
**FULL GENOME ANALYSIS AND FUNCTIONAL
CHARACTERIZATION OF MUTANTS OF HEPATITIS B VIRUS
ISOLATES FROM SOUTHERN AFRICAN BLACKS.**

Gerald Chiafiinii Kimbi

**Thesis submitted in compliance with the requirements for the degree of
Doctor of Philosophy in the Faculty of Health Sciences
at the University of the Witwatersrand**

Johannesburg 2005

DECLARATION

I, Gerald Chiafiinii Kimbi declare that this thesis is my own work. It is being submitted for the degree of Doctor of Philosophy in Medicine, in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at this or any other university.

.....

.....day of.....2005

DEDICATION

This thesis is dedicated to my father Kimbi Joseph, who taught me to persevere and to face challenges with faith, courage, prayer and humility. It is also dedicated to my mother Kimbi Prudencia, who taught me that success is not all I need in life, leaving the success is equally as important.

PUBLICATIONS AND PRESENTATIONS

The following articles and conference presentations arose from work presented in this thesis.

Research Articles

1. Ahn, S. H., Kramvis, A., Kawai, S., Spangenberg, H. C., Li, J., **Kimbi, G.**, Kew, M., Wands, J. and Tong, S. (2003). Sequence variation upstream of precore translation initiation codon reduces hepatitis B virus e antigen production. *Gastroenterology* **125**, 1370-8.
2. **Kimbi, G. C.**, Kramvis, A. and Kew, M. C (2004). Distinctive sequence characteristics of sub-genotype A1 isolates of Hepatitis B virus from South Africa. *J Gen Virol* **85**, 1211-20.
3. **Kimbi, G. C.**, Kramvis, A. and Kew, M. C (2005). Integration of Hepatitis B Virus DNA into chromosomal DNA during acute hepatitis B. *World J Gastroenterol* [In Press].

Presentations

1. 11th International symposium on Viral Hepatitis and Liver Diseases, Sydney, Australia, 6th to 10th, April 2003, Integration of HBV DNA in Acute Hepatitis B. **Kimbi G.C**, Kramvis A, Kew M.C.
2. International meeting of the molecular biology of the hepatitis B viruses, Bergamo, Italy, 7th to 10th, September 2003, Analysis of 23 Hepatitis B Virus isolates from South Africa. **Kimbi G.C**, Kramvis A, Kew M.C.
3. 2004 meeting of the molecular biology of the Hepatitis B virus, Marine Biological Laboratory, Woods Hole, Massachusetts, USA, 24th to 27th October 2004, Effect of the G1888A mutation of subgenotype A1 on the translation of the core protein. **Kimbi G.C**, Kramvis A, Kew M.C.
4. 8th Beatson International Cancer Conference, "Human Cancer: Modelling the Disease", Glasgow, Scotland, 19th to 22nd, June 2005, Characterisation of hepatitis B virus mutants in hepatocellular carcinoma in southern Africa. Skelton M, **Kimbi G.C**, Kramvis A, Kew MC.

ACKNOWLEDGMENTS

Any project, no matter how individual, will almost certainly require input, assistance or encouragement from others. My project is no exception. I would like to acknowledge the following for their support throughout this project;

Firstly, my supervisor Prof Anna Kramvis exemplifies the high quality scholarship to which I aspire. I would like to express my gratitude to her. Her expertise, understanding, and patience, added considerably to my postgraduate experience. I appreciate her immense knowledge and skill in many areas (e.g., vision, team management, and mentoring), and her assistance in writing reports (i.e., grant proposals, scholarship applications, journal articles and this thesis), which have on some occasions made me "GREEN" with envy.

I was particularly fortunate to have as co-supervisor Prof Michael C Kew. He is clearly an expert in his field. I have benefited tremendously from observing him, chatting to him and conducting research with him.

I would also like to thank Ms Alice Chen. Her help in tissue culture and plasmid construction could not have gone unnoticed.

Next I wish to thank the members of the Molecular Hepatology Research Unit (MRHU). Each individual provided insights that guided and challenged my thinking, substantially improving the finished product.

In addition to the technical and scientific assistance above, I received equally important assistance from my family and friends. My mother Mrs Kimbi Prodencia provided on going support throughout the thesis process (most of it was spiritual as she passed away). My father Mr Kimbi Joseph instilled me, from an early age with the desire and skills to obtain the Doctorate.

In conclusion, I recognize that this research would not have been possible without the financial assistance of the Poliomyelitis Research Foundation (PRF) and the University of the Witwatersrand.

ABSTRACT

Hepatitis B virus (HBV) infection is endemic in several parts of the world. Africa is one of the most affected continents with about 50 million chronically infected individuals.

Data on complete sequences of HBV genomes from southern Africa is limited. HBV genotypes show a distinct geographical distribution. Hence, studies based on genotypes from other regions of the world cannot be extrapolated to southern Africa.

The aim of this study was to amplify and characterize complete HBV genomes from southern African Blacks.

The complete genome of HBV isolates from eighteen asymptomatic carriers and five acute hepatitis patients was amplified and the resulting amplicons were cloned and sequenced. There were no differences between isolates from asymptomatic carriers and those from acute HBV infected patients, suggesting that differences in disease presentation may be as a result of host and environmental factors.

