Perspectives

HEPSEQ – AN INTEGRATED HEPATITIS B EPIDEMIOLOGY AND SEQUENCE ANALYSIS PLATFORM

R Myers (richard.meyers@hpa.org.uk)¹, S Gnaneshan¹, S Ijaz², R Tedder², M Ramsay³, J Green¹, HepSEQ Steering Committee

1. Bioinformatics Unit, Centre for Infections, Health Protection Agency, London, United Kingdom

2. Virus Reference Department, Centre for Infections, Health Protection Agency, London, United Kingdom

3. Immunisation Department, Centre for Infections, Health Protection Agency, London, United Kingdom

Hepatitis B virus (HBV) is a major human pathogen. The outcome of acute hepatitis B is variable but usually followed by a complete recovery. A small proportion of infections in adults and a higher proportion of infections in early childhood continue in a chronically infected state in which the virus persists in the liver. Patients with chronic hepatitis B usually have no initial symptoms of infection, but over time the major disease sequelae, cirrhosis and hepatocellular carcinoma, can develop. It is estimated that some 350 million people worldwide are currently chronically infected with HBV, but many more will have been infected and recovered.

HBV exists in its human host as eight clusters of viruses (genotypes A to H [1]), each cluster displaying a similarity of sequences and variable antigenicity (serotypes) [2]. Traditionally, HBV genotypes have a distinct geographical distribution, with genotypes A and D predominant in Europe, the Middle East and India, and genotypes B and C in Asia and the Far East. Genotype E is mainly found in West Africa, and genotypes F and H are associated with Central and South America. There is some evidence that different genotypes may be associated with alternative disease profiles and differing responses to antiviral therapy [3-5]. The influx of non-European HBV genotypes represents increasing population movement or migration and may have a significant effect on management of the disease within Europe. The mixing of genotypes within a European population may also promote increased genetic diversity within HBV (due to recombination), the epidemiological effects of which are impossible to predict. Therefore, surveillance of HBV genotypes is important within a national and a European context.

HepSEQ (<http://www.hepseqresearch.org>) [6] is a freely accessible web resource for the public health aspects of HBV management, with specific focus on epidemiological, virological, clinical, nucleotide sequence and mutational aspects of HBV infection. HepSEQ is comprised of a relational database with a web-enabled interface, which allows inserting and retrieving epidemiological and sequenced-based information. The database currently contains 1,769 patient records from acute and chronic hepatitis B cases in the United Kingdom (UK) and 2,182 sequences, of which 1,679 cover the surface/polymerase region and 497 cover the X/pre-core/core region. The HepSEQ website also provides access to tools that predict the genotype and characterise the genetic polymorphisms of sequences entered by the user. The HBV genotyper tool assigns a genotype to an HBV sequence provided by the user. Sequence alignment and statistical modelling

are used to ensure that predictions of the genotype are accurate. HepSEQ allows rapid and dynamic generation of genotypic data within the database and contains 43.6% A, 5.6% B, 12.4% C, 28.4% D and 9.3% E genotypes. The presence of approximately 20% non-A or non-D genotypes highlights the changing dynamic of HBV sequences within the UK and raises questions about the origins and transmission of these infections.

The majority of countries in Europe have a low prevalence of chronic hepatitis B infection, and the acquisition of hepatitis B is often associated with medical interventions or with specific adult risk behaviour such as sex between men, injecting drug use. As HepSEQ contains information on viruses associated with acute hepatitis B infections, it has the potential to further highlight patterns of disease transmission. The database is currently being enriched by attempting to link in exposure categories so that viruses associated with different routes of transmission in the UK can be identified. This will help to identify the emergence of new variants in specific sub-populations and to evaluate the impact of control measures to reduce the incidence, and therefore the prevalence of specific variants, in these populations.

Antiviral therapy targeting the reverse transcription function of polymerase is increasingly used in clinical practice to suppress viral replication. Six polymerase inhibitors have either been licensed or are in development for treatment of chronic hepatitis B: lamivudine, adefovir dipivoxil, entecavir, tenofovir, telbivudine and clevudine. Not surprisingly, long term use of mono or dual therapy in the face of continued replication is associated with viral mutational escape from the drug and this can be monitored by direct sequencing of the polymerase gene [7,8]. Additional variability in the genome has been shown to arise as a result of the natural emergence of strains which may have a selective advantage during the course of chronic HBV infection in a patient. These include pre-core mutants, deletions in the core gene, and mutations in the preS1 and pre S2 regions [9]. It is speculated that these variants are driven by the immune system but it currently remains unknown which, if any, are clinically significant.

