key words: HCV genotypes, 5' Untranslated region, Sequencing, Phylogenetic analysis, Pakistan

Introduction

Hepatitis C virus (HCV) infection is a serious health problem that is encountered in daily clinical practices. HCV has affected 170-200 million individuals throughout the

*Correspondence to: Shamim Saleha, Biotechnology and Genetic Engineering, Kohat University of Science and Technology (KUST) Kohat 26000, Pakistan E-mail: shamimsaleha@yahoo.com

world including 17 million individuals from Pakistan (Idress et al., 2009). Hepatitis C infection in early stages is asymptomatic and later on it may lead to advance liver diseases such as liver fibrosis, cirrhosis and hepatocellular carcinoma (HCC), (Berenguer et al., 2001). Of those individuals infected with HCV, 55 to 85% subsequently develop a chronic HCV infection. Moreover, 10 to 20% of patients with chronic HCV infection develop cirrhosis or hepatocellular carcinoma and thousands of people die every year globally due to HCV infection (Ashfaq et al., 2011).

HCV is linear, single stranded, positive sense RNA virus and classified as a separate genus Hepacivirus of the family Flaviviridae (Ndjomou et al., 2003). The size of HCV genome is approximately 9.6 kb. The 5' untranslated region (UTR) consists of about 340 nucleotides that forms stem loop structure and contains internal ribosomal entry site (IRES). The 5'UTR is the highly conserved region in HCV genome that is important for initiation of translation process and ribosomal binding (Beales et al., 2001). Downstream to the 5'UTR is a single large open reading frame (ORF) consists of about 9,000 nucleotides that encodes to a nearly 3,010 amino acids polyprotein precursor. Then it is cleaved by combination of viral and host proteases into structural and non-structural proteins essential for viral replication and for viron formation in post-translational process (Grakoui et al., 1993). From the 5' region of ORF one capsid protein, a putative protein P7 of unknown function and two envelop protein E1 and E2 are encoded. Moreover, six nonstructural (NS2, NS3, NS4A, NS4B, NS5A and NS5B) proteins are encoded from 3' region of ORF. At the end of ORF is a polypurine or polypyrimidine tail of about 50 nucleotides that make 3' UTR and is important in secondary structure formation (Tanaka et al., 1995). However, replication errors are greatly produced during replication by RNA-dependent RNA polymerase, consequently producing nucleotide sequence variability (Tanaka et al., 1992).

From all over the world complete or partial nucleotide sequences of HCV isolates have been reported. Comparison of these sequences revealed marked genetic heterogeneity existed within the HCV genome (Nicot et al., 2005; Rehman et al., 2011; Filippo et al., 2012; Wu et al., 2012; Aziz et al., 2013). HCV taxonomy is based upon phylogenetic analyses of nucleotide sequence data (Simmonds, 2013). HCV genotypes are substantially divergent in sequence from each other and fall into 7 phylogenetic clades, designated as genotypes (Tamura et al., 2011). Within these, a variable number of sub-groupings are apparent. HCV genotype 1 is most prevalent genotype in the United States, Europe and Japan (Manos et al., 2012; Qattan and Emery, 2012; Wu et al., 2012). HCV genotype 2 is dominant in Korea and Taiwan. HCV genotype 3 is predominant in Pakistan, India and Thailand while HCV genotype 4 is more prevalent in Saudi Arabia, Egypt, Lebanon, Iraq and Syria. HCV genotype 5 is predominant in South Africa while genotype 6 in Vietnam (Sievert et al., 2011).

Although the classification of new HCV isolates should preferably be based on complete genome sequences, provisional genotype assignments can be based on phylogenetic analysis of nucleotide sequences of noncoding and coding regions. It has been proved that phylogenetic analysis of HCV sequences has led scientists to better understand the origin and evolution of the virus and therefore estimate its impact on human health. In HCV genome the 5'UTR is correspondingly the slowest evolving genomic region (Cantaloube, 2006). From sequence analysis of the 5'UTR

of HCV, a wealth of phylogenetic information may be derived such as identification of genotype present in our population. The current study was designed to determine the HCV genotypes of isolates of HCV from chronically infected patients in Pakistan.

