Molecular characterisation of hepatitis A virus strains from clinical and environmental sources in South Africa

by

RACHIDA SAÏD

Submitted in partial fulfillment of the requirements for the degree

Magister Scientiae

MSc (Medical Virology)

Department of Medical Virology

Faculty of Health Sciences

University of Pretoria

Pretoria

December 2013

DECLARATION

I, Rachida Saïd, declare that this work was not copied or repeated from any other studies either from national or international publications. Procedures were carried out in accordance with the ethical rules as prescribed by the Faculty of Health Science Research Ethics Committee, University of Pretoria

Signature:.....

Date: 20/06/2014

ACKNOWLEDGEMENTS

All praises and thanks are to God Almighty, without Whom this work would not have been possible. I declare that there is none worthy of worship besides the One true God and that Muhammad (peace be upon him) is His last messenger.

I would like to thank:

My supervisor, Professor MB Taylor, for giving me the opportunity of being her student, for her guidance, advice and assistance throughout the project. It has been an honour having her as a supervisor and I learned a lot

Dr M Wolfaardt, my co-supervisor, for her advice, assistance and encouragement on this project

Dr J Mans and Mr I Botha for assistance with phylogenetic analysis and protein structure

Professor R Pintó, for assistance with protein 3D structure and her interest in the project

The entire enteric group for their assistance, encouragement and friendship

Dr A Glass from Lancet Laboratories and Dr S Korsman from NHLS-Cape Town for the collection of clinical specimens

The Poliomyelitis Research Foundation, South Africa, for financial assistance with a post-graduate bursary

My husband, my father, my family at large and my friends for their full support, kindness and comfort throughout this project

Research support:

i) This work is based, in part, on the research supported by the National Research Foundation of South Africa (Grant UID 83898). *Any opinion, finding and conclusion or recommendation expressed in this material is that of the author and the NRF does not accept any liability in this regard.*

ii) This study was funded, in part, by an ongoing Water Research Commission solicited research project co-funded by Department of Agriculture, Forestry and Fisheries "An investigation into the link between water quality and microbiological safety of fruit and vegetables from the farming to the processing stages of production and marketing (WRC Project no K5/1875/4, Water Research Commission, 2009)"

Molecular characterisation of hepatitis A virus strains from clinical and environmental sources in South Africa

by

Rachida Saïd

SUPERVISOR: PROF MB TAYLOR

CO-SUPERVISOR: DR M WOLFAARDT

DEPARTMENT: MEDICAL VIROLOGY

DEGREE: MAGISTER SCIENTIAE (MEDICAL VIROLOGY)

SUMMARY

Hepatitis A virus (HAV) is the most common cause of acute viral hepatitis worldwide. Hepatitis A virus is classified in the genus *Hepatovirus* of the family *Picornaviridae*. The virion has a 7.5 kb positive-sense single-stranded RNA genome and only one serotype of the virus has been identified so far. There are six genotypes of HAV distinguished by nucleotide sequence analysis of the VP1 region. Genotypes I, II and III, each with their subdivisions A and B, infect humans. Genotypes IV, V and VI infect non-human primates. These genotypes have a distinct geographical distribution. In 1997, it was established that genotype IA and IB prevail in South Africa (SA) with IB being predominant. However, no recent data have been published on the types of HAV circulating in SA even though the virus was detected in surface water sources and on fresh produce at the point of retail. There is an effective vaccine for the prevention of hepatitis A, but recent publications have reported the possible emergence of antigenic escape mutants of HAV. The aim of the study was therefore to

establish which genotypes and possibly variants thereof of HAV are circulating the community compared to the genotypes present in water and food sources.

To identify the different strains of HAV circulating in SA, detection and characterisation of HAV strains was performed on clinical and environmental sources. A total of 117 clinical specimens and 80 water samples (irrigation, surface and wastewater) and 20 fresh produce (tomatoes, lettuce, cabbage) samples were analysed. Samples were collected in SA and neighbouring countries. Hepatitis A virus was detected in 78% of the clinical specimens and in only 46% of the water samples. None of the fresh produce analysed in the study tested positive for HAV even though some of the produce were irrigated with HAV contaminated water. Molecular characterisation of the detected strains revealed the presence of genotype IB in SA. The SA strains have a unique nucleic acid sequence, namely a R298K mutation within the VP1 region and R71S substitution within the VP1/P2B junction, which distinguishes them from strains detected in the rest of the world and even in neighbouring countries. The characterised strains were also analysed for the presence of previously isolated antigenic escape mutants and/or novel ones. One possible vaccine escape mutant, with a I179T substitution next to the neutralisation residue Y181 in the VP1 region, was identified. A variety of strains with mutation(s) at or around the neutralisation sites were also identified as well as circulation of quasispecies among clinical HAV strains. The study revealed the circulation of a potential new HAV genotype among the SA simian population. The data provided by the study stresses the need to implement a proper surveillance system of the circulation of HAV strains in SA.

PRESENTATIONS AND PUBLICATIONS

Publications

Saïd R, Wolfaardt M, Taylor MB. Molecular characterisation of hepatitis A virus strains from water sources in South Africa. Water Science and Technology 2014;69:923-933.

National Presentations

Saïd R, Wolfaardt M, Taylor MB. Molecular characterisation of hepatitis A virus strains from water sources in South Africa [Presentation]. Faculty Day, Faculty of Health Sciences, University of Pretoria 28-29 August 2012: HW Snyman Building North, Pretoria

Saïd R, Wolfaardt M, Taylor MB. Molecular characterisation of hepatitis A strains from water sources in South Africa [Presentation]. Fountain of Knowledge UP Water Institute Members' Forum Conference 9 November 2012 Law Building, Hatfield Campus, University of Pretoria

Saïd R, Wolfaardt M, Taylor MB. Molecular characterisation of hepatitis A strains from water sources in South Africa [Presentation]. Third regional Conference of the Southern African Young Water Professionals 16-18 July 2013: Konservatorium, University of Stellenbosch, Stellenbosch, South Africa.

TABLE OF CONTENTS

		Page
DECLARATION		i
ACKNOWLEDGEMENTS		ii
SUMMARY		iv
PUBLICATIONS AND PRESENTATIONS		vi
TABLE OF CONTENTS		vii
ABBREVIATIONS AND SYMBOLS		xii
LIST OF FIGURES		xv
LIST OF TABLES		xvii
CHAPTER 1: LITERATURE REVIEW		
1.1	**GENERAL INTRODUCTION**	1
1.2	**VIROLOGY**	2
	1.2.1 Morphology	2
	1.2.2 Viral proteins	3
	1.2.3 Nucleic acid and genome organisation	3
	1.2.3.1 Noncoding regions	4
	1.2.3.2 The open reading frame	4
	1.2.4 Physicochemical and physical characteristics	5
	1.2.5 Classification	5
1.3	**CLINICAL ASPECT OF HEPATITIS A**	10
	1.3.1 Pathogenesis	10
	1.3.2 Clinical features	12
	1.3.2.1 Symptomatic icteric hepatitis A	12
	1.3.2.2 Immune response	14
	1.3.2.3 Biochemical parameters	15
	1.3.2.4 Complications of hepatitis A	15
	1.3.2.5 Extrahepatic manifestations	15

1.3.3 Laboratory diagnosis of hepatitis A 16
1.4 EPIDEMIOLOGY 17
1.4.1 Global epidemiology 17
1.4.2 Role of food and water in the transmission of hepatitis A virus 19
1.4.2.1 Water sources 19
1.4.2.2 Food sources 20
1.5 PREVENTION AND CONTROL OF HEPATITIS A 22
1.6 DETECTION AND CHARACTERISATION OF HEPATITIS A VIRUS IN WATER AND FOOD SOURCES 24
1.6.1 Sampling and sample preparation 24
1.6.1.1 Water matrices 24
1.6.1.2 Food matrices 26
1.6.2 Viral detection 27
1.6.3 Viral characterisation 28
1.7 RATIONALE FOR STUDY 29
1.8 HYPOTHESIS 30
1.9 AIM 30
1.10 SPECIFIC OBJECTIVES 30

CHAPTER 2: DETECTION AND GENOTYPING OF HEPATITIS A VIRUS STRAINS IN CLINICAL AND ENVIRONMENTAL SOURCES
2.1 INTRODUCTION 31
2.2 MATERIALS AND METHODS 32
2.2.1 Ethics statement 32
2.2.2 Clinical specimens 33
2.2.2.1 Specimens collection 33
2.2.2.2 RNA extraction 33
2.2.3 Water samples 33
2.2.3.1 Sample collection 33
2.2.3.2 Microbial indicator analysis 35
2.2.3.3 Viral recovery and concentration 35

2.2.3.3.1 Viral recovery from 10 L irrigation water samples 35
2.2.3.3.2 Viral recovery wastewater samples 37
2.2.3.4 Total nucleic acid extraction 37
2.2.4 Fresh produce samples 37
2.2.4.1 Samples collection 37
2.2.4.2 Viral recovery and concentration 38
2.2.4.3 Total nucleic acid extraction 39
2.2.5 Viral detection 39
2.2.5.1 Mengovirus detection 39
2.2.5.2 Hepatitis A virus detection 39
2.2.6 Hepatitis A virus characterisation 40
2.2.6.1 Complementary DNA (cDNA) synthesis 40
2.2.6.2 Amplification of the VP1 region 41
2.2.6.3 Amplification of the VP1/P2B junction 41
2.2.6.4 Sequencing and cloning 42
2.2.6.5 Nucleotide sequence analyses 43
2.2.6.6 Phylogenetic analysis 45
2.3 RESULTS 45
2.3.1 Microbial indicator analysis 45
2.3.2 Hepatitis A virus detection 45
2.3.2.1 Viral detection in clinical specimens 45
2.3.2.2 Viral detection in water samples 45
2.3.2.3 Viral detection on fresh produce 46
2.3.3 Hepatitis A virus characterisation 48
2.3.3.1 Viral characterisation in clinical specimens 48
2.3.3.2 Characterisation of viral strains from water samples 48
2.3.3.3 Genetic relatedness of hepatitis A virus from clinical and water sources 49
2.4 DISCUSSION 54

CHAPTER 3: GENETIC DIVERSITY OF HEPATITIS A VIRUS IN SOUTH AFRICA

3.1 INTRODUCTION 63

3.2 MATERIALS AND METHODS 64

3.2.1 Nucleotide sequence analyses 64

3.2.2 Detection of novel or predicted mutations 64

3.2.3 Phylogenetic analysis 65

3.2.4 Cloning of clinical strains and quasispecies analysis 65

3.3 RESULTS 66

3.3.1 Pairwise analysis 66

3.3.2 Phylogenetic analysis 68

3.3.2.1 Phylogenetic analysis of HAV strains from clinical specimens 68

3.3.2.2 Phylogenetic analysis of HAV strains from water samples 71

3.3.3 Quasispecies analysis 73

3.3.4 Amino acid sequence analysis 74

3.3.5 Search for antigenic escape variants 77

3.4 DISCUSSION 78

CHAPTER 4: GENERAL DISCUSSION 83

CHAPTER 5: REFERENCES 87

APPENDIX A 103

Reagents and buffers

APPENDIX B 105

B.1: ABSTRACT: Saïd R, Wolfaardt M, Taylor MB. Molecular characterisation of hepatitis A virus strains from water sources in South Africa [Presentation].

B.2: ABSTRACT: Saïd R, Wolfaardt M, Taylor MB. Molecular characterisation of hepatitis A strains from water sources in South Africa [Presentation].

B.3: ABSTRACT: Saïd R, Wolfaardt M, Taylor MB. Molecular characterisation of hepatitis A strains from water sources in South Africa [Presentation].

APPENDIX C 108

Publication: Saïd R, Wolfaardt M, Taylor MB. Molecular characterisation of hepatitis A virus strains from water sources in South Africa. Water Science and Technology

APPENDIX D 120

Ethical approval

ABBREVIATIONS AND SYMBOLS

5'NCR	5' noncoding region
AC-PCR	Antigen-capture PCR
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
BLAST-n	Basic Local Alignment Search Tool – nucleotide
bp	Base pairs
cDNA	Complementary DNA
CPE	Cytopathic effect
EIA	Enzyme immunoassays
g	Gram
GBEB	Glycine beef-extract buffer
GP	Gauteng province
h	Hour
HAV	Hepatitis A virus
HAVCR1	HAV cell receptor 1
HAVCR2	HAV cell receptor 2
HCl	Hydrochloric acid
HEV	Hepatitis E virus
HIV/AIDS	Human immunodeficiency virus/acquired immunodeficiency syndrome
IAC	Internal amplification control
ICC-PCR	Integrated cell culture–PCR
ICH-GCP	International Conference on Harmonisation–Good Clinical Practice
IC-PCR or IMC-PCR	Immunomagnetic capture RT-PCR
IEM	Immune electron microscopy
IgA	Immunoglobulin A
IgG	Immunoglobulin G

IgM	Immunoglobulin M
IRES	Internal ribosome entry site element
kb	Kilo bases
kDa	Kilo Dalton
KZN	KwaZulu-Natal
L	Litre
LAMP	Loop-mediated isothermal amplification
LP	Limpopo province
M	Molar
min	Minute
mL	Millilitre
MP	Mpumalanga
MPN	Most probable number
NaCl	Sodium Chloride
NaOH	Sodium hydroxide
NASBA	Nucleic acid sequence based amplification
nm	Nanometre
NW	North West
ORF	Open reading frame
PBS	Phosphate buffered saline
PEG	Polyethylene glycol
ppm	Parts per millions
qRT-PCR	Quantitative RT-PCR
RIA	Radioimmunoassay
RIFA	Radio-immunofocus assay
RT-PCR	Reverse transcription-polymerase chain reaction
s	Second
SA	South Africa
ssRNA	single-stranded RNA
$TCID_{50}$	Fifty per cent tissue culture infectious dose
TGBE	Tris-glycine beef extract
U	Units

USA	United States of America
USSR	Union of Soviet Socialist Republics
VP1	Viral protein 1
VP2	Viral protein 2
VP3	Viral protein 3
VP4	Viral protein 4
VPg	Virion protein, genome linked
WC	Western Cape
WHO	World Health Organization
µL	Microlitre

