

ABSTRACT :

The pandemic Hepatitis C concerns 170 million people (OMS 1999). It is caused by a virus of the *Flaviviridae* family. One difficulty to study this virus is that no effective culture system exists for it. It's known only by its genome which presents a genetic variability leading to viral quasi-species within the same host. Moreover, viral sequences have been classified into 6 clades and a lot of subtypes which are related with the disease severity or the resistance to interferon-ribavirin treatment. To analyze this variability and to establish correlations between sequences and pathology, we have developed an integrated viral and clinical bioinformatic tool which brings together analysis methods and a Hepatitis C virus (HCV) sequence database. The database, available on the Web (<http://hepatitis.ibcp.fr>), joins more than 10000 sequences monthly extracted from EMBL. It is divided into 5 components : (1) HCVWEB is the Web interface that gives access to other components, (2) HCVDB includes all the HCV sequence databases and associated management tools, (3) HCVFORM contains forms to deposit clinical data in the private database available to French Hepatitis Network members, (4) HCVSRS allows keywords queries, and (5) HCVSA is the sequence analysis component for which the NPS@ (<http://npsa-pbil.ibcp.fr>) server has been developed. NPS@ integrates in a user friendly and simple Web interface 28 bioinformatic methods (*e.g.* BLAST, CLUSTAL W, PHD, SOPM) for biological macromolecule sequence and structure analysis and 12 sequences, patterns and structures databases. Every day, 2000 analyses are carried out with NPS@. NPS@ and HCVDB form a coherent set and are reference tools to rationalize biologist and virologist research.