Materials and Methods

Patients

We included in this study 15 HCV chronically infected patient samples that could not be genotyped by assay we previously utilized in our study (Khan et al., 2014). To conduct this study approval was obtained from Ethical Committee of Department of Microbiology, Kohat University of Science Technology (KUST). Written informed consent was obtained from patients for participation in this study.

RNA Extraction and RT-PCR

HCV RNA was extracted from serum samples of patients by using QIAamp Viral RNA Kit (QIAGEN, Valencia, CA, USA) according to manufacturer's instructions. The amplification was carried out by using one-step RT- PCR kit (QIAGEN, Valencia, CA, USA). Primers (5′-ACCCGCCCCTAATAGGGGCG-3′) and (5′- CTTGTGGGCGACGGCTGGTG -3′) were designed by using online available software JustBio at http://www.justbio.com/ to amplify a 345 bp DNA fragment of the 5' UTR region HCV. RT-PCR was carried out in a total reaction volume of 50µl in which 10µl of extracted RNA and 0.6 µM of each primer were added. PCR cycling conditions consisted of 1 cycle at 50ºC for 30 minutes; followed by 1 cycle of initial denaturation at 95ºC for 15 min and 35 cycles of denaturation at 95ºC for 60 seconds; annealing at 62ºC for 40 seconds; extension at 72º C for 40 seconds and a final extension at 72ºC for 10 minutes. PCR products were sequenced in a specialized commercial laboratory at Institute of Biomedical and Genetic Engineering (IBGE), Islamabad Pakistan.

Phylogenetic Analysis

HCV genotypes were determined by sequence analysis of the 5' UTR of HCV genome isolated from 15 patients. Nucleotide sequences obtained from direct sequencing were imported to Molecular Evolutionary Genetics Analysis (MEGA) computer program version 6.06, which was used to align the sequences, according to reference sequences. Homology analysis of the HCV sample sequences with known nucleotide sequences present in National Center for Biotechnology Information (NCBI) was performed through standard nucleotide–nucleotide Blast (Basic Local Alignment Search Tool) software available at website http:// www.ncbi.nlm.nih.gov/BLAST.

Phylogenetic trees were inferred using the Neighbor-Joining (NJ) model and robustness of the tree branches was tested using bootstrap analysis (500 replicates) with MEGA 6.06.

Results

Sequence Analysis

Out of 15 samples, 5'UTR of HCV genome from only 9 (60 %) samples were successfully sequenced. However, among 9 only 6 (66.7 %) samples were assigned genotypes of HCV based upon sequence comparison with reference sequences from NCBI database, remaining 3(33.3%) failed to generate usable sequence information. Sequence alignment of 5'UTR of studied isolates (PK-KUST 01-06) with reference sequences of database highlighted some well-conserved regions as well as some nucleotide substitutions within 5'UTR of HCV genome (Figure 1). The data also revealed that 5' UTR was found to be 325 to 340 nucleotides in length. The sequences of all studied isolates showed 97 to 99 %

sequence conservation to isolates of genotypes existing in worlds as well as to previously reported isolates from Pakistan.

Phylogenetic analysis

Phylogenetic analysis of the HCV 5' UTR region was performed for studied isolates (PK-KUST 01-06) with other sequences reported from the rest of the world (Figure 1). Results of this study revealed that five isolates (PK-KUST 01-04, 06) that were assigned genotypes 3a based on the phylogenetic analyses, grouped together and coincided with their positions in the phylogenetic tree, showing high resemblance among themselves as well as to isolates previously reported from Pakistan. One isolate (PK-KUST 05) that was assigned genotype 4a based on the phylogenetic analyses, coincided with its position in the phylogenetic tree showing resemblance with genotypes 4a from rest of world (Figure 2).