LIST OF FIGURES

FIGURE		Page
Figure 1.1	Electron micrograph of hepatitis A virions	3
Figure 1.2	Genome organisation of hepatitis A virus, illustrating proteins encoded by the different genomic regions	4
Figure 1.3	Hepatitis A virus genome organisation depicting regions used for typing	7
Figure 1.4	Viral titre, immunological response and biochemical signs during HAV infection	13
Figure 2.1	Distribution of irrigation water and wastewater sampling sites among six provinces of South Africa.	34
Figure 2.2	Phylogenetic analysis of hepatitis A virus strains characterised from clinical specimens based on the sequence of the VP1 region (900 bp).	50
Figure 2.3	Phylogenetic analysis of hepatitis A virus strains characterised from clinical specimens based on the sequence of the VP1/P2B junction (~350 bp)	51
Figure 2.4	Phylogenetic analysis of hepatitis A virus strains characterised from water samples based on the sequence of the VP1 region (900 bp)	52
Figure 2.5	Phylogenetic analysis of hepatitis A virus strains characterised from water samples based on the sequence of the VP1/P2B junction (~350 bp).	53
Figure 2.6	Phylogenetic analysis of hepatitis A virus strains characterised from clinical and water sources based on the sequence of the VP1 region (900 bp)	55
Figure 2.6a	Phylogenetic tree highlighting strains of ‘SA major cluster’ as depicted by Figure 2.6.	56
Figure 2.7	Phylogenetic analysis of hepatitis A virus strains characterised from clinical and water sources based on the sequence of the VP1/P2B junction (~350 bp).	57
Figure 2.7a	Phylogenetic tree highlighting strains of ‘SA major cluster’ as depicted by Figure 2.7	58

Figure 3.1	Phylogenetic tree highlighting the strains GP_PT131S, GP_PT112S and GP_PT112S2 , WC_PT44S and WC_PT44S2 of 'SA and non-SA cluster' as depicted by Figure 2.2	69
Figure 3.2	Phylogenetic tree highlighting the strains GP_PT134S, WC_PT47S, WC_PT47S2, WC_PT49S and WC_PT49S2 of 'SA major cluster' as depicted by Figure 3.1	70
Figure 3.3	Phylogenetic tree highlighting the strains GP_PT134S, GP_PT112S, GP_PT112S2, WC_PT44S, WC_PT44S2 of 'SA minor clusters A and B' as depicted by Figure 2.3.	71
Figure 3.4	Phylogenetic tree highlighting the strains GP_PT131S, WC_PT47S, WC_PT47S2, WC_PT49S, WC_PT49S2 of 'SA major cluster' as depicted on Figure 3.3.	72
Figure 3.5	Quasispecies analysis of hepatitis A virus strains from five clinical specimens	73
Figure 3.6	3D model of VP1 structure highlighting neutralisation site at residue 181 (yellow) and recorded I179T mutation (red)	78

LIST OF TABLES

TABLE		Page
Table 2.1	Sequences of primers and probes used for mengovirus detection	39
Table 2.2	Sequences of primers and probes used for HAV detection	40
Table 2.3	Sequences of internal primers used to amplify the VP1 gene	41
Table 2.4	Sequences of primers used to amplify the VP1/P2B junction	42
Table 2.5	Hepatitis A virus strains used in phylogenetic analysis	44
Table 2.6	Summary of results of HAV detection in clinical specimens	46
Table 2.7	Summary of information on sampling site, results of microbial indicator and viral analysis of irrigation water samples	47
Table 2.8	Summary of results of HAV detection in wastewater samples by province	48
Table 3.1	Pairwise analysis (percentage of identities and divergence) results of HAV strains characterised from clinical specimens and water samples	66
Table 3.2	Genetic distances (percentage of identities and divergence) between type strains of HAV genotypes I to VI and KZN_20130530	67
Table 3.3	Amino acid substitutions observed at or near the neutralisation sites in the VP1 sequences characterised in this study	75
Table 3.4	Amino acid substitutions observed in the VP1/P2B sequences	76
Table 3.5	Similar amino acid substitutions recorded for the KZN_20130530 strain when compared to other hepatitis A virus genotypes	77

CHAPTER 1

LITERATURE REVIEW

1.1 GENERAL INTRODUCTION

Jaundice epidemics have plagued mankind since the ancient times of the Greeks, the Romans and the Chinese (Beard and Lemon, 1999; Nainan *et al*., 2006; Pintó and Saiz, 2007). The aetiological agent responsible for these epidemics remained elusive. At the beginning of the 20th century, a form of hepatitis was associated with certain forms of infectious jaundice epidemics (Dotzauer, 2008; Pintó *et al*., 2010). Subsequently, based on the route of transmission, two separate entities of hepatitis were identified, i.e. "infectious" and "serum" hepatitis (Pintó *et al*., 2010). In the second half of the 20th century, a specific group of viruses was identified as the aetiological agents responsible for both entities of hepatitis (Beard and Lemon, 1999; Nainan *et al*., 2006; Pintó *et al*., 2010). These viruses primarily infect hepatocytes, resulting in either acute or chronic liver inflammation, and as a result were given the name "hepatitis viruses" (Nainan *et al*., 2006; Tang, 2009). At present, there are five major hepatitis viruses that cause similar clinical manifestations, but they differ in their morphology, genomic organisation, taxonomy and mode of replication (Tang, 2009; Kumar *et al*., 2010; Pintó *et al*., 2010). These viruses can be grouped according to their predominant mode of transmission, namely enteric ("infectious" hepatitis) or parenteral ("serum" hepatitis) (Collier and Oxford, 2006). Parenterally transmitted hepatitis viruses are hepatitis B virus, hepatitis C virus and hepatitis D virus. They cause acute hepatitis with a high probability of developing chronic infection (Kumar *et al*., 2010) and can be spread via blood and blood products, sexual contact or vertically (mother to child) (Collier and Oxford, 2006; Tang, 2009). Enterically transmitted hepatitis viruses are hepatitis A virus (HAV) and hepatitis E virus (HEV). They are predominantly

transmitted via the faecal-oral route (Collier and Oxford, 2006; Pintó and Saiz, 2007; Tang, 2009; Pintó *et al*., 2010), either directly from person-to-person or indirectly from contaminated food and water sources (Nainan *et al*., 2006; Pintó *et al*., 2010). Infections with HAV or HEV are reported either as epidemic or sporadic cases (Kumar *et al.*, 2010).

Hepatitis A virus is the aetiological agent of hepatitis A (Feinstone *et al*., 1973; Beard and Lemon, 1999; Dotzauer, 2008). The virus was first visualised in 1973 by immune electron microscopy (IEM) (Feinstone *et al*., 1973; Beard and Lemon, 1999). Ten years later, the RNA genome of HAV was reverse transcribed and molecularly cloned (Ticehurst *et al*, 1983), and in 1987 the first full length genome of HAV was published (Beard and Lemon, 1999). Sequence analysis of the first HAV genome revealed that the genomic organisation of the virus was similar to that of picornaviruses. Thereafter, progress in cell culture and development of safe methods to inactivate viruses led to the production of a safe and very effective vaccine for the prevention of HAV infection (Martin and Lemon, 2006; Dotzauer, 2008).

1.2 VIROLOGY

1.2.1 Morphology

Hepatitis A virus is a small non-enveloped RNA virus that has an icosahedral capsid of approximately 27 nm in diameter (Figure 1.1) (Costa-Mattioli *et al*., 2003; Dotzauer, 2008; Spradling *et al*., 2009; Knowles *et al*., 2012). The capsid consists of 60 protomers (80-97 kDa each) that assemble via pentameric intermediates. A single protomer is made of three surface proteins 1B, 1C and 1D commonly known as viral protein 2, 3 and 1 (VP2, VP3 and VP1) (Knowles *et al*., 2012). The N-termini and β-barrel domain of the three major capsid proteins (VP1 to VP3) are essential for the correct conformation and stability of the viral capsid (Racaniello, 2007; Knowles *et al*., 2012).

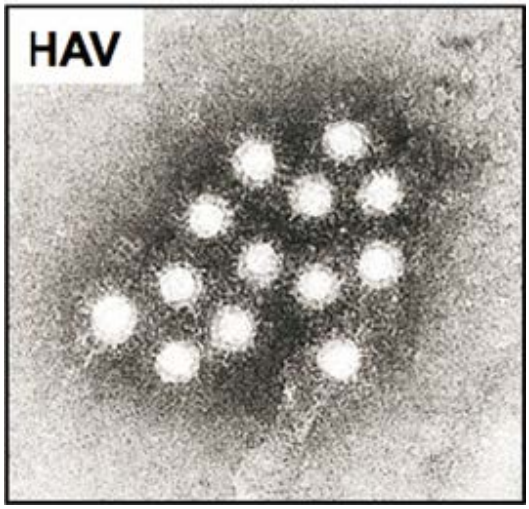

Figure 1.1: Electron micrograph of hepatitis A virions (Kumar *et al*., 2010)

1.2.2 Viral proteins

Infectious particles of HAV have four proteins: VP1, VP2, VP3 and VPg (virion protein, genome linked) (Cuthbert, 2001; Dotzauer, 2008; Knowles *et al*., 2012). The first three are involved in the synthesis of the capsid while the last one is a small protein (2.5 kDa) that functions in RNA synthesis initiation (Cuthbert, 2001).

1.2.3 Nucleic acid and genome organisation

Hepatitis A viral particles each contain a single molecule of positive-sense, single-stranded RNA (ssRNA) of approximately 7.5 Kb in size (Costa-Mattioli *et al*., 2003; Dotzauer, 2008; Spradling *et al*., 2009; Knowles *et al*., 2012). The RNA molecule consists of a 5' noncoding region (5'NCR) covalently linked to VPg, a single open reading frame (ORF) and a 3' noncoding region (3'NCR) that ends with a poly(A) tail (Figure 1.2) (Costafreda *et al*., 2006; Nainan *et al*., 2006; Kumar *et al*., 2010; Knowles *et al*., 2012).

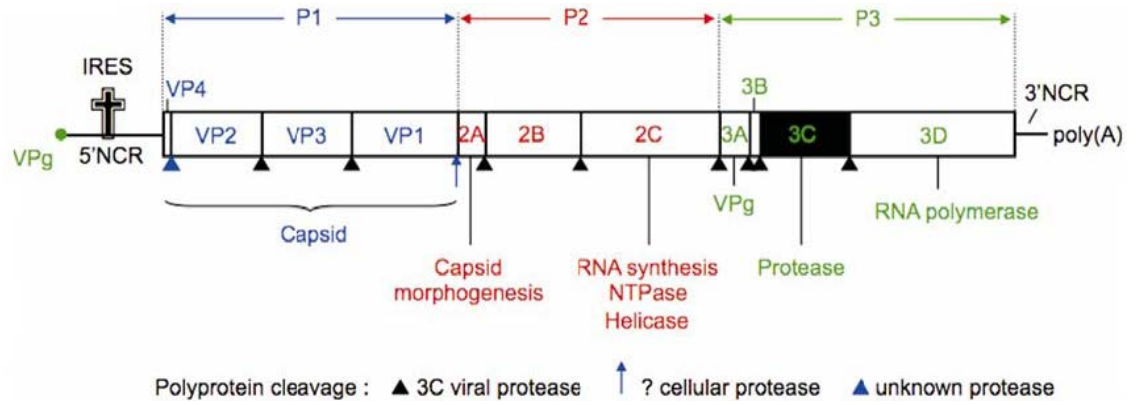

Figure 1.2: Genome organisation of hepatitis A virus, illustrating proteins encoded by the different genomic regions (Kumar *et al*., 2010).

1.2.3.1 Noncoding regions

The 5'NCR has a 5'-terminal domain involved in replication as well as an internal ribosome entry site element (IRES) that functions in translation initiation (Nainan *et al*., 2006; Knowles *et al*., 2012; Pan *et al*., 2012). Because of the important functions it fulfils, the 5'NCR is the most conserved region of the HAV genome. To detect HAV in clinical and environmental specimens, a real-time reverse transcription-polymerase chain reaction (RT-PCR) assay was developed with primers and probes that target the 5'NCR (Costafreda *et al*., 2006). Conversely, the 3'NCR is highly variable among strains, with up to 20% nucleotide divergence. Despite the high degree of divergence, the 3'NCR might down regulate HAV RNA synthesis (Hollinger and Emerson, 2007).

1.2.3.2 The open reading frame

The ORF of the HAV genome has three distinct regions: P1, P2 and P3 (Figure 1.2). Translation of the ORF gives a single polyprotein which is subsequently cleaved into structural and non-structural proteins by a unique protease called $3C^{pro}$ (Nainan *et al*., 2006; Kumar *et al*., 2010; Knowles *et al.,* 2012). The non-structural proteins are encoded from region P2 and P3 (Figure 1.2) and are involved in the replication cycle of the virus (Nainan *et al*., 2006; Kumar *et al*., 2010). The structural proteins encoded from region P1 are VP1 (33 kDa), VP2 (27 kDa), VP3 (29 kDa) and VP4 (17 amino acids) (Kumar *et al*., 2010). Even though VP4 is absent from the mature virion, it could play an important role

during HAV morphogenesis (Cristina and Costa-Mattioli, 2007; Hollinger and Emerson, 2007; Dotzauer, 2008; Knowles *et al.,* 2012).

1.2.4 Physicochemical and physical characteristics

Hepatitis A viruses have a buoyant density of ~1.33 g cm^{-3} and sedimentation coefficient of ~70S (empty capsid) or 160S (Dotzauer, 2008, Spradling *et al*., 2009). The structure and conformation of the capsid allows the virus to survive harsh conditions such as acidic pH and elevated temperatures (Hollinger and Emerson, 2007; Knowles *et al*., 2012). Infectivity is retained at pH 1.0 for 2 hours (h) at room temperature and only partially reduced after 10 to 12 h incubation at 60ºC under neutral conditions (Hollinger and Emerson, 2007; Dotzauer, 2008; Spradling *et al*., 2009). Hepatitis A viral particles can remain infectious for days or months in fresh water, seawater, wastewater, soils, marine sediment, cream-filled cookies and live oysters (Hollinger and Emerson, 2007). The absence of an envelope allows the virus to resist organic solvents like ether and chloroform, as well as drying, detergents and storage at 25ºC or at -20ºC. However, inactivation of infectious HAV particles is possible by chlorine (10 to 15 ppm residual chlorine concentration after 30 minutes (min); free residual chlorine concentration of 2.0 to 2.5 mg/L for 15 min) or chlorine containing compounds such as sodium hypochlorite (3 to 10 mg/L at 20ºC for 5 to 15 min), by autoclaving (121ºC for 20 min), by ultraviolet radiation (1.1 W at a depth of 0.9 cm for 1 min, or 197 μW/cm^2 for 4 min), by formalin (3% for 5 min at 25ºC, or 8% for 1 min at 25ºC) and by iodine (3 mg/L for 5 min) (Hollinger and Emerson, 2007; Spradling *et al*., 2009).

