

CC Lau 劉智超  
 WY Ha 夏慧茵  
 PYK Yue 余英傑  
 SH Chui 崔紹漢  
 AKK Chui 徐家強  
 RNS Wong 黃岳順

# A high-throughput detection method for hepatitis B virus mutations in the Chinese population

## Key Messages

1. Arrayed primer extension (APEX) has been successfully developed as a high-throughput method for single nucleotide polymorphisms (SNPs) present in the hepatitis B virus (HBV) genome.
2. Co-existence of wild-type and mutant genotypes in one SNP was successfully identified by APEX, which was not identified by conventional DNA sequencing.
3. Thirty-three patients were involved in this study; APEX results showed A1762T and G1764A double mutations and G1896A mutation are the most prevalent SNPs present in HBV genomes.
4. APEX can be developed as a SNP screening tool to monitor other infectious agents.

*Hong Kong Med J* 2008;14(Suppl 5):S4-7

Department of Biology, Hong Kong Baptist University, Kowloon, Hong Kong SAR, China

CC Lau, PYK Yue, RNS Wong  
 Research and Development Division, School of Chinese Medicine, Hong Kong Baptist University, Kowloon, Hong Kong SAR, China

WY Ha  
 Diagnostix Medical Centre Ltd, Hong Kong SAR, China

SH Chui  
 Department of Surgery, The Chinese University of Hong Kong, Shatin, NT, Hong Kong SAR, China  
 AKK Chui

RFID project number: 01030072

Principal applicant and corresponding author:  
 Prof Ricky NS Wong  
 Department of Biology, Hong Kong Baptist University, Kowloon, Hong Kong SAR, China  
 Tel: (852) 3411 7057  
 Fax: (852) 3411 5995  
 E-mail: rns Wong@hkbu.edu.hk

## Introduction

Hepatitis B virus (HBV) is the infective agent responsible for the widespread liver disease in humans. According to the World Health Organization, about 2 billion persons worldwide have been infected with HBV; more than 350 million have been chronically infected, of whom three-quarters are Asian.<sup>1</sup> Chronic infection has been associated with a high risk of developing liver cirrhosis and hepatocellular carcinoma (HCC),<sup>2</sup> and kills one million persons each year.<sup>1</sup>

Evidence suggests that chronic HBV infection is associated with the development of point mutations or single nucleotide polymorphisms (SNPs) in the four open reading frames (ORFs) of the HBV genome, namely the C, S, P and X genes (Fig 1). Studies show that the development of SNPs in these ORFs is correlated with the occurrence of HCC. Therefore, different methods have evolved for the detection of HBV SNPs to monitor disease development. Conventional methods are generally not practical for detecting multiple SNPs. Also, for clinical and diagnostic analysis, simple techniques with high-throughput potential, high reproducibility, sensitivity, and specificity are required.

DNA microarray technology has been applied not only to study gene expression, but also for large-scale sequence analysis and mutation detection. Here we describe a DNA-based method for rapid detection of multiple mutations in a microarray format, known as arrayed primer extension (APEX). This method is based on the hybridisation of HBV DNA to an array of complementary DNA spotted on a glass surface, followed by the incorporation of fluorescent-labelled dideoxynucleotides mediated by an enzyme (Fig 2). Thus, APEX can be regarded as the miniature form of DNA sequencing. With prior optimisation, detection of known SNPs is possible in a single reaction by designing hundreds to thousands of short specific sequence of DNAs on the slide. One advantage of using APEX for detecting SNPs is the high signal-to-noise ratio, which allows rapid and accurate determination; APEX is a promising technique for large-scale analysis under optimal conditions. It may also be possible to apply this technique to detect SNPs in other bacterial or viral genomes.

## Methods

This study was conducted from September 2004 to May 2006.

### *Selection of patients in the Chinese population for hepatitis B virus DNA extraction*

The blood samples from HBV-infected patients with HBV DNA levels equal to or higher than 10<sup>4</sup> copies/mL were selected and the viral DNA were extracted. The DNA amount and quality was quantified and checked, and finally samples from 33 patients were selected.

### *Polymerase chain reaction for hepatitis B virus DNA amplification*

The four ORFs of HBV were separately amplified from serum HBV DNA using polymerase chain reaction (PCR).