Seventeen of 23 isolates (73%) belonged to genotype A and six to genotype D, confirming the predominance of genotype A in the South African Black population. Following phylogenetic analysis, genotype A separated into two subgenotypes A1 and A2. Of the 17 genotype A isolates sequenced in the present study 15 belonged to subgenotype A1, confirming that subgenotype A1 is the dominant strain found in southern African Blacks. The mean nucleotide divergence of subgenotype A1 was

greater than that of subgenotype A2, suggesting that this subgenotype has been endemic for a longer time in the South African Black population.

The molecular characterization of the 15 subgenotype A1 and subgenotype A2 isolates confirmed the presence of the previously described amino acid residues that distinguish subgenotype A1 from the remainder of genotype A in the S and polymerase genes. Moreover, the large number of subgenotype A1 isolates sequenced allowed identification of additional nucleotide and amino acid changes that are characteristic of subgenotype A1 in the other open reading frames. In particular, mutations at positions 1809-1812 that alter the Kozak sequence of the precore/core open reading frame, and 1888A in the precore region were found exclusively in subgenotype A1 isolates. Unique sequence alterations of the transcriptional regulatory elements were also found in subgenotype A1 isolates. Amino acids differentiating the subgenotypes of genotype A from each other and from other non A genotypes were identified in the pres-S1 (Gln⁵⁴, Val⁷⁴, Ala⁸⁶, and Val⁹¹) region overlapping the spacer of the polymerase (Thr²³⁶, Gly²⁶⁸, Tyr²⁶⁹). These changes could affect binding of the virus to hepatocytes as well as viral DNA synthesis. Specific nucleotide changes within the core promoter/Enhancer II region, the encapsidation signal, S1 promoter, and the S2 promoter were also identified in subgenotype A1.

Double or triple nucleotide substitutions at position 1809-1812 were found only in subgenotype A1 isolates. These substitutions occur as a stable trait of this subgenotype. These substitutions were shown to reduce HBeAg translation by a ribosomal leaky

scanning mechanism. The most common of these variants, the 1809^T1812^T mutation resulted in a 20% reduction in HBeAg translation.

A G to A mutation at nucleotide 1888 of the precore region was also unique to subgenotype A1 isolates. This nucleotide substitution introduces an out-of-frame AUG, creating an overlapping upstream open reading frame (uORF), which terminates five nucleotides downstream from the core AUG. Functional studies were carried out using fusion plasmids in which the complete HBV core gene was fused at the amino terminal of a GFP expression plasmid. The G1888A mutation resulted in a 18.75% decrease in core protein translation. The decrease in core protein production was shown to result from leaky scanning at the 1888 AUG as well as stalling of terminating ribosomes at the stop of upstream open reading frames (uORF).

Genotype D was the second most abundant genotype found in South African blacks. Six complete HBV genotype D genomes were sequenced in this study and represent the first complete genomes of genotype D isolates from South Africa. All six genotype D isolates from South Africa clustered together in a separate minor clade (Dsa), following phylogenetic comparison with other genotype D isolates from GeneBank. Amino acids specific to Dsa were identified in the polymerase region. Specific nucleotide changes within the enhancer I/ X promoter and the S2 promoter were also identified in Dsa.

During the course of the study, in addition to amplifying the subgenomic portions of the HBV viral genome, chromosomal DNA was fortuitously amplified from the serum of three patients with acute HBV infection and one with fulminant hepatitis. In one of the acute hepatitis B patients, HBV DNA was shown to be integrated into chromosome 7q11.23 in the WBSCR1 gene within a region commonly deleted in patients with Williams-Beuren syndrome. The identified viral host junction was situated in the cohesive overlap region of the viral genome, at a preferred topoisomerase 1 cleavage site. This is one of the very few times that a HBV integrant has been described in the acute phase of the disease.

Because HBeAg functions as an immuno-regulator, reduction of HBeAg translation as a result of the 1809-1812 mutants could result in a more vigorous immune response during infection. However, the G1888A mutation may counterbalance this by reducing the production of core antigen, which is highly immunogenic.

Although HBV is hyperendemic in South African Blacks and there is a correspondingly high incidence of hepatocellular carcinoma, prior to this study there was a paucity of data on the complete genomes of strains circulating in this population. Genotype A, specifically subgenotype A1 was found to predominate over genotype D. The sequence of subgenotype A1 and South African genotype D isolates has been fully characterized. The higher intragroup divergence of subgenotype A1 would suggest that this strain is endemic for a longer time in the South African population whereas the lower intragroup divergence of genotype D isolates point to a more recent introduction

of these strains into the population. Work carried out in the Molecular Hepatology Research Unit, University of the Witwatersrand, has shown that subgenotype A1 is more hepatocarcinogenic than subgenotype A2 and other genotypes found in southern Africa. The variations and mutations in subgenotype A1, which may result in the low replication rate, high HBeAg-negativity and persistence of this strain, were confirmed and identified in this study and may contribute to the high hepatocarcinogenic potential of this strain. Fortuitously, HBV belonging to subgenotype A1 has been found integrated in the WBSCR1 gene of chromosome 7 in the serum of an acute hepatitis B patient, intimating that integration can take place during the early stages of infection. These findings are significant in our quest to understand, at a molecular level, the mechanisms involved in the pathogenesis of hepatocellular carcinoma in southern African Blacks, in relation to the HBV.