1.2.5 Classification

Hepatitis A virus is the sole member of the genus *Hepatovirus* within the family *Picornaviridae* (Costa-Mattioli *et al.*, 2002; Racaniello, 2007; Spradling *et al*., 2009; Knowles *et al*., 2012). The term “picornavirus” refers to viruses that are small in size (in the range of picometre [10^{-12} m]) and that have a RNA viral genome (Racaniello, 2007; Knowles *et al.*, 2012). Viruses that belong to the

Picornaviridae family infect vertebrates and include important human and animal pathogens such as poliovirus and foot-and-mouth-disease virus (Racaniello, 2007; Tapparel *et al.*, 2013; Knowles *et al*., 2012). In addition to the genus *Hepatovirus*, there are eleven more genera within the family *Picornaviridae*: *Enterovirus, Cardiovirus, Aphthovirus, Parechovirus, Erbovirus, Kobuvirus, Teschovirus, Sapelovirus, Senecavirus, Tremovirus, Avihepatovirus* (Racaniello, 2007; Knowles *et al*., 2012). The genus *Avihepatovirus* is the last genus that was added to the family and it groups viruses that cause hepatitis in young ducklings (Knowles *et al*., 2012; Li *et al*., 2013). The genus *Hepatovirus* share similar genomic organisation and protein expression patterns with other genera in the family *Picornaviridae*. However, a number of unique characteristics distinguishes HAV from other picornaviruses; i.e. i) the fourth structural protein of HAV, 1A (VP4), is very small, does not appear to be myristylated at its N-terminus and is absent from the mature virion (Knowles *et al*., 2012); ii) HAV has one serotype and one immunodominant antigenic site (Hollinger and Emerson, 2007); iii) the mature HAV virion may not have the canyon, a structure that other picornaviruses use to attach to cellular receptors (Dotzauer, 2008).

A strain of HAV can be defined as an isolate that is genetically distinguishable from other HAV isolates (Hollinger and Emerson, 2007). Even though growth characteristics could be used to differentiate HAV strains, classification based on nucleotide sequence analyses of specific genomic regions (Figure 1.3) are most reliable (Nainan *et al*., 2006; Hollinger and Emerson, 2007). A group of HAV strains with ≥85% or >92.5% nucleotide sequence identity, over the VP1/P2A region, are classified within the same genotype or subgenotype, respectively (Robertson *et al*., 1992; Taylor, 1997; Vaughan *et al*., 2013a).

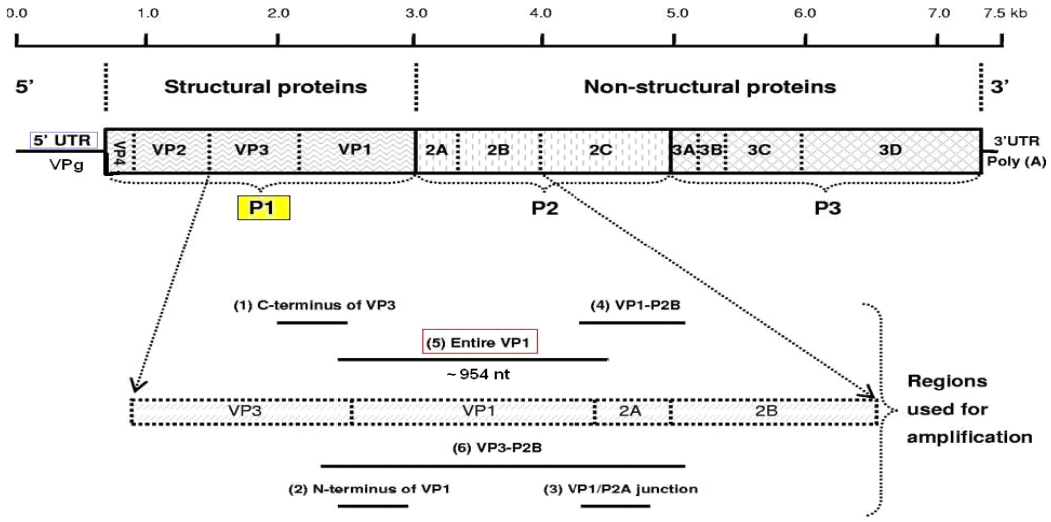

Figure 1.3: Hepatitis A virus genome organisation depicting regions used for typing (adapted from Nainan *et al*., 2006)

Naturally occurring HAV can infect humans or non-human primates. Based on the nucleic acid sequence analysis of the entire VP1 region (900 nucleotides) six genotypes, with their subdivisions (A and B), have been defined: I, II and III grouping HAVs associated with humans, and IV, V and VI grouping HAV associated with non-human primates (Costa-Mattioli *et al*., 2002). The VP1 region was chosen to genotype HAV strains because it contributes to the immunodominant antigenic site and is more discriminatory when compared to the VP1/P2A junction that was previously used for genotyping (Figure 1.3) (Costa-Mattioli *et al*., 2002). The regions VP1/P2B or N-terminus of VP1 can also be used for genotyping (Figure 1.3). The only difference between the new classification of HAV strains and the previous one based on the VP1/P2A junction is the type strain of genotype VII is now classified as subgenotype IIB (Costa-Mattioli *et al*., 2002, 2003; Lu *et al*., 2004). In addition, sequence diversity over the VP1 region differs slightly when compared to VP1/P2A junction. Consequently, for characterisation based on the entire VP1 region strains belonging to the same subgenotype should have >88.9% nucleotide sequence identity and >94.4% amino acid identity (Costa-Mattioli *et al*., 2002). Genetic similarities between genotypes should range from 76.5% to 89.4% at nucleotide level and 89.5% to 99.3% at amino acid level.

Hepatitis A virus strains have a distinct geographical distribution that impact greatly on the epidemiology of the virus worldwide (Robertson *et al*., 1992; Nainan *et al*., 2006; Vaughan *et al*., 2013a). One or more genotype or subgenotype may occur at the same geographical location. In the case where more than one subgenotype prevails in a specific region, unequal distribution of the different HAV types among environments could be observed (Vaughan *et al*., 2013a). During the investigation of a hepatitis A outbreak in Rio de Janeiro, Brazil, both subgenotype IA and IB where identified, with IA exclusively detected in serum specimens while IB was present in water samples and saliva of some patients (Amado *et al*., 2011). This, together with other reports (Villar *et al*., 2007; Prado *et al*., 2012; Vaughan *et al*., 2013a) indicates the need to sample different environments when assessing HAV strain predominance in a specific region.

Of the three genotypes that cause hepatitis A in humans, genotype II is the least represented in the world (Robertson *et al.*, 1992; Nainan *et al*., 2006, Endo *et al*., 2007; Vaughan *et al*., 2013a). Hepatitis A virus strains belonging to genotype II has been isolated in the Netherlands, France and Sierra Leone (Robertson *et al.*, 1992; Nainan *et al.*, 2006). Subgenotype IIA is endemic to Benin (Desbois *et al*., 2010). However, Benin may be the only West African country harbouring HAV subgenotype IIA, since only subgenotype IA has been detected in Nigeria and Cameroon (Forbi *et al*., 2012, 2013; Vaughan *et al*., 2013a).

Genotype I is commonly found worldwide, with subgenotype IA more prevalent than IB (Robertson *et al*., 1992; Nainan *et al.,* 2006). Subgenotype IA predominates in Europe, Asia, North and South America (Costa-Mattioli *et al*., 2003; Nainan *et al*., 2006; Sulbaran *et al*., 2010; Vaughan *et al*., 2013a). Data on the types of HAV circulating in Africa are scarce. Available reports indicate circulation of genotype I on the African continent, with IA detected in Nigeria (Forbi *et al*., 2013), Cameroon (Forbi *et al*., 2012) and Tunisia (Gharbi-Khelifi *et al*., 2012) and IB prevailing in South Africa (SA) (Robertson *et al*., 1992; Kedda *et al*., 1995; Taylor, 1997). With the exception of Tunisia, subgenotype IB

prevails in the Mediterranean region (Gharbi-Khelifi *et al*., 2006, 2012) and the Middle East (Normann *et al*., 2008; Nejati *et al*., 2012). A putative subgenotype IC has been proposed, but has not been confirmed. Subgenotype IC groups strains detected during the period 2005-2009 in Catalonia, Spain (Pérez-Sautu *et al*., 2011b).

The third HAV genotype that includes strains infecting humans is genotype III. This genotype may also include simian HAV strains, as the type strain of subgenotype IIIA was recovered from a feral owl monkey in Panama (Robertson *et al*., 1992; Nainan *et al*., 2006; Endo *et al*., 2007). Since then, subsequent isolates of IIIA were all of human origin and have been detected in numerous countries of Asia and Europe, Madagascar and the United States of America (USA) (Robertson *et al*., 1992; Nainan *et al.,* 2006; Endo *et al*., 2007). Strains of subgenotype IIIB are also of human origin and have been associated with cases of hepatitis A in Japan and Denmark (Endo *et al*., 2007). The three genotypes that group simian strains have only been isolated from non-human primates found in the Philippines (genotype IV), Kenya (genotype V) and Indonesia (genotype VI) (Nainan *et al*., 2006; Cristina and Costa-Mattioli, 2007).

Hepatitis A virus is antigenically stable with a single serotype and two biotypes (Nainan *et al*., 2006; Knowles *et al*., 2012; Pintó *et al*., 2012). For this reason, infection with HAV confers lifelong immunity. As for people who have not being previously exposed, there are several effective vaccines that have been developed after the successful propagation of HAV in cell culture (Provost and Hilleman, 1979; Shouval, 2011). The presence of a single serotype is one of the main reasons why HAV was removed from the genus *Enterovirus* and classified in a genus on its own (Hollinger and Emerson, 2007). Previous reports suggested that a total of 15% of the HAV's capsid surface residues are encoded by rare codons (Pintó *et al*., 2012; Vaughan *et al*., 2013a). These rare codons are exposed on the surface of the capsid and strategically located near or at the epitope regions. The immunodominant site and the glycophorin A binding site are two of the three main epitopes that define the antigenic structure of the HAV capsid (Pintó *et al*., 2012; Vaughan *et al*., 2013a). The

presence of a third epitope has been hypothesised but not yet confirmed (Pintó *et al*., 2012). Substitution of the rare codons is negatively selected because of their location on the capsid as well as strict maintenance of structure integrity (Aragonès *et al*., 2008, 2010; Pintó *et al*., 2012; Vaughan *et al*., 2013a). Negative selection is also observed under specific immune pressure (Pérez-Sautu *et al*., 2011a) which might explain the presence of a single serotype and the limited number of naturally occurring antigenic variants. Antigenic escape mutants have been isolated from partially immunised patients. Sequence analysis of the genomic regions of these mutants showed the following amino acid substitutions: V72I in VP3, and V166G, V171A, Y181S, R189T, A280V and A280E in VP1 (Pérez-Sautu *et al*., 2011a; Pintó *et al*., 2012). These residues are located on the immunodominant site. Hepatitis A virus antigenic variants that have a mutation at the immunodominant site or glycophorin A binding site (G217D, or K221E, or K221M in VP1) are less fit than wild-type viruses (Costafreda *et al*., 2012; Pintó *et al*., 2012). Probably because a conformational change at the glycophorin A binding site might make the virus more susceptible to blood clearance mechanisms (Pintó *et al*., 2012).

1.3 CLINICAL ASPECT OF HEPATITIS A

1.3.1 Pathogenesis

The exact sequence of events that occurs between HAV infection and manifestation of clinical disease is unknown. However, the genome variation and lineage of HAV, together with host factors such as age, sex, race and genes of the major histocompatibility locus, may play an important role in clinical manifestation of hepatitis A (Vaughan *et al*., 2013a). Studies in simians have suggested that the virus follows an enterohepatic cycle model of replication inside the host (Cuthbert, 2001; Pintó *et al*., 2012). According to the model, after ingestion of contaminated materials, HAV enters the bloodstream through the stomach or the intestine in order to gain access to the liver. The virus replicates in the hepatocytes and is released via the biliary canalicula into the intestine. Once in the intestine, either the virus is shed in faeces or it re-

enters the bloodstream and repeats the cycle. During experimental studies, HAV was not only detected in the blood and faeces, but in the saliva, suggesting a possible early replication in the salivary gland (Hollinger and Emerson, 2007). This model of replication seems plausible because the capsid of HAV can survive the acidic pH of the stomach, the action of intestinal proteases and detergent mechanisms (Pintó *et al*., 2012). Many questions still remain concerning the infectious cycle of HAV. The exact mechanism by which the virus enters the bloodstream is still unknown as well as factors that determine tissue tropism (Hollinger and Emerson, 2007; Pintó *et al*., 2012).

Hepatitis A virus primarily targets hepatocytes of primates, but may also infect crypt cells of the small intestine and Kupffer cells of the liver (Hollinger and Emerson, 2007; Knowles *et al*., 2012; Pintó *et al*., 2012). The replication cycle of HAV within the hepatocyte is similar to that of other picornaviruses (Racaniello, 2007; Pintó *et al*., 2012). The infectious viral particle interacts with a receptor at the surface of a liver cell, possibly the HAV cell receptor 1 (HAVCR1) before internalisation and uncoating of the viral genome (Spradling *et al*., 2009; Pintó *et al*., 2012). Two additional receptors, HAVCR2 and the asialoglycoprotein receptor may also play a role during HAV cell entry (Vaughan *et al*., 2013a). Inside the cytoplasm, interaction between the ribosome and the IRES within the 5'NCR of the viral RNA initiate a cap-independent translation of the RNA genome. The synthesised polyprotein is co- and post-translationally processed by the viral protease to produce the RNA-dependent polymerase and other proteins necessary to the synthesis of a negative-strand copy of the genome of HAV. The negative-strand copy of the genome will be used as a template for the synthesis of positive-strand RNA molecules that will either be translated into more proteins or used for more RNA synthesis. Virions are formed once capsid particles formed by the assembly of structural proteins enclose a copy of the RNA molecule. Viral progeny is secreted into the bile, and then transported to the small intestine. Hepatitis A virus is not cytopathic and any liver damage observed is attributed to the host immune response rather than to the replication of the virus (Dotzauer, 2008; Kumar *et al*., 2010). A novel cell release mechanism known as “membrane hijacking” has been

proposed for HAV (Feng *et al*., 2013; Vaughan *et al*., 2013a). Viral progeny may be released from the host cell in an enveloped form. Thus, two populations of HAV may be present in an infected individual: enveloped and non-enveloped virions. The presence of the first one might explain the ability of HAV to escape the immune system and replicate to high titres prior to the clinical phase of hepatitis A.