### *Preparation of short specific sequences of DNA and array printing*

Short specific sequences of DNA for detecting the 30 SNPs in the four ORFs

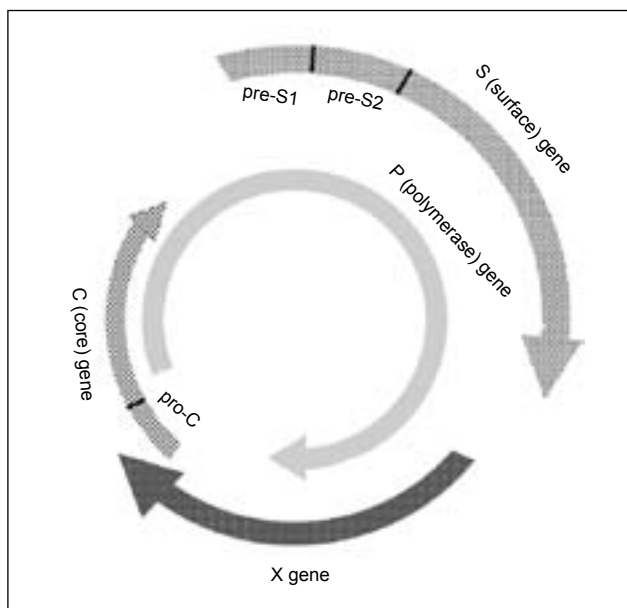


Fig 1. Schematic representation of the HBV genome<sup>3</sup>

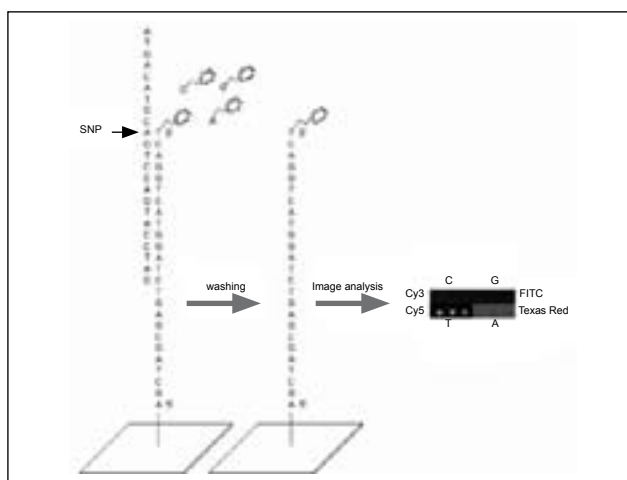


Fig 2. Principle of arrayed primer extension

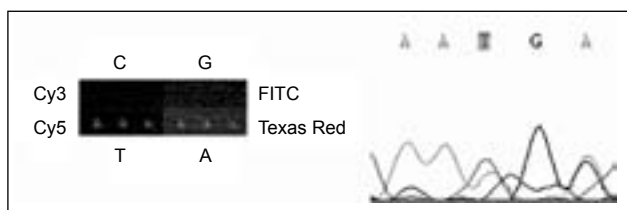


Fig 3. Typical comparisons of arrayed primer extension (APEX) single nucleotide polymorphism (SNP) genotyping and DNA sequencing

The coexistences of wild-type and mutant genotypes in one SNP of a single individual were successfully genotyped by APEX, but the same cases were not detected by DNA sequencing

were designed according to the genomic sequence of the HBV (GenBank accession no. NC\_003977). Each short

specific sequence was spotted in triplicate onto special glass slides.

### Template preparation for APEX

The purified PCR products were enzymatically fragmented.

### Arrayed primer extension

Under optimal conditions, the fragmented PCR products were allowed to bind to the short complementary sequence of DNA in the presence of four different fluorescent-labelled dideoxynucleotides and DNA polymerases. A standard APEX reaction (but without an HBV PCR product) was used to check if there were self-extending oligonucleotides or non-specific background signals.

### Image capture and analysis

The slides were scanned using a confocal laser scanner. The intensity of each spot on the microarray was quantified and subsequently analysed. The specificities and sensitivities were calculated according to the following equations:

Specificity = number of true negatives / (number of true negatives + number of false positives)

Sensitivity = number of true positives / (number of true positives + number of false negatives)

### DNA sequencing

The sequences of the PCR products were determined by DNA sequencing, which acted as the 'gold standard'.

## Results

### Single nucleotide polymorphisms genotyping

This study involved 33 patients infected with HBV. The results were compared with the 'gold standard' DNA sequencing data for method validation. Typical comparisons of APEX SNP genotyping and DNA sequencing on different nucleotides are shown in Fig 3. Among the 30 SNPs present on the whole HBV genome, specificities ranged from 100 to 66% and sensitivities from 100 to 59%.

### Coexistence of wild-type and mutant genotypes in a single nucleotide polymorphism

The coexistences of wild-type and mutant genotype in one SNP of a single individual were successfully genotyped by APEX. However, the same cases were not discernable by DNA sequencing (Fig 3). These cases happened most frequently on SNP G1896A in the C gene (9 of 33 patients), while there was no significant co-existence for other SNPs.

### Prevalence study by arrayed primer extension

According to the APEX results, the most frequent mutation (23 of 33) was G1896A in the C gene. All patients carried the G1896A mutant SNP, which creates a stop codon TAG in codon 28 of the precore region. The other 'hotspots' were A1762T and G1764A double mutations present in the precore region, in 22 and 20 patients respectively. In addition, some frequent mutations are also present in the C

gene. These included A1979G (codon 27), A2159G (codon 87), A2189C (codon 97) and C2288A (codon 130), which comprised 18%, 27%, 27%, and 30% of the respective mutations. For the P gene, the SNPs of A739G and G741T which lead to lamivudine resistance in the YMDD motif were only found in three and two patients respectively. For the S gene, one patient expressed A530G and another carried a G546A mutation, showing that these were minor mutations. For the X gene, the eight nucleotide deletions (nt 1768 -1775) commonly present in the COOH terminal were not found in our pool of samples. However, 33% (11/33) of our patients had T1464C/G mutations.