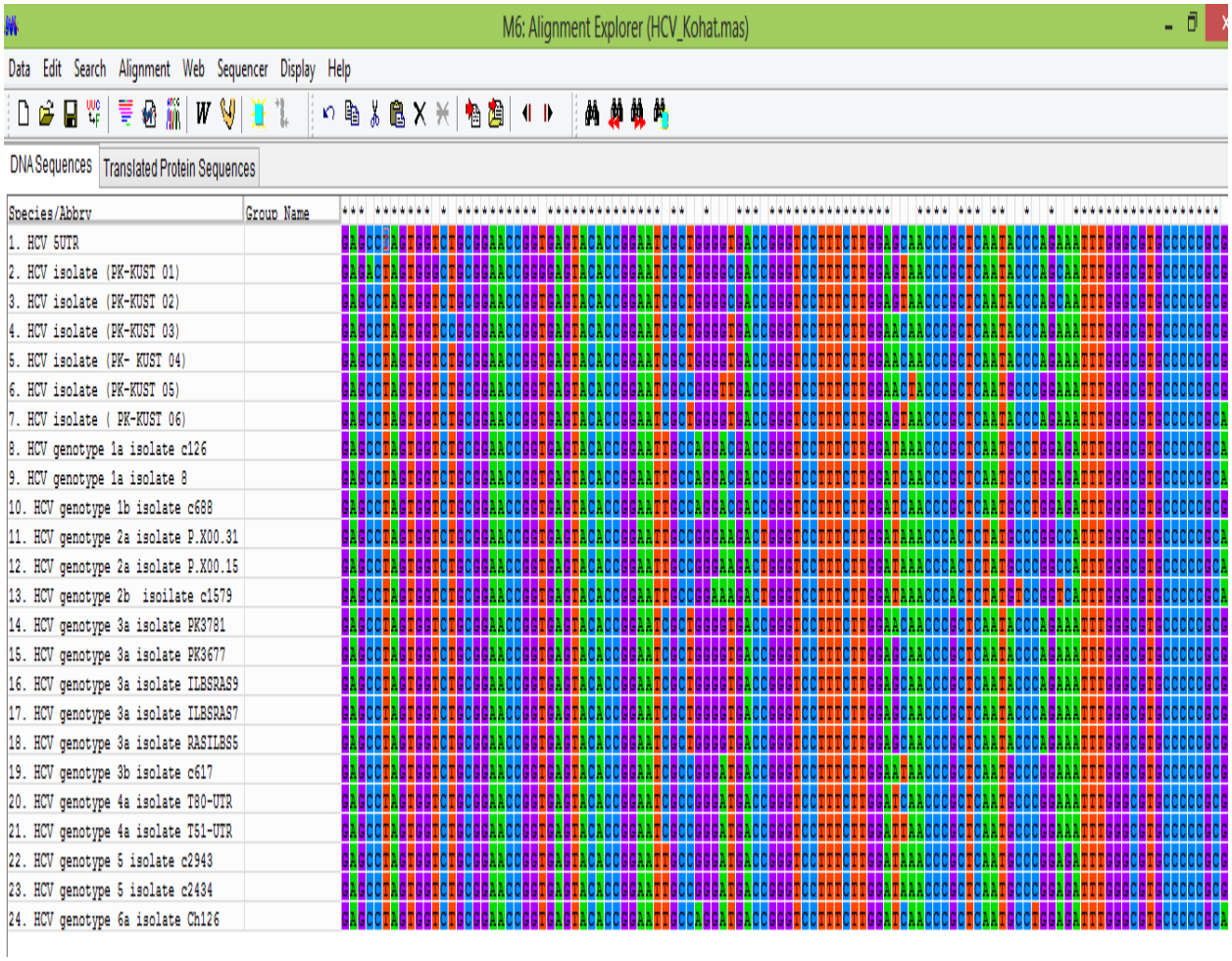

Fig. 1. Multiple sequence alignment of HCV 5'UTR sequences of studied isolates and reference sequences of HCV genotypes retrieved from NCBI.