1.3.2 Clinical features

Clinically, hepatitis A is defined as an acute, self-limiting infection of the liver (Spradling *et al*., 2009; Pintó *et al*., 2010). The clinical presentation of the infection can be classified in one of the three following categories: inapparent asymptomatic or subclinical infections, and symptomatic anicteric or icteric infections (Hollinger and Emerson, 2007; Dotzauer, 2008; Spradling *et al*., 2009). In asymptomatic and subclinical hepatitis A, infection can only be differentiated at the biochemical level (Spradling *et al*., 2009). While in symptomatic hepatitis A, the icteric and anicteric infections can also be differentiated on the basis of development or not of jaundice respectively (Spradling *et al*., 2009). The severity of symptoms depends on the age of the individual at the time of exposure and the presence or not of an underlying chronic liver disease (Collier and Oxford, 2006; Nainan *et al*., 2006; Pintó *et al*., 2010). In children under the age of 5, infection with HAV is asymptomatic in 70% of cases (Spradling *et al*., 2009). While approximately 70% of older children and adults infected with HAV develop a symptomatic icteric infection (Dotzauer, 2008; Spradling *et al*., 2009).

1.3.2.1 Symptomatic icteric hepatitis A

There are four phases of symptomatic icteric hepatitis A, i.e. the incubation period, prodromal or preicteric phase, icteric phase and the convalescent period (Hollinger and Emerson, 2007). Hepatitis A virus has an incubation period of 15 to 50 days with a mean of approximately 30 days (Beard and Lemon, 1999; Cuthbert, 2001; Hollinger and Emerson, 2007; Dotzauer, 2008; Tang, 2009; Pintó *et al*., 2012). During the incubation period, the infected individual is

asymptomatic despite active virus replication in the liver (Hollinger and Emerson, 2007). The viruses are found in faeces, sera and to a lesser extent in saliva (Hollinger and Emerson, 2007; Pintó *et al*., 2012). Faecal shedding of HAV, reaches peaks at approximately 10^{11} genome copies/g of faeces just prior to the onset of symptoms (Figure 1.4) at which time an infected individual is highly infectious (Pintó *et al*., 2010, 2012).

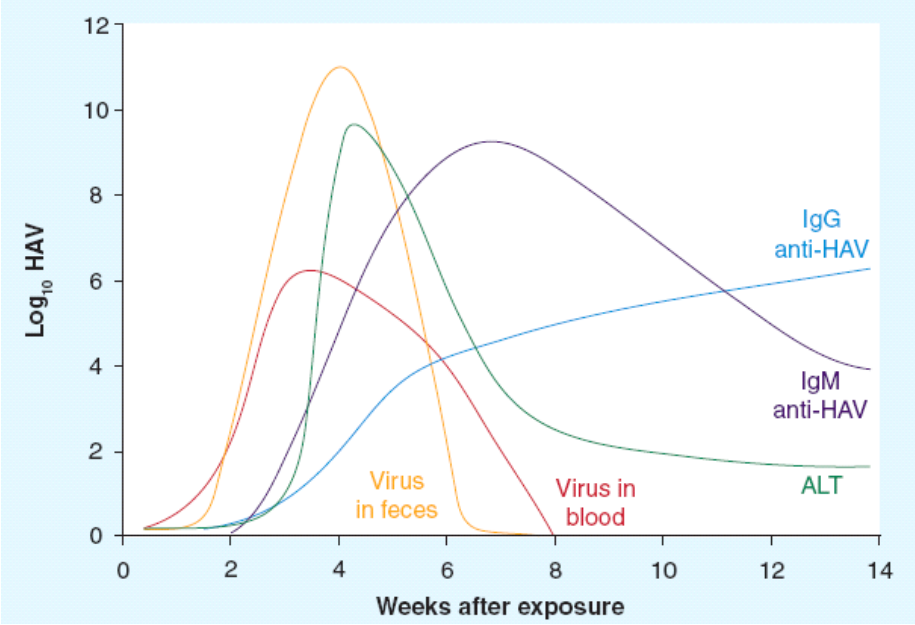

Figure 1.4: Viral titre, immunological response and biochemical signs during HAV infection (Pintó *et al*., 2012)

Viruses are shed in faeces for approximately three weeks after the onset of symptoms (Martin and Lemon, 2006; Pintó *et al*., 2012). Prolonged faecal shedding, up to 10 weeks or six months after onset, has also been documented, but mainly in infected premature infants (Beard and Lemon, 1999; Martin and Lemon, 2006; Dotzauer, 2008; Spradling *et al*., 2009). In addition, a high viraemia occurs during the incubation period, reaching peaks of up to 10^7 genome copies/mL of sera just before the onset of symptoms (Figure 1.4) (Pintó *et al*., 2012). Contrary to faecal shedding, viraemia can last for an average of six weeks after symptoms appear (Bower *et al*., 2000; Costafreda *et al*., 2006; Pintó *et al*., 2012). The fact that faecal shedding and viraemia reach peaks in the absence of symptoms, have great epidemiological significance.

The prodromal period occurs abruptly and is characterised by abdominal pain anorexia, diarrhoea, fatigue, malaise, myalgia, nausea and vomiting (Hollinger and Emerson, 2007; Dotzauer, 2008; Spradling *et al*., 2009; Tang, 2009; Pintó *et al*., 2012). Diarrhoea is more common in younger children (Hollinger and Emerson, 2007; Spradling *et al*., 2009). The prodromal phase is short and can vary from a day to more than a week (Hollinger and Emerson, 2007; Dotzauer, 2008). The prodromal symptoms resolve at the onset of the icteric phase.

The icteric phase is characterised by jaundice which ranges from 2 to 22 days, with a mean duration of three weeks. During this period, the infected individual sees the appearance of dark urine due to bilirubinuria, clay-coloured stools and yellowish colouration of the skin and sclera (Beard and Lemon, 1999; Hollinger and Emerson, 2007; Dotzauer, 2008; Tang, 2009). Once clinical symptoms have resolved, it takes three to six weeks for an infected individual to completely recover. However, fatigue might still remain for another two to four months (Dotzauer, 2008).

1.3.2.2 Immune response

Three classes of antibodies are elicited during HAV infection, namely immunoglobulins A, G and M (IgA, IgG and IgM). Immunoglobulin M and G neutralises HAV particles (Pintó *et al*., 2012). Neutralising anti-HAV IgM antibodies, detected in the sera at the onset of symptoms, last for 6 to 12 months and decreases as the levels of anti-HAV IgG antibodies increases (Figure 1.4). Anti-HAV IgA antibodies are also produced at the onset of symptoms and have been detected in sera, faeces and saliva, but the exact function they fulfil has yet to be identified (Nainan *et al*., 2006). The neutralising anti-HAV IgG antibody response occurs last, but confers a lifelong immunity against subsequent re-infection with HAV even though titres may become undetectable after decades (Beard and Lemon, 1999; Pintó *et al*., 2012).

1.3.2.3 Biochemical parameters

Biochemical abnormalities are also observed in the course of symptomatic HAV infection. The levels of aminotransferases such as alanine aminotransferase (ALT) (Figure 1.4) and to a lesser extent aspartate aminotransferase (AST) in the serum are elevated and may indicate severity of the disease. High levels of serum bilirubin and alkaline phosphatase activity also accompany clinical signs (Beard and Lemon, 1999; Hollinger and Emerson, 2007; Dotzauer, 2008; Spradling *et al*., 2009).

1.3.2.4 Complications of hepatitis A

Individuals infected with HAV might have a prolonged course of infection or relapse at least once after complete resolution of the symptoms. In the first case, jaundice and elevated serum bilirubin can persist for up to 17 weeks causing a form of hepatitis known as cholestatic hepatitis A that is characterised by extensive itchiness (Dotzauer, 2008). In relapsing hepatitis A, anti HAV IgM antibodies reappear in the serum, the infection is usually milder when compared to the first episode and viruses are again detected in faeces and serum (Hollinger and Emerson, 2007; Dotzauer, 2008; Shouval, 2011). Patients with either prolonged or relapse courses of infection should be considered infectious (Cuthbert, 2001).

Occasionally, in patients with underlying chronic liver disease, such as chronic hepatitis B or hepatitis C, acute hepatitis A might result in fulminant hepatitis A. Fulminant hepatitis A is characterised by rapid onset of liver failure and possibly coma. This type of hepatitis mainly occurs in patients over 50 years of age and there is a 30 to 60% chance of full recovery, although liver transplantation may be required in some cases (Hollinger and Emerson, 2007; Dotzauer, 2008; Spradling *et al*., 2009).

1.3.2.5 Extrahepatic manifestations

Extrahepatic manifestations of hepatitis A are uncommon. Haemolysis, pancreatitis, arthralgia and autoimmune hepatitis have been reported in patients with hepatitis A (Hollinger and Emerson, 2007; Spradling *et al*., 2009; El-sayed

and El-karaksy, 2012). The relationship between hepatitis A and these complications is yet to be clearly defined.

1.3.3 Laboratory diagnosis of hepatitis A

Hepatitis A is clinically similar to other types of viral hepatitis. The different types of acute viral hepatitis can only be distinguished by virus-specific laboratory diagnosis. The diagnostic gold standard for HAV infection is the detection of anti-HAV IgM antibodies in sera (Amado *et al*., 2008; Dotzauer, 2008). Alternatively, hepatitis A can also be diagnosed by detecting anti-HAV IgG antibodies which are present within two weeks of onset of jaundice. However, anti-HAV IgM will also have to be tested for in order to differentiate between past and current infection (Cuthbert, 2001; Nainan *et al*., 2006). Radioimmunoassays (RIA), enzyme immunoassays (EIA), immunoblotting and dot blot immunogold filtration are some of the methods used to detect anti-HAV IgM antibodies (Nainan *et al*., 2006).

Hepatitis A virus can be isolated from clinical specimens using cell culture. However, HAV propagation *in vitro* is a laborious time consuming process that is not always fruitful. Wild-type HAV grows poorly in cell culture and rarely produces cytopathology. After a period of four to eight weeks of culturing, HAV antigen can be detected by immunological assays such as fluorescent focus assays or *in situ* hybridisation, RIA, EIA, radio immunofocus assays (RIFA), or molecular-based techniques such as RT-PCR and antigen-capture PCR (AC-PCR) (Nainan *et al*., 2006). For these reasons, HAV isolation by cell culture is rarely used for diagnostic purposes.

Detection of HAV RNA, in sera and stool, is an alternative method for diagnosing hepatitis A but is mainly used for research purposes (Hollinger and Emerson, 2007). Alternative specimen such as saliva has been proposed for the diagnosis of HAV infection (Oba *et al*., 2000; Amado *et al.,* 2008, 2011).

1.4 EPIDEMIOLOGY

1.4.1 Global epidemiology

Hepatitis A is predominantly acquired through direct contact with an infected individual (Nainan *et al*., 2006; Rodríguez-Lázaro *et al*., 2012; Pintó, *et al*., 2010, 2012). Epidemiological studies of hepatitis A have identified and defined distinct patterns of the disease worldwide, attributed to the strong link between HAV infection risk and socio-economic indicators, the age-dependant clinical expression of the disease and lifelong immunity (Jacobsen and Wiersma, 2010; Franco *et al*., 2012). Hepatitis A virus infection is highly endemic in lower socio-economic communities such as those found in less developed countries and in some regions of developing countries. In these areas, conditions such as poor levels of sanitation, household crowding and inadequate water supplies prevail. These living conditions facilitate the spread of HAV infection among young children living in these areas and almost all adults are immune to the disease (Jacobsen, 2009; Jacobsen and Wiersma, 2010; Franco *et al*., 2012). According to the World Health Organization (WHO), highly endemic areas include most of Africa, Asia and Central and South America (Jacobsen, 2009). In contrast, HAV has a low to moderate endemicity in industrialised countries, where sanitary conditions are relatively good. As a result, most adults who have not been infected during childhood are susceptible to the infection, leading to the occurrence of hepatitis A outbreaks (Pintó, *et al*., 2010). Areas of low endemicity of hepatitis A include North America, Western Europe, Australia and Japan (Jacobsen, 2009; Jacobsen and Wiersma, 2010; Franco *et al*., 2012). In these areas, cases of hepatitis A can still be recorded especially among specific adult risk groups, namely international travellers to endemic regions, intravenous drug users and men with a history of homosexual behaviour (Nainan *et al*., 2006; Stene-Johansen *et al*., 2007; Tortajada *et al*., 2009, 2012). Occasional foodborne or waterborne outbreaks also add to the number of recorded cases (Dotzauer, 2008; Franco *et al*., 2012). High risk groups also includes the elderly, people with human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS), medical staff working in neonatal

intensive care units, children attending day-care centres, their families and the staff working at these centres (Beard and Lemon, 1999; Cuthbert, 2001; Dotzauer, 2008).

In areas with improved sanitation and access to clean water, an epidemiological shift from intermediate to low prevalence is noted, with a consequent increase in the susceptibility to infection of the older children and adults with a resultant increase in symptomatic cases and disease severity (Jacobsen and Wiersma, 2010). Regions, such as the Mediterranean basin, where an epidemiological shift is recorded can have an increasing burden from infection with HAV. The growing number of symptomatic cases can impact on the development of the affected area. During the symptomatic phase, an infected individual is unable to work, go to school or participate in social activities in order to prevent possible person-to-person transmission that will spread the infection further. At a personal level this could be costly for the infected individuals and their families. At the level of the country an increase in the number of symptomatic cases of hepatitis A could have a serious impact on the economy especially in the trade and tourism sectors or in terms of loss of productivity in general (Jacobsen and Wiersma, 2010). Hepatitis A is hyperendemic in SA with characteristics of both developing and industrialised countries (Taylor, 1997). However, a change in the epidemic vulnerability of the SA population can be expected with the current trends in urbanisation (Venter, 2004).

The physical properties of the HAV capsid allows it to survive outside its host long enough to provide a common source of transmission. Outbreaks of hepatitis A due to contaminated food and water sources have been well documented (Nainan *et al*., 2006; Rodríguez-Lázaro *et al*., 2012; Pintó, *et al*., 2010, 2012). Blood-borne transmission of HAV is a possible but uncommon cause of infection (Kedda *et al*., 1995; Bower *et al*., 2000; Pintó and Saiz, 2007; Vaughan *et al*., 2013a). Cases of hepatitis A were reported in neonates and adults after blood transfusion and in haemophilic patients who received contaminated factor VIII concentrate (Beard and Lemon, 1999; Bower *et al*., 2000; Cuthbert, 2001). Because it lacks an envelope, HAV can survive the

solvent-detergent method used to inactivate viruses during the preparation of factor VIII concentrates (Cuthbert, 2001; Spradling *et al*., 2009).