## Discussion

Liver cancer is the third commonest cause of cancer mortality in Hong Kong, of which HBV infection is one of the major causes. Many studies show that SNPs present on the HBV genome correlate with liver cancer development. Therefore, large-scale population studies to determine the prevalence of SNPs, its natural course, and response to treatment are urgently needed. Using the APEX strategy to detect mutations in the HBV genome will be applicable in this regard, as it allows multiplex screening of different SNPs in a single experiment. By designing short specific sequences of DNA, different SNPs in the HBV genome could be genotyped accordingly. One feature of APEX is the design of a redundant set of short specific sequences of DNA against the same SNPs, to ensure the highest statistical significance and the correct interpretation of signals. Binding efficiency of HBV DNA to short complementary sequences of DNA is increased by fragmentation. In the presence of DNA polymerases, fluorescent-labelled ddNTPs corresponding to the SNPs are extended accordingly. The labelled short specific sequence of DNA is subsequently visualised by a laser confocal scanner. The advantage of APEX is that all the SNPs present in an ORF can be detected in a single-step reaction.

According to our results, the most common co-existence of wild-type and mutant SNPs within individual patients are A1762T, G1764A and G1896A present in the C gene. However, these coexistences are undetectable by DNA sequencing. This is due to the ambiguity in base-calling by the instrument, during the presence of double peaks. This limitation affects the reliability of direct sequencing for the identification of nucleotides in these SNPs, at least under automated conditions. The coexistence of wild-type and mutant genotypes may imply the transition stage of chronic infection or that one viral species is dominant over the others.<sup>3</sup> Thus, the detection of both dominant and non-dominant viral species by APEX, which can be missed by DNA sequencing, might be significant for monitoring disease development.

The APEX results also showed that A1762T, G1764A and G1896C are the most frequent SNPs present in the precore/basal core promoter region as well as the whole

HBV genome. The presence of G1896A mutation causes suppressed external core antigen (HBeAg) secretion. As HBeAg is a major humoral and cellular target, HBeAg negativity is regarded as an indicator of immunological clearance of HBV and rapid viral replication. The A1762T and G1764A double mutations suppress the transcription of precore mRNA and arrest HBeAg production. Therefore, the detection of this double mutation is useful for monitoring disease development.

In addition, some frequently encountered SNPs present in the C gene may disturb the level of HBeAg expression, causing evasion of the host's immune clearance and thus directly affect the activity of HBV-induced liver disease.

For the S gene, no significant SNPs were detected in our samples. However, since the SNPs present in this region lead to HBV replication in vaccinated patients, the detection of SNPs present on the S gene will be beneficial for vaccine development.

For the P gene, although no significant SNPs were detected in our samples, the detection of SNPs present in it was crucial for patients who receive lamivudine therapy during the course of HBV infection. Lamivudine is a reverse-transcriptase inhibitor and has antiviral activity against HBV. However, recent studies reported the emergence of lamivudine-resistant HBV during treatment. Therefore the detection of lamivudine-resistance-related SNPs in this region can be an alternative means of monitoring the development of lamivudine-resistant strains of HBV.

For the X gene, re-sequencing by APEX was used to detect the presence of HBx COOH-terminal in our study. The design of an APEX array for comparison sequencing is straightforward and complementary to the gene of interest.<sup>4</sup> According to our results, none of the patients had HBV X gene COOH-terminal deletions. However, studies demonstrated that the HBx COOH-terminal is responsible for controlling host cell viability and proliferation. In addition, a prevalent HBV X gene mutant T1464C/G present in Taiwanese patients with liver cirrhosis and HCC was also included in our system. Approximately 33% (11/33) of patients had the T1464C/G mutation. Development of this mutant might represent a strategy of the virus to escape immune surveillance, so a reliable genotyping method will be beneficial for studying the process of multiple-step hepatocarcinogenesis.

In conclusion, the simultaneous and high-throughput screening of SNPs in the HBV genome by APEX enables large-scale diagnostic analysis, which is an alternative to genotyping by DNA sequencing.

## Acknowledgements

This study was supported by the Research Fund for the Control of Infectious Diseases, Food and Health Bureau,

Hong Kong SAR Government (RFCID: 01030072).

## References

1. Hepatitis B. World Health Organization. Available from: WHO website: <http://www.who.int/mediacentre/factsheets/fs204/en/>. Accessed 10 Oct 2006.
2. Tiollais P, Pourcel C, Dejean A. The hepatitis B virus. *Nature* 1985;317:489-95.
3. Chen WN, Oon CJ. Pre-S deletion mutants co-exist with wild-type virus at low viral replication stage in Singapore chronic hepatitis B virus carriers. *Curr Microbiol* 2002;44:145-7.
4. Shumaker JM, Tollet JJ, Filbin KJ, Montague-Smith MP, Pirrung MC. APEX disease gene resequencing: mutations in exon 7 of the p53 tumor suppressor gene. *Bioorg Med Chem* 2001;9:2269-78.