Full length genome characterisation of Hepatitis B virus isolates at Dr George Mukhari hospital in Pretoria, South Africa.

By

Rindidzani Edith Magobo

Submitted in fulfilment of the requirement for the degree of

MASTER OF MEDICAL SCIENCE (MEDICAL VIROLOGY)

In the

Department of Virology

School of Pathology, Faculty of Health Sciences

University of Limpopo, Medunsa Campus

Supervisor: Dr SG Selabe

Department of Virology

Medunsa Campus

Pretoria

SEPTEMBER 2011

DECLARATION

I, **Rindidzani Edith MAGOBO**, hereby declare that this work, unless where acknowledged, is my own and has not been submitted at any other institution for a degree purpose It is being submitted in fulfilment of the MSc (Med) in Medical Virology, in the Department of Virology at the University of Limpopo, MEDUNSA campus.

…… ……..

Signature of candidate Date

DEDICATION

This dissertation is dedicated to my precious son Tanganedzani, my parents Malindi and Luvhengo Magobo, my sisters Ndivhuwo, Ňaledzani, Tshianeo, my only brother Phathutshedzo for their never ending love, support and encouragement.

ACKNOWLEDGEMENTS

First and foremost I would like to give thanks to God almighty who gave strength to pull through all difficult situations I came across during my study period. I would like to express my sincere appreciation and gratitude to following people who made contribution during the course of my studies and without whose assistance this work would not have been a success:

Dr SG Selabe and Professor MJ Mphahlele for their superb supervision and all the support, motivation, courage and inspiration you gave. Thanks a lot for your guidance throughout the study.

To all the postgraduate students in the Department of Virology, your motivation, encouragement and assistance are highly appreciated.

Mr NJ Rakgole for technical assistance; training in laboratory techniques and assistance in analysis of the results.

Dr A Lukhwareni for being a good friend, brother and a mentor, and the scientific input from the beginning of this project to its very end, thanks a million times.

Special thanks to Ms L Netshifhefhe for proof reading this dissertation.

I also thank financial assistance from National Research Foundation (NRF), Poliomyelitis Research Foundation (PRF) and the University of Limpopo Postgraduate Funding.

Mashudu Netshikweta for being a loving, supportive and understanding partner throughout the process of dissertation preparation.

Finally, I would like to extend my appreciation to my family for putting their trust in me, their support, love and for giving courage; most importantly to my precious son for being the light at the end of the tunnel when things were not looking good.

ABSTRACT

Introduction: Sub-Saharan Africa is a region with hepatitis B virus (HBV) hyperendemicity with more than 8% HBsAg prevalence. An estimate of two billion people has been reported to carry HBV markers. HBV was associated with about 25% of annual deaths in Africa. HBV possesses a DNA polymerase which lacks proofreading mechanism. This results in highly variabillty and genetic diversity which poses a challenge for the diagnosis and therapeutic management of HBV infection. High mutation rate of HBV also has great implications on the development of drug resistant mutations. Moreover, HBV diversity represents a challenge for the sensitivity of immunological and molecular diagnostic assays. A number of studies on HBV full length genome have been conducted particularly in developed countries. Limited studies are available in Africa and South Africa. In South Africa, few studies have been done analysing the complete genome of HBV isolates from patients with asymptomatic carriers and fulminant hepatitis B (Owideru et al, 2001a; Owideru et al, 2001b; Kimbi et al, 2004; Kramvis et al, 2002). This study was aimed at characterising the full-length genome of HBV isolates at Dr George Mukhari Hospital, Pretoria, South Africa, with a view of developing a PCR-based technology for amplification and characterisation of HBV strains with different serological profiles. The technology, if successfully developed, will contribute in understanding the molecular mechanisms resulting in various HBV variants or isolates.

Methods: The study design was exploratory. Four stored serum samples collected from HBV infected patients at Dr George Mukhari hospital, Pretoria, were used to develop the molecular technology and test the hypothesis. HBV serology was previously performed targeting 5 HBV serological markers; HBsAg, anti-HBs, anti-HBc, HBeAg and anti-HBe using Elecsys version; HBV DNA quantification was done using Cobas Amplicor HBV DNA monitor assay, HBV DNA was extracted and subjected to nested PCR assay targeting HBV full length genome as two overlapping fragments: fragment A (1670 bp) and fragment B (1868 bp). The generated PCR products for both fragments were cloned into the pGEM T easy vector and 2 clones were selected from each sample. The plasmids were purified using Invisorb® Spin Plasmid Mini *Two* and the clones were recovered by PCR assay. The sample PCR products and the clone PCR products were purified and sequenced using

SpectruMedix SCE2410 genetic analysis system. HBV genotyping was performed using the NCBI web-based genotyping tool. Phylogenetic analysis was done using MEGA 4 software to confirm HBV genotypes.

Results: Serology results were as follows: All samples were HBsAg positive, Anti-HBs negative, anti-HBc positive and anti-HBe negative. Sample B1121 and sample 6 were HBeAg positive while samples B452 and 5 were HBeAg negative. A total of 12 PCR products were sequenced (4 study samples and 8 clones [2 clones each sample]). In total, 7 HBV full length genome sequences were deduced from this study, with 3 sequences belonging to genotype A, 2 to genotype C and 2 to genotype D.