1.4.2 Role of food and water in the transmission of hepatitis A virus

1.4.2.1 Water sources

Hepatitis A virus waterborne outbreaks are scarce, but water may constitute a major vehicle in the transmission of the virus (Nainan *et al*., 2006; Bosch *et al*., 2011; Rodríguez-Lázaro *et al*., 2012). Recently, a viral hepatitis outbreak was investigated in the Malwa region of Punjab, India. The epidemiological investigation found that the outbreak was caused by HAV and HEV as a result of faecal contamination of drinking water (Arora *et al*., 2013). An infected individual sheds approximately 10^{11} HAV particles per gram of faeces (Pintó *et al*., 2010). These viral particles are not always effectively removed or inactivated by current methods of sewage treatment (Vantarakis and Papapetropoulou, 1999; Rodríguez-Lázaro *et al*., 2012). Consequently, the highly resistant virus could be released into the environment from treated and untreated wastewater and contaminate various water bodies such as surface waters and groundwater. The virus can survive for extended period of time in water environments. This could pose a serious health hazard since ingesting a minimum of 10 infectious HAV particles can cause disease (Rodríguez-Lázaro *et al*., 2012). Hepatitis A virus can persist for up to 60 days in tap water, at least 42 days in river water, more than 56 days in groundwater and a maximum of 92 days in seawater (La Rosa *et al*., 2012; Rodríguez-Lázaro *et al*., 2012). In addition, HAV can be present in sludge and pose a risk of contamination especially if used as manure (Rodríguez-Lázaro *et al*., 2012). Hepatitis A virus contamination of wastewaters (treated or not), surface waters and drinking waters have been reported (La Rosa *et al*., 2012), as well as outbreaks linked to wells (Bowen and McCarthy 1983; La Rosa *et al*., 2012), recreational (lakes, pool, thermal pool/spa) (Mahoney *et al*., 1992; Sinclair *et al*., 2009) and drinking water (Craun *et al*., 2010).

Surface waters are used as source waters for water treatment plants by municipalities in many countries (Rodríguez-Lázaro *et al*., 2012). The general scheme for surface water treatment intended for human consumption consists of sedimentation, filtration and/or disinfection. Disinfection using a free residual chlorine concentration of 2.0 to 2.5 mg/L for 15 minutes can inactivate HAV particles in water (Hollinger and Emerson, 2007). However, the quality of source water has a great impact on the safety of drinking water produced (Lodder *et al*., 2010; Rodríguez-Lázaro *et al*., 2012). The amount of viruses present in the source water can influence the efficiency of drinking water processes. Drinking water outbreaks caused by HAV have been reported (Craun *et al*., 2010).

1.4.2.2 Food sources

Viral illnesses attributed to food are reported worldwide and HAV is one of the leading causes of viral foodborne infection (Koopmans and Duizer, 2004; Bosch *et al*., 2011; Rodríguez-Lázaro *et al*., 2012). During the period 2000-2008, 3 576 cases of hepatitis A associated with food were reported in the USA (Scallan *et al*., 2011). Globalisation of the food market and changes in food consumption habits has led to the increase in the incidence of foodborne viral diseases, in particular for hepatitis A (Berger *et al*., 2010; Rodríguez-Lázaro *et al*., 2012).

Fresh produce can be contaminated at various points throughout the chain of production (Fiore, 2004; Berger *et al*., 2010; Rodríguez-Lázaro *et al*., 2012). At the pre-harvest stage, the produce can be contaminated directly through irrigation with water contaminated with faeces containing HAV. During post-harvesting processing food products can also be exposed when polluted water is used for washing. The use of contaminated water to mix agrichemicals or inadequate sanitary facilities for food handlers at the point of harvest, at commercial outlets or at preparation facilities also poses risk of spreading the virus (Fiore, 2004; Sun *et al*., 2012). The importance of infected food handlers and contaminated food surface preparation in the transmission of HAV is increasingly being recognised (Mbithi *et al*., 1992; Deboosere *et al*., 2012). The

virus can survive on human hands for at least four hours and be transferred with ease to clean animate and inanimate surfaces.

Food products that are minimally processed such as shellfish, fresh produce and soft fruits are important vehicles in the transmission of HAV (Cuthbert, 2001; Wheeler *et al*., 2005; Croci *et al*., 2008; Pintó *et al*., 2009, 2010). Shellfish grown in contaminated water can concentrate the virus and transmit it to the consumer. To date, the largest recorded outbreak of hepatitis A involved the consumption of sewage-contaminated clams and approximately 300,000 people became ill (Halliday *et al*, 1991; Nainan *et al*., 2006; Pintó *et al*., 2010). Hepatitis A virus can survive on the surface of fresh produce for extended periods of time and remain infectious during normal conditions of commercial and household storage (Croci *et al*., 2002; Rodríguez-Lázaro *et al*., 2012). The virus can survive for at least five days on the surface of lettuce leaves (Bidawid *et al*., 2001; Croci *et al*, 2002) and on green onions (Sun *et al*., 2012), and for a maximum of four days on carrots (Croci *et al*, 2002). Washing contaminated lettuce leaves and carrots was found to be ineffective at removing all HAV particles (Croci *et al*, 2002). Foodborne outbreaks of hepatitis A have been linked to contaminated lettuce (Rosenblum *et al*., 1990) and green onions (Centers for Disease Control and Prevention (CDC) foodborne outbreak online database; Dentinger *et al*., 2001; Wheeler *et al*., 2005). Contaminated semidried tomatoes (Gallot *et al*., 2011; Donnan *et al*., 2012), frozen berries (Hutin *et al*., 1999; Lassen *et al*., 2013) and raw oysters (Desenclos *et al*., 1991; Bialek *et al*., 2007) have caused multistate outbreak of hepatitis A, where one contaminated batch was distributed to geographically distinct regions and consumers got infected. Various biocides exist to disinfect food items, but the properties of the HAV capsid often requires the application of doses higher than those recommended by the manufacturer and that could affect the health of consumers (Deboosere *et al*., 2012). Thus preventing HAV contamination of food items could be a good approach to decrease foodborne outbreaks of hepatitis A.

1.5 PREVENTION AND CONTROL OF HEPATITIS A

During the year 2010, the WHO adopted resolutions that would help control and prevent viral hepatitis in the world in general and in Africa in particular (Lavanchy, 2012; Mihigo *et al*., 2013). One of the first steps was to raise more awareness by marking World Hepatitis Day on 28th July. It was also realised that infection with HAV can be prevented by using available vaccines according to the epidemiology of the location or by simply educating the public and improving living conditions (Lavanchy, 2012; Mihigo *et al*., 2013).

Access to safe drinking water and proper disposal of sewage, coupled with public education on good hygiene practices like careful and regular hand-washing can reduce the spread of HAV in the absence of a vaccination programme (Lavanchy, 2012). Improvements in living standards had a great impact in the decrease of infection rate in developed countries as well as in regions where an epidemiological shift was recorded (Franco *et al*., 2012).

Hepatitis A is a vaccine preventable disease (Fitzsimons *et al*., 2010; Jacobsen and Wiersma, 2010). Two types of monovalent HAV vaccines are used worldwide, namely inactivated HAV vaccines and live attenuated vaccines. The first type is used in most countries and given in two doses at 6 – 36 month intervals. The second type of vaccine (a single dose) is manufactured and mainly used in China and occasionally in India (Shouval, 2011; Lavanchy, 2012; Xu and Wang, 2013). The licensed vaccines are highly immunogenic and safe to use. These vaccines prevent infection with HAV for a minimum of 15 – 30 years, with a high probability of lifelong immunity provided that the vaccination scheme of two doses has been completed. However, to date there are no long-term post-marketing surveillance studies to tell if that is the case or if there is a need for immune booster and after what period of time (Vandamme *et al*., 2003; Bovier *et al*., 2010; Lavanchy, 2012). Vaccines against HAV can also be given in combination with other vaccine like hepatitis B or typhoid fever vaccines (de Jong *et al*., 2007; Shouval, 2011; Lavanchy, 2012).

Immunisation programmes against HAV infection are not universal. According to the WHO, epidemiological data of a region, where vaccination is being considered, is required to assess cost-benefit and sustainability of different prevention strategies (Lavanchy, 2012). In highly endemic regions, most of the population is immune to HAV following asymptomatic infection, thus vaccination is not recommended (de Jong *et al*., 2007; FitzSimons *et al*., 2010; Jacobsen and Wiersma, 2010; Franco *et al*., 2012). In low endemic regions, immunisation programmes should target high-risk population groups such as travellers to highly endemic region, intravenous drug users, patients with chronic liver disease or with blood-clotting disorders receiving blood-derived products, immunocompromised persons (HIV/AIDS patients), men with a history of homosexual behaviour, raw sewage workers and individuals working at day-care centres or as caretakers of non-human primates (de Jong *et al*., 2007; FitzSimons *et al*., 2010; Jacobsen and Wiersma, 2010; Shouval, 2011). Vaccination schemes should be completed for the high-risk population groups, especially the immunocompromised. Because of their impaired immune system, they do not develop enough immunity against HAV and repeated exposure to high concentration of the virus could select for vaccine escape variants (Pérez-Sautu *et al*., 2011a; Pintó, *et al*., 2012). In countries with intermediate endemnicity of HAV universal childhood vaccination should be considered in adolescent and young adults that have a high probability of developing a severe form of hepatitis A (FitzSimons *et al*., 2010; Jacobsen and Wiersma, 2010).

Hepatitis A vaccines are designed and recommended for pre-exposure prophylaxis (de Jong *et al*., 2007). The vaccines can also be administered for post-exposure prophylaxis, but should be used together with pooled intramuscular immunoglobulin for long-term protection (de Jong *et al*., 2007; Lavanchy, 2012). Pooled immunoglobulin provides passive immunity to hepatitis A and can be used for pre- or post-exposure prophylaxis. In the case of post-exposure prophylaxis, immunoglobulin should be administered within 14 days of exposure. Immunoglobulin may not prevent HAV infection in the case

of exposure up to four weeks, but may decrease disease severity in high-risk contact (de Jong *et al*., 2007).

1.6 DETECTION AND CHARACTERISATION OF HEPATITIS A VIRUS IN WATER AND FOOD SOURCES

The detection of enteric viruses in environmental samples is a complicated multistage process (Wyn-Jones, 2007; Mattison and Bidawid, 2009). Enteric viruses are present in low numbers in the environment (food or water) (Mattison and Bidawid, 2009; Bosch *et al*., 2011). Thus, analytical methods in food and environmental virology have to concentrate low numbers of viral particles and eliminate inhibitors that could impact on the downstream processes (Mattison and Bidawid, 2009). Three main steps are considered when analysing environmental samples: sampling and sample preparation, viral detection and viral characterisation.

1.6.1 Sampling and sample preparation

Methods used to sample for virological analysis of either water or fresh produce have to ensure that the sample is statistically representive of the sampling environment. Sampling for the virological analysis of water has been well documented (Bosch *et al*., 2011) while a consensus has not been reached yet with regard to food sampling (Croci *et al*., 2008; Bosch *et al*., 2011). For example, in the case of fruit or vegetables, a 10-100 g sample taken from a field, crate or punnet box in triplicate has been proposed (Croci *et al*., 2008; Bosch *et al*., 2011). Once sampling has been done a number of methods exist to extract and concentrate viruses from the different matrices.

1.6.1.1 Water matrices

The efficiency of recovery processes of viruses from water matrices depend on factors like the pH, conductivity, turbidity, presence or not of particulate matter and organic acids (Richards *et al*., 2004; Bosch *et al*., 2011). A wide range of methods have been developed to successfully recover and concentrate enteric

viruses from water, namely adsorption-elution using charged membranes or filters or glasswool or glass powder, the use of polyethylene glycol (PEG), ultrafiltration, ultracentrifugation, magnetic beads (Mattison and Bidawid, 2009; Bosch *et al*., 2011).

The adsorption-elution method, using charged membranes or filters, has been successfully used to recover and concentrate entroviruses, polioviruses, rotaviruses, HAV, noroviruses, adenoviruses, astroviruses and bacteriophages from water and wastewater samples (Wyn-Jones, 2007; Mattison and Bidawid, 2009). Although bacteria can be recovered at the same time, the system requires sample clarification steps prior to processing and may not be easily adapted for the in-line recovery of viruses from water. A variation of the adsorption-elution method using glass wool instead of filters has been successfully implemented to recover human enteric viruses from surface waters in SA (Taylor *et al*., 2001; Venter, 2004; Vivier *et al*., 2004; Van Heerden *et al*., 2005; Bosch *et al*., 2011). The glass wool adsorption-elution method prevents transport of large volumes of water reducing the potential risk of health hazard. The method has been coupled to secondary concentration by PEG/sodium chloride (NaCl) precipitation to increase the efficiency of recovery (Vilaginès *et al*., 1993; Mans *et al*., 2013). Viruses can be precipitated out of solution by the addition and dissolution of PEG and NaCl to the sample eluate. Viral DNA or RNA can be extracted from the resulting pellet for detection methods. Concentration of viruses using PEG is inexpensive, simple to carry out and can be used to test sewage, faecal, fruit and vegetable samples. One disadvantage of using PEG is the overnight incubation period. The precipitation period has been reduced by the use of liquid PEG_{8000} (ISO/TS 15216-2:2013[E]).

Ultrafiltration and ultracentrifugation can concentrate viruses based on particle size and weight, respectively (Mattison and Bidawid, 2009). Both methods have been successfully applied to detect enteroviruses, HAV and noroviruses from environmental samples. Both of these methods are fast and can co-concentrate bacteria and parasites. However, there are several drawbacks to the use of either method to concentrate viruses. Processing requires small

volumes of samples: <1000 mL when using membrane for ultrafiltration, up to 15 L when using ultrafilters and ~10 mL – 1 L for ultracentrifugation (Wyn-Jones, 2007; Ikner *et al*., 2012; Prata *et al*., 2012). The processing of turbid water is slow and requires a prefiltration step to remove larger debris and cells. The prefiltration step can result in loss of viruses due to membrane clogging by non-filterable components (Wyn-Jones, 2007; Ikner *et al*., 2012; Prata *et al*., 2012). In addition, the equipment required for ultrafiltration and ultracentrifugation are expensive and the probability of co-concentrating inhibitors is high (Wyn-Jones, 2007; Ikner *et al*., 2012).

Magnetic beads coated with antibodies have been used to extract human enteric viruses from environmental samples (Mattison and Bidawid, 2009). Coated beads bind to viral particles and the bead-virus complexes are magnetically separated from the sample and washed thoroughly. The main advantage of using magnetic beads technology is that it yields concentrated and purified viral particles, but the method is expensive and highly specific. However, concentration of HAV particles using magnetic beads may not be limited by the specificity of the technique as a single serotype of the virus has been identified.