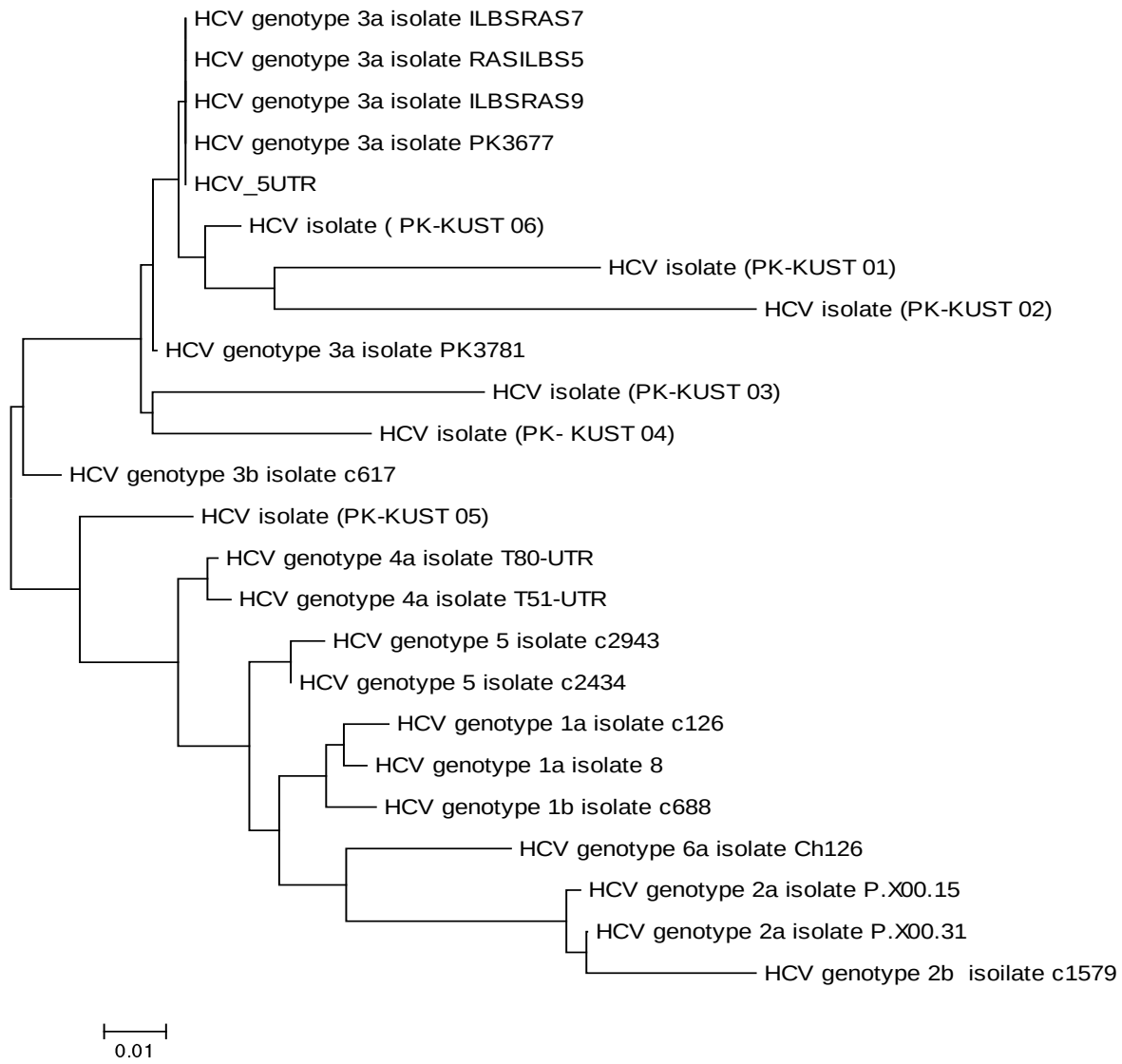

Fig. 2. Phylogenetic neighbor-joining tree of studied isolates and reference isolates of HCV genotypes.

Discussion

Determination of HCV genotypes by direct sequencing of HCV 5' UTR of isolates is an efficient approach as it does not require an further sample processing steps and utilizes amplified products obtained from one-step, non-nested PCR reaction. Consequently it eliminates the delays and cost involved in carrying out further amplification reactions. Additionally, the direct sequencing of PCR products provides more detailed sequence information than other genotyping assays. This additional information could prove to be quite useful in the detection of novel HCV genotypes (Germer et al., 1999; Shier et al., 2014).

Sequencing data of HCV genome indicates that the 5' UTR of 324 to 341 nucleotides is the most conserved region among HCV strains (Beales et al., 2001). This high degree of conservation within the 5'UTR has made it ideal region for detection of sequence based genotypes of HCV. A number of studies utilized this region for

determination of HCV genotypes (Germer et al., 1999; Ahmadi Pour et al., 2006; Anjum et al., 2013; Shier et al., 2014).