3 HBV genotypes were detected from this study; genotype A, C and D with subgenotype A2, C1 and D1 respectively. Mutations were observed throughout the genome. In the pre-S/S open reading frame (ORF), the most significant findings were the detection of mutations within the “a” determinant site and major hydrophilic region (MHR). These mutations included Y161F,E164G observed in sample B1121 and B1121C1 belonging to subtype A1; 2 amino acid insertion at aa 161-162 in sample 5 belonging to subtype C1. Drug resistance associated mutations were identified in the polymerase gene, these included M204T and L217R which are associated with adefovir resistance, M204T also resulted in a change from tryptophan (W) to arginine (R) at aa 196 on the overlapping surface gene on sample B452 C1. Basal core promoter (BCP) and pre core/core mutations were detected in study isolates; specifically the BCP double mutation (1762/1764) was seen in 8 isolates which belonged to subtype C1 (5) and D1 (3) and the pre-core stop codon mutations (G1896A) in 4 isolates (2 belonging to subtype C1 and the other 2 to D1). Other changes observed included a 48 nucleotides deletion in the pre-core gene, 6 nucleotides insertion in the HBx gene of all subtype D1 isolates and a 3 nucleotides deletion in subtype C1 clone.

Conclusion: This study successfully optimised a PCR-based technology for the amplification and characterisation of HBV full length genome. 3 HBV genotypes were detected with subtypes A2, C1 and D1. However, the detected subtypes are rarely detected in South Africa. The detection of subtype A2 may confirm its Southern

African origin. Drug resistance associated mutations were observed in this study. These included the adefovir resistance mutation which the current study confirmed it is a naturally occurring mutation as it was detected in adefovir therapy naïve patient. The BCP and pre-core/core mutations were detected in genotype C and D isolates; however, their association with serological profile and clinical outcomes could not be deduced. Unique or novel mutations were seen in the study isolates, these included 48 nucleotides deletion in the pre core gene, 3 amino acids insertion in the RNase H and 8 amino acids deletion in the RT domain of polymerase gene. To our knowledge, these mutations have not been identified or reported in the literature. The detection of 6 nucleotide insertion in the HBx gene was reported for the first time in South African isolates. Further analysis is required to ascertain the biological significance of the unique mutations detected in this study.

Table of contents

Page No.

List of Abbreviations

µl	Microliter
3TC	Lamivudine
Aa	Amino acid
ADV	Adefovir dipivoxil
AIDS	Acquired immune deficiency syndrome
ALT	Alanine transaminase
Anti-HBc	Antibody to core antigen
Anti-HBe	Antibody to hepatitis B e antigen
Anti-HBs	Antibody to surface antigen
AST	Aspartate transaminase

BCP	Basal core promoter
CAESER	Canada, Australia, Europe and South African research
CAH	chronic active hepatitis
cccDNA	Covalently closed circular DNA
CD	Cluster of differentiation
CDC	Centers for Disease Control and Prevention
CPH	chronic persistent hepatitis
dATP	deoxyadenine triphosphate
dCTP	Deoxycytosine triphospate
dGTP	Deoxyguanine triphosphate
DNA	Deoxyribonucleic acid
dNTP	Dinucleotidephosphate
DR	Direct repeat
DRC	Democratic Republic of Congo
dTTP	Deoxythymine triphosphate
E. coli	Eschericia coli
EPI	Expanded Programme of Immunisation
ER	Endoplasmic reticulum
ETV	Entecavir
FHB	Fulminant hepatitis B
FTC	Emtricitabine
HAART	Highly active antiretroviral therapy
HAV	Hepatitis A virus
HBcAg	Hepatitis B core antigen
HBeAg	Hepatitis B “e” antigen
HBsAg	Hepatitis B surface antigen
HBV	Hepatitis B virus

HBx	Hepatitis B x gene
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
HDV	Hepatitis D virus
HEV	Hepatitis E virus
HIV	Human immunodeficiency virus
HLA	Human leukocyte antigen
IgG	Immunoglobulin G
IgM	Immunoglobulin M
kDa	Kilo Dalton
LCR	Ligase chain reaction
LdT	Telbivudine
LFT	Liver function test
MHC	Major histocompatability complex
MHR	Major hydrophilic region
mRNA	Messanger RNA
Nt	Nucleotide
ºC	Degree celsius
OHB	Occult hepatitis B
ORF	Open reading frame
PCR	Polymerase chain reaction
pgRNA	Pregenomic RNA
pre-S1,	Pre surface 1
pre-S2	Pre surface 2
RNA	Ribonucleic acid
RT	Reverse transcriptase
USA	United States of America

WHO	World health organisation
YMDD	Tyrosine-Methionine-Aspartate-Aspartate

Nucleotides

A	Adenine
C	Cytosine
G	Guanine
T	Thymine

Amino acids

A	Alanine
R	Arginine
N	Asparagine
D	Aspartic acid
C	Cysteine
E	Glutamic acid
H	Histidine
I	Isoleucine
L	Leicine
K	Lysine
M	Methionine
F	Phenylalanine
P	Proline
S	Serine
T	Threonine
W	Tryptophan

Y	Tyrosine
V	Valine

List of tables

List of figures