1.6.1.2 Food matrices

The principle behind virus release from food matrices depends on the type of food matrix and the suspected route of contamination. When analysing shellfish, viral particles should be extracted from digestive tissues, while in the case of fruit and fresh produce viral elution from the surface is more appropriate (Croci *et al*., 2008; Bosch *et al*., 2011). The tris-glycine beef extract (TGBE) buffer has been successfully used to release viruses from the surface of fruit or vegetables (Dubois *et al*., 2007; Bosch *et al*., 2011). Once eluted, viruses can be concentrated from the eluate using precipitation by PEG, ultrafiltration, ultracentrifugation and magnetic beads. A direct elution and/or protease treatment can also be used to extract viruses from food products without additional concentration step. However, this method is less sensitive and is

only recommended for food products with suspected high viral load like shellfish (Mattison and Bidawid, 2009).

1.6.2 Viral detection

Viral detection is accomplished by molecular-based or non-molecular based methods. Infectious viral particles can be detected by propagation in cell culture and quantified by fifty per cent tissue culture infectious dose ($TCID_{50}$), the most probable number (MPN), or plaque assays (Wyn-Jones, 2007; Mattison and Bidawid, 2009; Bosch *et al*., 2011). However, cell culture methods rely on the formation of a cytopathic effect (CPE) and HAV grows poorly in cell culture and seldom produce CPE. Therefore, other techniques such as RIA, EIA, RIFA or AC-PCR have to be applied to detect infected cells. Electron microscopy, IEM, EIA, electrophoresis and chromatography are alternative non-molecular based methods that require a minimum of 10^5 viral particles for detection and are therefore not applicable to most environmental samples (Mattison and Bidawid, 2009). Molecular based methods are more sensitive, specific and reproducible and have been successfully implemented for viral detection in food and environmental virology (Mattison and Bidawid, 2009; Bosch *et al*., 2011).

The first step in the detection of viral genomes is nucleic acid extraction from viral concentrates (Mattison and Bidawid, 2009; Bosch *et al*., 2011). A variety of extraction protocols exist and commercial kits are also available. The majority of these kits use an extraction system based on the Boom's method according to which lysis with guanidium thiocyanate is followed by nucleic acid binding to silica beads (Boom *et al*., 1990; Mattison and Bidawid, 2009; Bosch *et al*., 2011). The bound nucleic acid is purified before elution to further reduce the amount of inhibitors.

Molecular methods can be used to detect viruses in food and environmental samples. These methods are commonly used by many laboratories, because of their sensitivity and specificity (Wyn-Jones, 2007; Mattison and Bidawid,

2009). Molecular assays commonly used for viral genome detection are PCR and RT-PCR, (Mattison and Bidawid, 2009). These molecular systems are available in real-time formats. The real-time format PCR and RT-PCR assays combine PCR amplification and detection of amplified product in a single reaction mix. The detection can be qualitative or quantitative. Real-time PCR assays are susceptible to inhibition, thus the need for proper controls. Inhibitors of PCR can be removed by adding an antigen capture stage during PCR analyses (AC-PCR) or concentrating viral particles with immunomagnetic beads (immunomagnetic capture RT-PCR [IC-PCR or IMC-PCR]) (Deng *et al*., 1994; Jothikumar *et al*., 1998; Casas and Suñén, 2002). Inhibition during PCR analyses can also be overcome by diluting the sample or using an alternative molecular technique, like the nucleic acid sequence based amplification (NASBA) or loop-mediated isothermal amplification (LAMP) (Mattison and Bidawid, 2009; Bosch *et al*., 2011). Nucleic acid sequence based amplification has been successfully developed for the detection of HAV in environmental samples (wastewater, lettuce and blueberries) (Jean *et al*., 2001). Despite the high speed and sensitivity, viral detection by PCR cannot differentiate between infectious and non-infectious viruses. The integrated cell culture–PCR (ICC-PCR or ICC/real-time PCR) assay is a possible solution, where molecular detection of viruses is coupled to cell culture infection (Venter *et al*., 2007; Wyn-Jones, 2007; Bosch *et al*., 2011). Jiang *et al* (2004) and Cromeans and Sobsey (2004) used an ICC/strand-specific real-time PCR to differentiate infectious and non-infectious HAV in water samples.

1.6.3 Viral characterisation

Detected viruses can be classified by serotyping or nucleotide sequence analysis. Hepatitis A virus strains can be characterised by nucleotide sequence analysis of specific genomic regions (Costa-Mattioli *et al*., 2002; Nainan *et al*., 2006; Vaughan *et al*., 2013a) to determine genetic relatedness among detected strains (Hollinger and Emerson, 2007), and possibly trace the origin of an outbreak (Pintó *et al*., 2010). Viral genetic analyses have been applied to the investigation of outbreaks of hepatitis A associated with green onions (Wheeler

et al., 2005), lettuce (Cuthbert, 2001), semidried tomatoes (Gallot *et al*., 2011; Donnan *et al*., 2012), frozen berries (Hutin *et al*., 1999; Lassen *et al*., 2013) and raw oysters (Desenclos *et al*., 1991; Bialek *et al*., 2007; Shieh *et al*., 2007). In the later three cases, molecular characterisation was able to trace back the source of contamination and link outbreaks occurring in geographically separated regions. However, caution should be exercised when using viral genomic regions to trace the origin of outbreaks. Viral fitness and mutations (although limited) can result in the characterisation of different strains in potential source virus, i.e. contaminated water or food, and in infected recipients (Pintó *et al*., 2010; Vaughan *et al*., 2013b). A solution would be to use whole genome sequences of HAV strains for molecular source tracking and use the different genomic regions to identify transmission among clinical cases with identical epidemiological association (Vaughan *et al*., 2013b).

1.7 RATIONALE FOR STUDY

Hepatitis A virus has been detected in surface water sources in SA (Taylor *et al*., 2001; Venter, 2004) and the risk of infection for the different communities has been quantified (Venter *et al*., 2007). Hepatitis A virus has also been detected on fresh produce at the point of retail in SA (Netshikweta, 2011). Therefore, in SA, faecally-contaminated food and water is a possible source of HAV infection. Furthermore in the 1990s it was established that, compared to the rest of the world where genotype IA prevails (Robertson *et al*., 1992; Nainan *et al*., 2006), genotype IB predominates on the African continent (Robertson *et al*., 1992) including SA (Taylor, 1997) but not in Nigeria (Forbi *et al*., 2013), Cameroon (Forbi *et al*., 2012) and Tunisia (Gharbi-Khelifi *et al*., 2012) where IA prevails. Within genotype IB a unique cluster of SA-related strains has been identified with some of these strains implicated in hepatitis A in a group of German tourists who contracted the infection in SA (Faber *et al*., 2009). Although HAV is an antigenically stable virus with an effective inactivated vaccine, recent publications have reported the detection of “escape” mutants which could result in vaccination failure (Pérez-Sautu *et al*., 2011b; Pintó *et al*., 2012). Since the studies of Taylor (1997) and Venter (2004), there have been

no molecular epidemiological investigations in SA to establish whether or not different genotypes or mutant strains are emerging or circulating in the country. In addition to monitoring HAV strains in clinical specimens, the analysis of faecally polluted water and food samples provides a passive surveillance system as to which HAV strains are circulating in the South African communities. This investigation will identify high risk water and fresh produce sources, as well as provide valuable HAV genotype data for molecular source tracking in the event of an outbreak.

1.8 HYPOTHESIS

In contrast to early investigations, multiple genotypes and mutant strains of HAV are now circulating in the SA community.

1.9 AIM

The aim of this investigation is to establish which genotypes, and possibly variants thereof, of HAV are circulating the community compared to the genotypes present in water and food sources in SA.

1.10 SPECIFIC OBJECTIVES

The specific objectives of this investigation are:

i) To recover and detect HAV in environmental samples, i.e. fresh produce and faecally-contaminated surface water;
ii) To detect HAV by real-time RT-PCR in clinical specimens (stools and serum);
iii) To genotype, by nucleotide sequence analysis, HAV strains detected in clinical and environmental sources;
iv) To compare the nucleotide sequences of the different HAV strains by pairwise comparison to each other and published sequences for HAV, and to perform phylogenetic analysis in order to establish the genetic relatedness between strains from the different sources;
v) To analyse nucleotide sequences for the presence of novel mutations.

CHAPTER 2

DETECTION AND GENOTYPING OF HEPATITIS A VIRUS STRAINS IN CLINICAL AND ENVIRONMENTAL SOURCES

The irrigation water and fresh produce samples in this Chapter formed part of a collaborative Water Research Commission (WRC) solicited research project co-funded by Department of Agriculture, Forestry and Fisheries
"An investigation into the link between water quality and microbiological safety of fruit and vegetables from the farming to the processing stages of production and marketing (WRC Project no K5/1875/4, Water Research Commission, 2009)"

2.1 INTRODUCTION

Hepatitis A virus is a genetically stable RNA virus that infects both human and non-human primates (Nainan *et al*., 2006; Spradling *et al*., 2009; Knowles *et al*., 2012; Pintó *et al*., 2012). However, the genome of HAV is diverse enough to allow characterisation and classification of strains based on nucleotide sequence analysis. Nucleotide sequence analysis of the full VP1 region is used to distinguish among the six genotypes of HAV. The junction VP1/P2B has also been successfully used in molecular epidemiological investigations (Robertson *et al*., 1992; Wheeler *et al*., 2005; Nainan *et al*., 2006; Bialek *et al*., 2007; Shieh *et al*., 2007; Vaughan *et al*., 2013b). The six genotypes of HAV have a distinct geographical distribution which makes it possible to use molecular source tracking to identify a common source of exposure or contamination (Bosch *et al*., 2001; Sánchez *et al*., 2002; Pintó *et al*., 2009).

Hepatitis A virus is one of the leading causes of viral foodborne infection (Bosch *et al*., 2011; Rodríguez-Lázaro *et al*., 2012). Given the globalisation and

complexity of food chain production, source tracking in the case of a foodborne outbreak is a tedious and not always fruitful process. For this reason and many others, the best method to prevent foodborne outbreak is prevention. Prevention of hepatitis A foodborne outbreak could be through identification of possible source(s) of contamination and or exposure. Given that HAV is predominantly transmitted via the faecal-oral route, effluents and sludge of wastewater plants, as well as surface water and groundwater are potential sources of contamination of irrigated fruits and vegetables at pre- and post-harvest points (Okoh *et al*., 2010; Rodríguez-Lázaro *et al*., 2012). In addition, analysing environmental samples is a form of passive surveillance which aims to identify the different types of HAV circulating in the surrounding communities and the potential sources of exposure. These data together with the analysis of HAV strains from clinical specimens will help determine if environmentally detected viruses are clinically relevant.

The objective of this chapter was to determine the genetic relatedness of HAV strains detected in clinical specimens and in environmental samples.

2.2 MATERIALS AND METHODS

2.2.1 Ethics statement

The protocol for this study (number: 211/2012) was reviewed and approved by the Research Ethics Committee, Faculty Health Sciences, University of Pretoria. The Research Ethics Committee complies with the International Conference on Harmonisation–Good Clinical Practice (ICH-GCP) guidelines and has US Federal wide Assurance. There was no physical contact with patients from which clinical specimens were obtained and the identity of these patients was not revealed in the study.

2.2.2 Clinical specimens

2.2.2.1 Specimens collection

Clinical specimens (sera and stool) were collected from anti-HAV IgM positive individuals who presented with signs and symptoms of hepatitis A. Specimens were collected from SA and neighbouring countries. These specimens were obtained from the National Health Laboratory Service (Virology Diagnostic Laboratory, Tshwane Academic Division; Groote Schuur Hospital, Cape Town, Western Cape; Tygerberg Hospital, Western Cape) and Lancet Laboratories.

2.2.2.2 RNA extraction

Total RNA was extracted from 140 µL of sera and 140 µL of a 10% stool suspension (0.3 g of stool diluted in 2 mL of nuclease-free water [Promega Corp., Madison, WI]) using the QIAamp® Viral RNA Mini Kit (Qiagen, Hilden, Germany). The kit includes a RNA carrier and is based on the spin column technology. It is a solid phase extraction where nucleic acid from cell or virus lysates binds to a solid phase, namely silica, under certain conditions. The bound virus is washed several times to remove inhibitors and extraneous material before being eluted from the solid phase. Extraction was carried out as per manufacturer's instructions. Total RNA was eluted in 60 µL elution buffer, aliquoted and stored at -70°C.

2.2.3 Water samples

2.2.3.1 Sample collection

From January 2012 to August 2012, water samples (10 L) (*n=26*) were collected monthly (except in February and July) from seven sites near Polokwane, Limpopo province (LP) (WRC project) (Figure 2.1). These samples originated from rivers, dams and boreholes used for domestic purposes by local communities and to irrigate fresh produce, namely tomatoes, on a commercial farm.

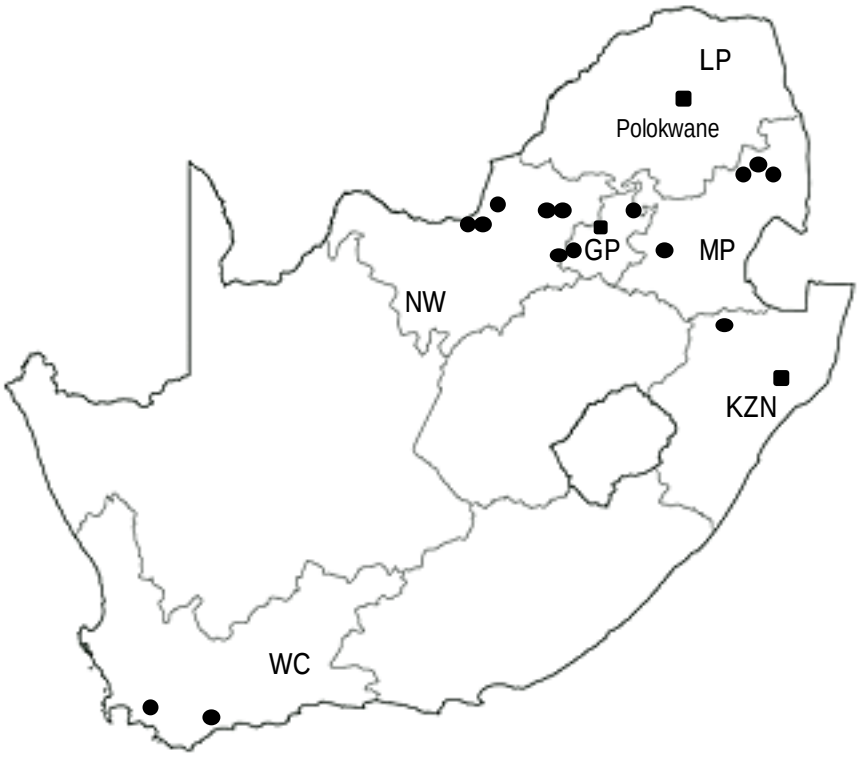

Figure 2.1: Distribution of irrigation water (■) and wastewater (●) sampling sites among six provinces of South Africa. Gauteng (GP); Kwazulu-Natal (KZN); Limpopo (LP); Mpumalanga (MP); North West (NW); Western Cape (WC) (adapted from Murray *et al.*, 2013).