The necessity to recognise and interpret HBV anti-viral resistance mutations has driven the development of a mutation annotation tool within HepSEQ. This publicly available tool takes a polymerase nucleotide sequence as input, performs a genotype assignment and compares both the nucleotide and translated amino acid sequence

to the appropriate genotype consensus sequence. The alignment of both nucleotide and amino acid query sequence and the genotype consensus are displayed within the web browser along with a table of mutations that are recognised as causing resistance to HBV inhibitors. When resistance mutations within the query sequence are detected, these mutations are listed along with the inhibitors that they impact upon and a literature citation for that mutation.

As well as providing the user with additional information, continuous dynamic assessment of the relationships between HBV polymorphisms and treatment histories may allow new resistance mutations to be identified. The mutation annotator tool demonstrates the clinical usefulness of HepSEQ and as a consequence of this we are constantly updating the tool to incorporate new information as it becomes available.

The resistance reporting aspect within HepSEQ has proved hugely successful and we are in the process of developing further the clinical and surveillance aspects of the application. Repeated resistance testing generates longitudinal sequence and clinical datasets for individual patients. HepSEQ has the capacity to capture this repeat testing information and we are designing interfaces that will present temporal data such as viral load, resistance mutations, treatment history and alanine transaminase (ALT) levels for a single patient. This interface will provide a single unified mechanism for collating, presenting and interpreting all available patient data, thus facilitating clinical treatment decisions.

Monitoring the prevalence of HBV drug resistance mutations and their transmission, coupled with correlates of disease progression, requires an extensive epidemiological data set that is representative of the population. At present, only HepSEQ contains such a dataset, which covers only the UK. However, these analyses are critical for the whole of Europe, and we are actively seeking European partners to facilitate this. We envisage that the tools within HepSEQ will provide an invaluable resource for other countries to analyse and interpret these data.

We hope that other national centres in Europe will be encouraged to contribute information on hepatitis B cases. This will allow the tracking of specific strains across Europe, and identify links between specific risk groups due to migration and travel. In addition, the increased use across Europe of selective and universal hepatitis B vaccination and of antiviral therapy may contribute to the emergence of specific mutations. Tracking these strains will have particularly important implications for the prevention and management of hepatitis B in Europe. The identification of increasing numbers of clinically important mutations amongst acute cases in one country may provide warning to neighbouring countries.

Acknowledgements

This work was supported by funding from the Department of Health (UK) and by the Health Protection Agency (UK).

References

1. Norder H, Couroucé AM, Coursaget P, Echevarria JM, Lee SD, Mushahwar IK, et al. Genetic diversity of hepatitis B virus strains derived worldwide: genotypes, subgenotypes, and HBsAg subtypes. *Intervirology*. 2004;47(6):289-309.
2. Couroucé AM, Lee H, Drouet J, Canavaggio M, Soulier JP. Monoclonal antibodies to HBsAg: a study of their specificities for eight different HBsAg subtypes. *Dev Biol Stand*. 1983;54:527-34.

3. Ding X, Mizokami M, Yao G, Xu B, Orito E, Ueda R, et al. Hepatitis B virus genotype distribution among chronic hepatitis B virus carriers in Shanghai, China. *Intervirology*. 2001;44(1):43-7.
4. Fujie H, Moriya K, Shintani Y, Yotsuyanagi H, Iino S, Koike K. Hepatitis B virus genotypes and hepatocellular carcinoma in Japan. *Gastroenterology*. 2001;120(6):1564-5.
5. Kao JH, Wu NH, Chen PJ, Lai MY, Chen DS. Hepatitis B genotypes and the response to interferon therapy. *J Hepatol*. 2000;33(6):998-1002.
6. Gnaneshan S, Ijaz S, Moran J, Ramsay M, Green J. HepSEQ: International Public Health Repository for Hepatitis B. *Nucleic Acids Res*. 2007;35(Database issue):D367-70.
7. Tenney DJ, Levine SM, Rose RE, Walsh AW, Weinheimer SP, Discotto L, et al. Clinical emergence of entecavir-resistant hepatitis B virus requires additional substitutions in virus already resistant to Lamivudine. *Antimicrob Agents Chemother*. 2004 ;48(9):3498-507.
8. Lai CL, Dienstag J, Schiff E, Leung NW, Atkins M, Hunt C, et al. Prevalence and clinical correlates of YMDD variants during lamivudine therapy for patients with chronic hepatitis B. *Clin Infect Dis*. 2003; 36(6):687-96.
9. Günther S. Genetic variation in HBV infection: genotypes and mutants. *J Clin Virol*. 2006; 36 Suppl 1:S3-S11.

This article was published on 8 May 2008.

Citation style for this article: Myers R, Gnaneshan S, Ijaz S, Tedder R, Ramsay M, Green J, HepSEQ Steering Committee. HepSEQ – an Integrated Hepatitis B Epidemiology and Sequence Analysis Platform. *Euro Surveill*. 2008;13(19):pii=18866. Available online: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=18866>