In this study sequence analysis of 5'UTR of studied isolates with nucleotide reference sequences retrieved from NCBI Databank highlighted some well-conserved regions. Five studied isolates showed 97 to 99 % sequence conservation with references from genotypes 3a in Pakistan and rest of the world. Some of these studied isolates were 100 % identical to each other and coincided with their positions in the tree (Figure 2), showing high resemblance among themselves. However, one isolate coincided with position of genotypes 4a in the tree reported from world, showed resemblance with them. The sequence conservation in 5' UTR of HCV genome is significant in determination of genotypes by sequence analysis from those samples that may not be typed with more commonly used assays. Sequence analysis of 5'UTR of these studied isolates with nucleotide reference sequences from genotypes of HCV showed some nucleotide substitutions within 5'UTR. The sequence variations were previously reported by Yasmeen et al. (2009) in the 5'UTR of HCV genotype 3a Pakistani isolates in genotypes from rest of world. Similar sequence variability has been observed by Filippo et al. (2012) in the 5'UTR of HCV from multi-transfused patients in Colombia.

Conclusion: The 5' UTR of HCV genome is unique and conserved region in studied Pakistani isolates. Based on our findings the direct sequence analysis of the 5'UTR of HCV genome is a sensitive and efficient means of HCV genotyping in a clinical setting and it may be adopted for HCV genotyping in routine diagnostic procedure.

Competing interests

The author(s) declare that they have no competing interests.

Acknowledgements: This study was financially supported by department of Microbiology, Kohat University of Science and Technology. The authors acknowledge all those patients who participated in the study.

References

Ahmadi Pour, M. H., Keivani, H., Sabahi, F. and Alavian, S. M.2006. Determination of HCV Genotypes, in Iran by PCR-RFLP. Iranian J. Publ. Health, 35(4):54-61.

Anjum, S., Ali, S., Ahmad T., Afzal, M. S., Waheed, Y., Shafi, T., Ashraf, M. and Andleeb, S. 2013. Sequence and structural analysis of 3' untranslated region of Hepatitis C Virus, genotype 3a, from Pakistani isolates. Hepat. Mon. 13(5):e8390.

Ashfaq, U.A., Qasim, M., Yousaf, M.Z., Awan, M.T., and Jahan, S. 2011. Inhibition of HCV 3a genotype entry through host CD81 and HCV E2 antibodies. J. Transl. Med. 9:194.

Aziz, R.D., Kazım, S., İlkay, B., and Nursel, D. 2013. The different distribution of HCV genotypes in eastern black sea region of Turkey. J. Microb. Biochem. Technol. 5: 4.

Beales, L. P., Rowlands, D. J., and Holzenburg, A. 2001. The internal ribosome entry site (IRES) of hepatitis C virus visualized by electron microscopy. RNA, 7: 661-670.

Berenguer, M., Lopez-Labrador, F. X. andWright, T. L. 2001. Hepatitis C and liver transplantation. J. Hepatol. 35: 666–678.

Cantaloube, J. F., Laperche, S., Gallian, P., Bouchardeau, F., Lamballerie, X., and Micco, P. 2006. Analysis of the 5' noncoding region versus the NS5b region in genotyping hepatitis C virus isolates from blood donors in France. J. Clin. Microbiol. 44: 2051-2056.

Filippo, D. D., Fabian, C. M., Beltran, M., Maria, P. A., Sergio, J., Juan, C. R., Gonzalo, C. and Maria, C. N. 2012. Molecular characterization of Hepatitis C Virus in multi-transfused Colombian patients. Virol. J. 9: 242.

Grakoui, A., Wychowski, C., Lin, C., Feinstone, S. M. and Rice, C. M. 1993. Expression and identification of hepatitis C virus polyprotein cleavage products. J. Virol. 67 (3): 1385-1395.

Germer, J. J., Rys, P. N., Thorvilson, J. N., and Persing, D. H.1999. Determination of Hepatitis C Virus Genotype by Direct Sequence Analysis of Products Generated with the Amplicor HCV Test. J. Clin. Microbiol. 37(8): 2625–2630.