In May 2013 and in August 2013 additional water samples (10 L) were collected from a commercial cabbage farm in KwaZulu-Natal (KZN) (*n=1*) and from a small-scale farm in Gauteng (GP) (*n=2*).(Figure 2.1). The sample collected in KwaZulu-Natal originated from a dam used for irrigation, post-harvest washing and agrichemicals preparation. The samples collected in Gauteng originated from a borehole used for irrigation purposes and a river close to the cabbage fields. The irrigation water samples were transported in cooler bags with cold packs to the Department of Medical Virology, Faculty of Health Sciences, University of Pretoria. The temperature and pH of each sample were recorded upon arrival. The samples were either processed immediately or stored at 4°C before being processed within 24 h of receipt.

From August 2010 to December 2011, wastewater samples (500 mL) (*n=51*) were collected from the outflow of wastewater treatment works in three

provinces: GP (*n=9*), Mpumalanga (MP, *n=20*) and North West (NW, *n=16*) (Figure 2.1). In addition, sludge was collected in a treatment plant in KZN (*n=2*) and in the Western Cape (WC, *n=2*); two samples were of unknown provincial origin (Figure 2.1). The wastewater samples were referred to the Department of Medical Virology, Faculty of Health Sciences, University of Pretoria for microbial indicator analysis. These samples were selected for this study as they had thermotolerant (faecal) coliform counts greater than 10^6 colony forming units (cfu) per 100 mL (Murray *et al*., 2013).

Strains of HAV from 11 surface water samples (10 L) from GP (*n=6*) and from Kenya (*n=5*) samples which had tested positive for HAV were included in this study. The samples from GP were referred to the Department of Medical Virology, Faculty of Health Sciences, University of Pretoria for routine virological analysis for enteric viruses. The samples from Kenya were from a SA/Kenya Research Co-operation project (UID 82467) funded by the National Research Foundation. The project aimed to detect clinically relevant viruses in surface water.

2.2.3.2 Microbial indicator analysis

The *Escherichia coli* and thermotolerant (faecal) coliform counts, expressed as cfu/100 mL were determined for the irrigation water and the wastewater samples, respectively. The choice of method for the different matrices was specified in the protocols of each sample type. The membrane filtration technique coupled with m-ColiBlue24® broth (Millipore Corp., Billerica, MA) or M-Fc agar (Selecta Media, Quantum Biotechnologies [Pty] Ltd, Ferndale, SA) (American Public Health Association, American Water Works Association, Water Environment Federation, 2005) was used to determine the *E. coli* and thermotolerant (faecal) coliform counts, respectively.

2.2.3.3 Viral recovery and concentration

2.2.3.3.1 *Viral recovery from 10 L irrigation water samples*

Prior to recovery, 100 µL of mengovirus (5×10^4 copies /10 µL) was added to each sample as a process control to monitor the efficiency of viral recovery and

downstream nucleic acid extraction processes. The virus was evenly spread in the sample by thorough mixing.

Viruses were recovered from the 10 L irrigation water samples using a glass wool adsorption-elution method based on the method of Vilaginès *et al* (1993) as described by Venter (2004) and Mans *et al* (2013). The glass wool column was prepared by compressing 3 x 5 g glass wool (Glass wool Bourre 725 QN, Ouest Isol, Alizay, France) into a Perspex column (260 mm x 30 mm) to a final density of 0.5 g/cm^3. A steel sieve grid (pore size of ~1 mm^2) was added to the column, after every 5 g of glass wool. The column was then soaked with sterile distilled water and pre-treated with 40 mL 1 M hydrochloric acid (HCl) (Merck KGaA, Darmstadt, Germany), 100 mL sterile distilled water and 40 mL 1 M sodium hydroxide (NaOH) (Merck), respectively (Appendix A). After pre-treatment, 200 mL of sterile distilled water was run through the column to adjust the pH to pH 7.0, in order to restore the positive charge of the glass wool. Each water sample was filtered through the glass wool column (one column per sample) by negative pressure at a rate of 10 L/h. The positive charge of the glass wool attracts and traps the negatively charged viruses. The adsorbed viruses were eluted from the glass wool with 100 mL of sterile glycine beef-extract buffer (GBEB, pH 9.5) (Appendix A). The eluting solution (GBEB) reverses the ionic charge of viruses, so as to release them from the glass wool. The eluting solution was left in contact with the glass wool for 10-15 min, and then eluted from the column by positive pressure. Thereafter, the pH of the eluate was neutralised to pH 7.0 with 1 M HCl.

For the secondary concentration, eluted viruses were concentrated by PEG_{8000} (Amresco, Solon, OH) and NaCl (Merck) (PEG_{8000}/NaCl) precipitation (ISO/TS 15216-2:2013[E]) (Appendix A). To concentrate viruses, 0.25 volume of 5 x PEG_{8000}/NaCl solution was added to the eluates. The solution was mixed by inversion and shaken for 60 min at 4°C. After mixing, the solution was centrifuged at 10 000 x g for 30 min at 4°C, then the supernatant discarded and the pellet re-suspended in 10 mL of phosphate buffered saline (PBS, pH 7.2,

[Sigma-Aldrich Co., St Louis, MO]). Aliquots (1 mL and 9 mL) of the concentrates were stored at -20°C for downstream analyses.

2.2.3.3.2 *Viral recovery wastewater samples*

Viruses were recovered from 75 mL to 100 mL wastewater samples to a volume of 2 mL in PBS, using PEG_{8000}/NaCl precipitation (ISO/TS 15216-2:2013[E]). No prior viral recovery was performed as the sample volumes were too small. Virus concentrates were also stored at -20°C.

2.2.3.4 Total nucleic acid extraction

Total nucleic acid was extracted from 1 mL of virus concentrates using the MagNA Pure LC Total Nucleic Acid isolation Kit (Large volume) on the MagNA Pure LC automated platform (Roche Diagnostics GmbH, Mannheim, Germany), or the semi-automated NucliSENS® EasyMAG® platform (BioMérieux, Marcy l'Etoile, France). Nucleic acid extraction on the NucliSENS® EasyMAG® instrument (BioMérieux) was done for the irrigation water samples while the MagNA Pure LC System (Roche Diagnostics) was used for virus concentrates of the wastewater samples. With both methods, extracted nucleic acid was eluted in 100 µL of elution buffer and stored at -70°C.

2.2.4 Fresh produce samples

2.2.4.1 Samples collection

Minimally processed food samples which are usually eaten raw were purchased from informal, e.g. street vendors, and formal commercial outlets, e.g. supermarkets and greengrocers, and collected from commercial farms. A representative sample is defined as three items of a specific vegetable/fruit collected from different areas of a field, punnet or box of produce. The following was included in the study:

i) Leafy green vegetables purchased from a GP street vendor: lettuce (*n=1*);
ii) Tomatoes (*n=9*) purchased in Pretoria, GP;
iii) Cabbages (*n=4*) purchased in Pretoria;

iv) Irrigated tomatoes (*n=4*) from a commercial farm in LP;

v) Cabbages were also collected at commercial farms in KZN (*n=1*) and in GP (*n=1*).

2.2.4.2 Viral recovery and concentration

Prior to recovery, the samples were prepared as follows:

i) Tomatoes, weighing at least 25 g, were seeded with 100 µL mengovirus (5 x 10^4 copies /10 µL). The viral suspension was dispersed onto as much of the surface as possible and allowed to air-dry in a biosafety cabinet used for food processing only;

ii) Leaves of lettuces and cabbages (three leaves per food item) were cut to fit in 145 mm petri dishes (Sterilin Ltd., Newport, United Kingdom) petri dishes and seeded with mengovirus as described previously. The first three leaves of the food product were removed and the next three used for analysis.

The recovery technique that was used is the one described by Croci *et al* (2008) and optimised by Netshikweta (2011). Briefly, the seeded food item was placed into a Ziplock or hybridisation bag (Roche Diagnostics) into which 40 mL of TGBE, (pH 9.5) was added. Air was withdrawn, the bag was sealed and the buffer distributed evenly over the entire bag. In the case of tomato samples, while the item was in the bag along with the buffer, the surface of the sample was gently rubbed to maximise the chance of recovering any virus that might have attached to the surface. Thereafter, the bags, containing the food item and buffer, were shaken on a shaking incubator for 20 min at room temperature (~25°C). After elution, the solution was poured into a sterile 50 mL centrifuge tube. The volume was recorded and the pH adjusted to pH 7.2 with 1 M HCl. The eluates were either stored at 4°C or at -20°C depending on whether the precipitation step will be carried out the next day or in at least two days time, respectively. Viruses were concentrated from the eluates by PEG_{8000}/NaCl precipitation to a final volume of 2 mL in PBS. Aliquots of the concentrates (1000 µL each) were stored at -20°C until further analyses.

2.2.4.3 Total nucleic acid extraction

Total nucleic acid was extracted from 1 mL of virus concentrates using the automated MagNA Pure LC instrument (Roche Diagnostics) or the semi-automated platform of the NucliSENS® EasyMAG® instrument (BioMérieux). In either method, extracted nucleic acid was eluted in 100 µL of elution buffer, aliquoted and stored at -70°C.

2.2.5 Viral detection

2.2.5.1 Mengovirus detection

Extracted nucleic acid (5 µL) from food and irrigation water samples was tested for mengovirus using a one-step real-time RT-PCR assay [mengo@ceeramTools™ Kit (Ceeram s.a.s, La Chappelle-Sur-Erdre, France)]. The cycling conditions of the assay were as follows: 45°C for 10 min, 95°C for 10 min and 45 cycles of 95°C for 15 seconds (s), 60°C for 45 s. The primers (Table 2.1) included in the kit, amplify a 99 bp fragment within the 5' noncoding region (5' NCR) of the mengovirus genome (Pintó *et al*. 2009).

Table 2.1: Sequences of primers and probes used for mengovirus detection (Pintó *et al*., 2009).

Oligonucleotide	**Nucleotide sequence (5'-3')**
Forward primer (Mengo110)	gCg ggT CCT gCC gAA AgT
Reverse Primer (Mengo209)	gAA gTA ACA TAT AgA CAg ACg CAC AC
Probe (Mengo147)	FAM-ATC ACA TTA CTg gCC gAA gC-TAMRA

2.2.5.2 Hepatitis A virus detection

Extracted nucleic acid (5 µL) from clinical specimens and environmental samples (fresh produce, irrigation water and wastewater) was tested for HAV using a one-step real-time RT-PCR assay. Detection of HAV RNA in clinical specimens was performed using hepatitisA@ceeramTools™ Clinical Kit (Ceeram s.a.s) while hepatitisA@ceeramTools™ Environmental Kit (Ceeram s.a.s) was used for food and irrigation water samples. Hepatitis A virus detection in wastewater samples was performed using a RNA Ultrasense™ One-step quantitative RT-PCR (qRT-PCR) system kit (Invitrogen, Carlsbad,

CA). Both assays were carried out in a LightCycler 2.0 (Roche Diagnostics) and used primers which amplified a 172 bp fragment within the most conserved region of the HAV genome, the 5' NCR (Table 2.2).

Table 2.2: Sequences of primers and probes used for HAV detection (Costafreda *et al*., 2006)

Oligonucleotide	**Nucleotide sequence (5'-3')**
Forward primer (HAV68)	TCA CCg CCg TTT gCg TAg
Reverse Primer (HAV240)	ggA gAg CCC Tgg AAg AAA g
Probe (HAV150)	FAM-TTA ATT CCT gCA ggT TCA gg-TAMRA

The hepatitisA@ceeramTools™ Kits (Ceeram s.a.s) includes an internal control to monitor the amplification process. The RNA Ultrasense™ One-step qRT-PCR system kit (Invitrogen) does not include an internal control, but a positive control, the cell culture adapted HAV strain HM175 43c was included in each run to monitor the amplification process. The conditions for the assay were different for each kit. For the hepatitisA@ceeramTools™ (Environmental and clinical) Kit (Ceeram s.a.s) the cycling conditions of the assay were as follows: 50°C for 1 h, for the reverse transcription reaction, followed by a hot start of 95°C for 10 min and 45 cycles of 95°C for 15 s, 60°C for 1 min and 70°C for 1 min (Costafreda *et al*., 2006). The cycling conditions of the assay for the RNA Ultrasense™ One-step qRT-PCR system kit (Invitrogen) were as follows: 50°C for 45 min, for the reverse transcription reaction, followed by a hot start of 95°C for 15 min and 50 cycles of 95°C for 15 s for denaturation, 60°C for 1 min for annealing and 65°C for 1 min for extension (Netshikweta, 2011).

2.2.6 Hepatitis A virus characterisation

2.2.6.1 Complementary DNA (cDNA) synthesis

Using random hexamers primers (Roche Diagnostics), cDNA (20 µL) was synthesised from 10 µL of RNA extracts from samples that tested positive for HAV. The synthesis was performed using a Thermo Scientific RevertAid™ Premium Reverse Transcriptase kit (Thermo Scientific, Waltham, MA). The

protocol used was a slight modification of that of the manufacturer's instructions, as 50 U of the reverse transcriptase was used instead of 200 U.

2.2.6.2 Amplification of the VP1 region

The full length VP1 region (900 bp) was amplified from cDNA (5 µL or 1 µL) by conventional PCR, using primers published in Costa-Mattioli et al (2002) (Table 2.3).

