# Sequence Analysis of the IRES-Loop III Region of Hepatitis C Virus

**Takayoshi Sasano** uaa@bi.a.u-tokyo.ac.jp
**Jun-Ichi Sagara** jun@bi.a.u-tokyo.ac.jp
**Shugo Nakamura** shugo@bi.a.u-tokyo.ac.jp
**Mitsunori Ikeguchi** ike@bi.a.u-tokyo.ac.jp
**Kentaro Shimizu** shimizu@bi.a.u-tokyo.ac.jp

Department of Biotechnology, The University of Tokyo
1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan

## 1 Introduction

Human hepatitis C virus (HCV) causes chronic infection of the liver and has been linked to the development of hepatocellular carcinoma. The viral genome consists of a single-stranded, positive-sense RNA approximately 9,500 nucleotides in length with an open reading frame. The 5' noncoding region (5'NCR), which varies in length from 332 to 341 nucleotides, is followed by a long open reading frame encoding a polyprotein of about 3,000 amino acids. The process by which this region initiates translation is controlled by an internal ribosomal entry segment (IRES) that occupies most of the 5'NCR. HCV is classified into seven genotypes, and the IRES of HCV genotype 3 is known to reduce translation efficiency. The loop III region of IRES are thought to influence translation efficiency, but had not yet been compared between genotypes. We therefore analyzed the IRES loop IIIa–IIId region by using principal component analysis (PCA) and multidimensional scaling analysis (MDS) in an attempt to compare the sequences of the region of all genotypes.

## 2 Material and methods

*HCV-5'NCR Sequences.* A total of 92 sequences were collected from Genbank: 36 of genotype 1, 24 of genotype 2, 10 of genotype 3, 14 of genotype 4, 5 of genotype 5, 2 of genotype 6, and 1 of genotype 7.
*Computer Programs for Sequence Analysis.* Nucleotide sequences were aligned using the CLUSTAL W program. After alignment, sequences containing loop IIIa–IIId region of the HCV IRES were analyzed by the program of PCA and MDS developed by J.-I. Sagara [2].

## 3 Results

Almost all the bases in the loop IIIa–IIId region of different isolates are conserved, but the results of this analysis show that the bases in some parts are not. The sequence of the stem between loop IIIa and loop IIIb (stem IIIab) is different in each genotype.

HCV genotype 3 is clearly different from others. In stem IIIcd, the nucleotides 247 to 249 are respectively “U”, “C”, and “A”, and in stem IIIab the 175th, 178th, and 221th nucleotides are respectively “C”, “U”, and “A” (Fig. 2). *Note*: These nucleotide positions are illustrated in Fig. 1.

# 4 Discussion

E. Buratti *et al.* have reported that translation efficiency of HCV genotype 3 is inferior to genotypes 1 and 2 [1]. Our analysis shows that the IRES of HCV genotype 3 IRES is clearly different from that of the other HCV genotypes. This difference is expected to affect the translation efficiency.

In this work we used only the loop IIIa–IIId region, but we are also going to analyze other loop regions.

# References


[1] Buratti, E., Gerotto, M., Pontisso, P., Alberti, A., Tisminetzky, S.G., and Baralle, F.E., In vivo translational efficiency of different hepatitis C virus 5'-UTRs, *FEBS Letters*, 411:275–280, 1997.

[2] Sagara, J.-I., Shimizu, S., Kawabata, T., Nakamura, S., Ikeguchi, M., and Shimizu, K., The use of sequence comparison to detect 'identities' in tRNA genes, *Nucleic Acids Research*, 26:1974–1979, 1998.


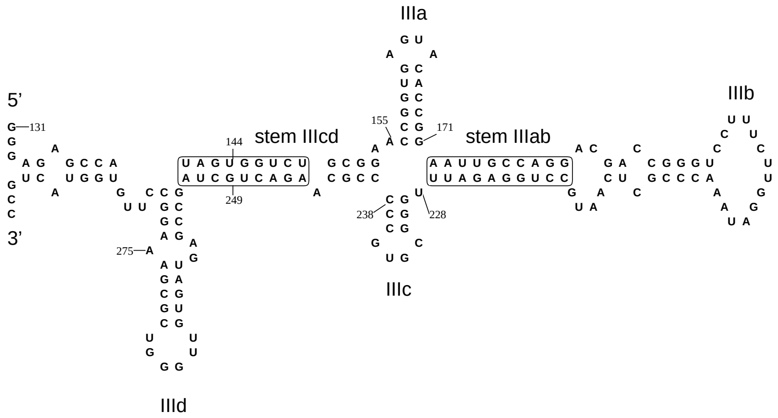


Figure 1: Model secondary structure of IRES-loop III of HCV genotype 1.

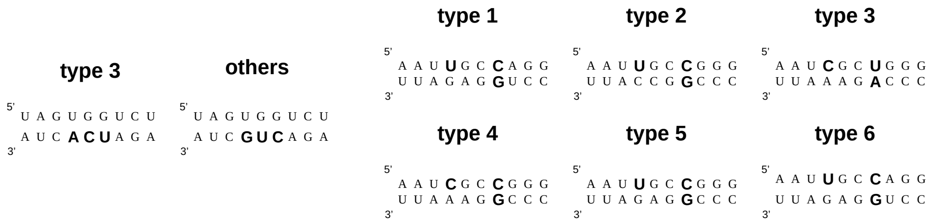


Figure 2: Difference in genotypes: stem IIIcd (left) and stem IIIab (right).

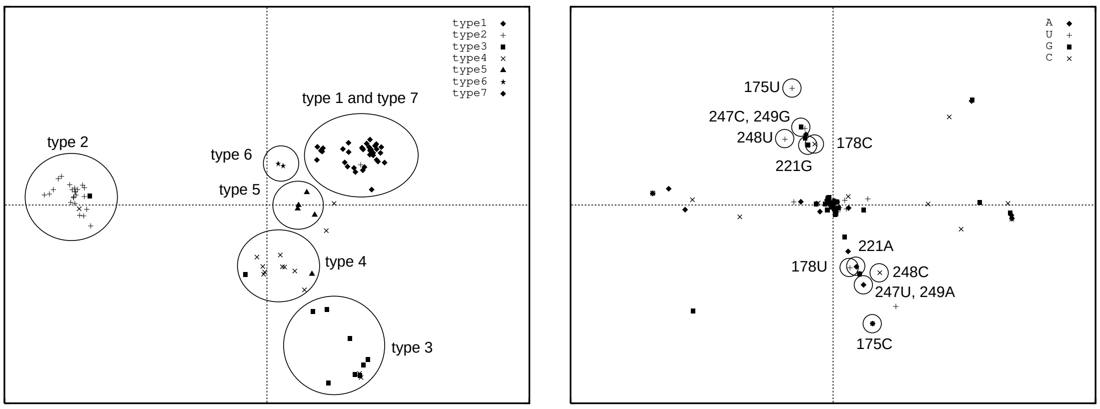


Figure 3: Results of PCA and MDS: sequence plot (left) and base plot (right).