Idrees, M., Rafique, S., Rehman, I., Akbar, H., Yousaf, M. Z., Butt, S., Awan, Z., Manzoor, S., Akram, M., Aftab, M., Khubaib, B. and Riazuddin, S. 2009. Hepatitis C virus genotype 3a infection and hepatocellular carcinoma: Pakistan experience. World J. Gastroenterol. 15(40): 5080–5085.

Khan, N., Akmal, M., Hayat M., Umar, M., Ullah, A., Ahmad, A., Rahim, K., Ali, S., Bahadar, S., and Saleha, S. 2014. Geographic Distribution of Hepatitis C Virus Genotypes in Pakistan. Hepat. Mon. 14(10): e20299.

Manos, M. M, Shvachko, V. A., Murphy, R.C., Arduino, J. M. and Shire, N. J. 2012. Distribution of hepatitis C virus genotypes in a diverse US integrated health care population. J. Med. Virol. 84 (11): 1744-1750.

Ndjomou, J., Pybus, O. G. and Matz, B. 2003. Phylogenetic analysis of Hepatitis C Virus isolates indicates a unique pattern of endemic infection in Cameroon. J. Gen. Virol. 84(9): 2333–2341.

Nicot, F., Legrand-Abravanel, F., Sandres-Saune, K., Boulestin, A., Dubois, M., Alric, L., Vinel, J. P., Pasquier, C. and Izopet, J. 2005. Heterogeneity of hepatitis C virus genotype 4 strains circulating in south-western France. J. Gen. Virol. 86: 107–114.

Qattan, I. and Emery, V. 2012. HCV genotyping in chronic hepatitis C patients. E. S. J. 8(27):1-2.

Rehman, L., Ihasn, U., Ali, J., Khan, I. A., Iqbal, A., Khan, S., Khan, S.H., Zaman, K., Khan, N., Swati, Z. A., and Jahangiri, A. T. 2011. Active Hepatitis C infection and HCV genotypes prevalent among the IDUs of Khyber Pakhtunkhwa. Virol. J. 8: 327.

Shier, M. K., El-Wetidy, M. S., Ali, H. H, Al-Qattan, M. M. 2014. Characterization of Hepatitis C Virus genotypes by direct sequencing of HCV 5'UTR region of isolates from Saudi Arabia. PLoS ONE, 9(8): e103160.

Sievert, W., Altraif, I., Razavi, H. A., Abdo, A., Ahmed, E. A., Alomair, A., Amarapurkar, D., Chen, C. H.,Dou, X., and El Khayat, H. 2011. A systematic review of hepatitis C virus epidemiology in Asia, Australia and Egypt. Liver Int. 31(2): 61–80.

Simmonds, P. 2013. The origin of hepatitis C virus. Microbiol. Immunol. 369: 1-15.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. 2011. MEGA5 molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28(10): 2731-2739.

Tanaka, T., Kato, N., Cho, M. and Shimotohno, K. 1995. A novel sequence found at the 3` terminus of hepatitis C virus genome. Biochem. Biophys. Res. Comm. 215: 744-749.

Tanaka, T., Kato, N., Nakagawa, M., Ootsuyama, Y., Cho, M. J., Nakazawa, T., Hijikata, M., Ishimura, Y. and Shimotohno, K. 1992. Molecular cloning of hepatitis C virus genome from a single Japanese carrier: sequence variation within

the same individual and among infected individuals. Virus Res. 23(1-2): 39 53.

Yasmeen, A., Siddiqui, A. A., Hamid, S., Sultana, T., Jafri, W. and Persson, M. A. 2009. Genetic variations in a well conserved 5'-untranslated region of hepatitis C virus genome isolated in Pakistan. J. Virol. Methods, 160: 38-47.

Wu, X., Robotham, J. M., Lee, E., Dalton, S., Kneteman, N. M., Gilbert, D. M., and Tang, H. 2012. Productive hepatitis C virus infection of stem cell-derived hepatocytes reveals a critical transition to viral permissiveness during differentiation. PLoS Pathog. 8: e1002617.