Table 2.3: Sequences of primers used to amplify the VP1 gene (Costa-Mattioli *et al*., 2002; Nainan *et al*., 2006)

PCR	**Primer**	**Nucleotide Sequence (5'-3')**
First round	HAV1 (Forward)	gTT TTg CTC CTC TTT ATC ATg CTA Tg
	HAV2 (Reverse)	AgT CAC ACC TCT CCA ggA AAA CTT
Nested	2172P (Forward)	gCT CCT CTT TAT CAT gCT ATg gAT
	3125N (Reverse)	CCT gCA TTC TAT ATg ACT CT

The total volume of the reaction mix was 25 µL and contained the following: 5 x KAPATaq™ HotStart buffer (Kapa Biosystems, Cape Town, SA), 10 mM dNTPs, and 10 µM each of the forward and reverse primers, respectively, 25 mM $MgCl_2$, and 1.25 units of KAPATaq™ HotStart DNA polymerase (Kapa Biosystems). The conditions under which the VP1 region was amplified were: pre-denaturation at 94°C for 3 min, follow by 45 cycles of 94°C for 30 s, 50°C for 30 s, 72°C for 1 min 30 s, and, final extension at 72°C for 5 min. If no positive amplicons were obtained after the first round of PCR, a nested PCR was performed and the VP1 gene was amplified using internal primers (Costa-Mattioli *et al*. 2002; Nainan *et al*. 2006) (Table 2.3) and 0.5 µL of the first PCR reaction as template. The assay conditions for the second round of PCR are essentially the same, except for the annealing temperature that decreased from 50°C to 48°C and the number of cycles that went from 45 to 35.

2.2.6.3 Amplification of the VP1/P2B junction

The VP1/P2B junction (~350 bp), was amplified from the same cDNA (5 µL) used to amplify the VP1 region using primers published in Nainan et al (2006) (Table 2.4).

Table 2.4: Sequences of primers used to amplify the VP1/P2B junction (Nainan *et al*., 2006)

PCR	Primer	Nucleotide Sequence (5'-3')
First round	2870P (Forward) 3381N (Reverse)	gAC AgA TTC TAC ATT Tgg ATT ggT CCA TTT CAA gAg TCC ACA CAC T
Nested	2896P (Forward) 3289N (Reverse)	CTA TTC AgA TTg CAA ATA CAA T AAC TTC ATT ATT TCA TgC TCC T

The total volume of the reaction mix was 25 µL and contained the following: 5 x KAPATaq™ HotStart buffer (Kapa Biosystems), 10 mM dNTPs, and 10 µM each of the forward and reverse primers respectively, 25 mM $MgCl_2$, and 1.25 units of KAPATaq™ HotStart DNA polymerase (Kapa Biosystems). The conditions under which the VP1/P2B junction was amplified were: pre-denaturation at 94°C for 3 min, follow by 35 cycles of 94°C for 30 s, 45°C for 45 s, 72°C for 1 min, and, final extension at 72°C for 5 min. If no PCR products were visible after the first round of PCR, a nested PCR was performed using internal primers (Nainan *et al*. 2006) (Table 2.4) and 0.5 µL of the first PCR reaction as template. The conditions for the second round of VP1/P2B amplification are similar to the ones used for the first round PCR, except for the annealing temperature that increased from 45°C to 48°C.

2.2.6.4 Sequencing and cloning

Amplicons were analysed by 2% agarose gel (SeaKem® LE Agarose, Lonza, Rockland, ME) electrophoresis and visualised by ethidium bromide staining and UV illumination. Positive amplicons were purified directly or from the gel using the Zymogen DNA Clean & Concentrator-25™ Kit or the Zymoclean™ Gel DNA Recovery Kit (Zymo Research, Irvine, CA), respectively. The primers used to amplify the different regions were used to sequence positive amplicons directly after amplification, in both the forward and reverse direction, using the ABI Prism BigDye® Terminator v3.1 Cycle sequencing Kit on an ABI 3130 automated analyser (Applied Biosystems), as per manufacturer's instructions. In cases where more than one strain were detected in a sample, purified PCR products were cloned using the cloneJET™ PCR cloning kit (Fermentas Life Sciences, Burlington, Ontario, Canada) as per manufacturer's instructions.

Colony PCR was performed on a minimum of 10 randomly selected clones per sample, using pJet 1.2/blunt specific primers. The total volume of the reaction mix was 20 µL and included the following: 5 x reaction buffer, 10 mM dNTPs, 1.5 mM $MgCl_2$, 0.2 pMol/µL of the forward primer, 0.2 pMol/µL of the reverse primer and 0.2 µL of GoTaq® (Promega Corp., Madison, WI). The conditions of the assay were as follows: pre-denaturation at 95°C for 3 min, follow by 35 cycles of 94°C for 30 s, 60°C for 30 s, 72°C for 1 min 30 s, and, final extension at 72°C for 10 min. After colony PCR, approximately 10 clones with the correct insert were selected and sequenced using the pJet 1.2/blunt specific primers.

2.2.6.5 Nucleotide sequence analyses

Basic sequence manipulation and verification were performed by using Sequencher™ v4.10.1 (Gene Codes Corporation, Ann Arbor, MI) and BioEdit Sequence Alignment Editor (v6.0.5). The identity of the edited sequences was verified in GenBank by using the Basic Local Alignment Search Tool (BLAST-n) program (Altschul *et al*., 1997). Multiple alignments of the verified nucleic acid sequences together with reference sequences for HAV and closely matched sequences from the output BLAST search were created using MAFFT version 7.110 (http://mafft.cbrc.jp/alignment/server/). After alignments, nucleic acid sequences were translated into protein using BioEdit Sequence Alignment Editor. Pairwise comparisons of nucleotides sequences of HAV reference types retrieved from Genbank and top hit BLAST search (Table 2.5) with the characterised strains were performed by using the two-parameter model of Kimura in Mega 5.10.

A group of HAV strains with ≥85% or >92.5% nucleotide sequence identity over the VP1/P2B junction were classified within the same genotype or subgenotype, respectively as was defined by Robertson et al. (1992). A group of HAV strains with nucleotide sequence identity ranging from 76.5% to 89.4% or >88.9% were classified within the same genotype or subgenotype, respectively as was defined by Costa-Mattioli et al., (2002).

Table 2.5: Hepatitis A virus strains used in phylogenetic analysis.

Genotype	Strain	Genbank Accession no.	Country of origin	Genomic region
IB	HM175	M14707	Australia	VP1, VP1/P2B
	CFH-HAV	HQ246217	Sierra Leone	VP1, VP1/P2B
	HAF 203	AF268396	Brazil	VP1, VP1/P2B
	L-A-1	AF314208	China	VP1, VP1/P2B
	BCN02	DQ504423	Spain	VP1
	BCN60	HQ401237	Spain	VP1
	BCN147	HQ401265	Spain	VP1
	BCN152	HQ401266	Spain	VP1
	MBB	M20273	North Africa	VP1, VP1/P2B
	JVR	U68692	SA	VP1, VP1/P2B
	126	EU416264	Germany	VP1/P2B
	166	EU416265	Germany	VP1/P2B
	034	EU416266	Germany	VP1/P2B
	-	EU930199	Germany	VP1/P2B
	-	DQ163904	Hungary	VP1/P2B
	-	EF190998	Hungary	VP1/P2B
	-	KC876797	Denmark	VP1/P2B
	SA-2009-Tom[a]	-	SA	VP1/P2B
	SA-2009-Let[a]	-	SA	VP1/P2B
IC	BCN70	HQ401240	Spain	VP1
	Shellfish08-106	HQ401253	Spain	VP1
IA	GBM	X75215	Germany	VP1, VP1/P2B
	HAV5	EU131373	Uruguay	VP1, VP1/P2B
	VDM	U66489	SA	VP1, VP1/P2B
	LU38/WT	AF357222	China	VP1, VP1/P2B
	H2	EF406357	China	VP1, VP1/P2B
	M2	AY974170	Cuba	VP1, VP1/P2B
	BCN31	HQ401230	Spain	VP1
	BCN143	HQ401264	Spain	VP1
	MSM08-09-StrainE	HQ401258	Spain	VP1
	MSM08-09-219	HQ401259	Spain	VP1
	923200[b]	-	SA	VP1/P2B
	923359[b]	-	SA	VP1, VP1/P2B
IIA	CF53/Berne	AY644676	Germany	VP1, VP1/P2B
IIB	SLF88	AY644670	Sierra Leone	VP1, VP1/P2B
IIIA	Sim27	FJ227135	India	VP1, VP1/P2B
IIIB	HA-JNG06-90F	AB258387	Japan	VP1, VP1/P2B
IV	Cy145	M59286	Philippines	VP1, VP1/P2B
V	AGM-27	D00924	Kenya	VP1, VP1/P2B

a: Netshikweta, 2011

b: Taylor, 1997

2.2.6.6 Phylogenetic analysis

Phylogenetic trees were constructed from the nucleic acid sequence alignments by using the Kimura two-parameter model from the neighbor-joining method and validated by 1000 bootstrap replicates. The genotype of each strain was determined by how well it clustered with reference sequences on the phylogenetic tree as well as nucleotide identity percentages.

2.3 RESULTS

2.3.1 Microbial indicator analysis

The irrigation water samples (59% [17/29]) had *E. coli* counts between 2 and 60 000 cfu/mL. The two borehole water samples tested negative for *E. coli* (Table 2.7).

2.3.2 Hepatitis A virus detection

2.3.2.1 Viral detection in clinical specimens

A total of 117 clinical specimens were collected (Table 2.6): 19 of which were collected form WC between 1997 and 2004; the remaining 108 specimens were collected from SA and non-SA regions between October 2012 and August 2013. Hepatitis A virus was detected in 81% (86/106) of the SA specimens and in 45% (5/11) of the non-SA specimens.

2.3.2.2 Viral detection in water samples

Hepatitis A virus was detected in 62% (18/29) of the irrigation water samples (Table 2.7). Mengovirus was detected in 55% (16/29) of the irrigation water samples. Of these 16 samples, nine tested positive for both HAV and mengovirus and in the remaining seven samples the process control was detected in the absence of HAV. Mengovirus was not detected in 45% (13/29) of the irrigation water samples of which 9/13 tested positive for HAV.

Table 2.6: Summary of results of HAV detection in clinical specimens

Geographical region	Age range	Total specimen collected	HAV detection
Gauteng	2 - 63	40	35 (78%)
Mpumalanga	3 - 23	4	3 (75%)
North West	4 - 25	6	6 (100%)
KwaZulu-Natal	29	1	1 (100%)
Western Cape	1 - 41	51	38 (75%)
Limpopo	5 - 35	3	2 (67%)
Eastern Cape	2	1	1 (100%)
SA-Total	**-**	**106**	**86 (81%)**
Kenya	4 - 28	6	2 (33%)
Mozambique	16	1	0
Swaziland	24 – 25	2	2 (100%)
Tanzania	26	1	1 (100%)
Zambia	-	1	0
Non SA-Total	**-**	**11**	**5 (45%)**
Total	**-**	**117**	**91 (78%)**

Hepatitis A virus was detected in 37% (19/51) of the wastewater samples (Table 2.8). The highest percentage of HAV detection was among wastewater samples from GP (56%), while the lowest detection rate was samples from KZN (0%) and WC (0%) (Table 2.8).

2.3.2.3 Viral detection on fresh produce

A total of 20 fresh produce samples were collected at pre- (*n=6*) and post-harvest (*n=14*) points. Hepatitis A virus was not detected on any of the fresh produce sample, while all of these samples tested positive for mengovirus.

Table 2.7: Summary of information on sampling site, results of microbial indicator and viral analysis of irrigation water samples

Site no	Date	Sample source	pH/ Temp	*E. coli* cfu/ 100 mL	Real-time RT-PCR analysis	
					HAV	Mengovirus
1	2012-01-31[c]	Bloed River	7/22	ND	+	-
	2012-08-13		7.5/13.5	344	-	-
5	2012-01-31[a]	River water below the Polokwane sewage works	ND	ND	+	+
	2012-08-13[b]		7.5/13.5	60000	-	+
7	2012-01-31[a]	River water on fresh produce farm	7/22	ND	+	+
	2012-03-27[a]		7.5/16	61	+	+
	2012-05-16[c]		7.5/11	1600	+	-
	2012-06-27[c]		7.5/11	600	+	-
	2012-08-13[c]		7.5/10.5	5100	+	-
9	2012-01-31[b]	Borehole water from cement dam A at farms offices	7/22	ND	-	+
	2012-04-24[a]		7/13	ND	+	+
	2012-08-13		7.5/13	0	-	-
16B	2012-01-31[b]	Dam B at pump site on fresh produce farm	7/22	ND	-	+
	2012-08-13[c]		7.5/12	8	+	-
17B	2012-01-31[b]	Dam C at pump site on fresh produce farm	7/22	ND	-	+
	2012-03-27[a]		7.5/16	180	+	+
	2012-05-16[c]		7.5/11.5	1150	+	-
	2012-06-27		7.5/11.5	20	-	-
	2012-08-13		8/12	150	-	-
23	2012-01-31[a]	River water below Seshego sewage works in the township	7/23	ND	+	+
	2012-08-13[c]		7.5/19	25600	+	-
Pivit	2012-03-27[a]	Irrigation pivot point A	7.5/16	66	+	+
	2012-05-16[c]		7.5/12	3600	+	-
KBBE	2012-04-24[b]	Borehole water from cement dam D	7/14	ND	-	+
KPL2	2012-04-24[a]		7/13	ND	+	+
KPL1	2012-06-27[c]	Irrigation pivot point B	7.5/13	2	+	-
KZN	2013-05-30[a]	Dam on cabbage farm	7/13	24	+	+
GP	2013-08-22[b]	Borehole water on cabbage farm	7/24	0	-	+
	2013-08-22[b]	River water on cabbage	7/18	40	-	+

a: samples that tested positive for both HAV and mengovirus
b: samples that tested negative for HAV but positive for mengovirus
c: samples that tested positive for HAV but negative for mengovirus
ND: not done. Samples arrived >24 h post collection.

Table 2.8: Summary of results of HAV detection in wastewater samples by province

Province	Total sample collected	HAV detection
Gauteng	9	5 (56%)
Mpumalanga	20	9 (45%)
North West	16	4 (25%)
KwaZulu-Natal	2	0
Western Cape	2	0
Unknown province	2	1 (50%)
Total	**51**	**19**

2.3.3 Hepatitis A virus characterisation

2.3.3.1 Viral characterisation in clinical specimens

Hepatitis A virus strains were genotyped from 67 HAV positive specimens. A total of 67 VP1 sequences and 64 VP1/P2B sequences were analysed. All the strains typed from clinical specimens clustered within genotype IB with >92.5% i.e. 96.5% and 95.8% nucleotide sequence identity to subgenotype IB type strain HM175 over the VP1 region and the VP1/P2B junction, respectively. The sequenced strains form unique SA clusters (SA major cluster) supported by high bootstrap values on both trees (Figures 2.2 [100%]; Figure 2.3 [98%], respectively).

Hepatitis A virus strains, grouping outside the ‘SA major cluster’, are strains characterised from specimens from SA and outside SA (Figure 2.2). The 'SA and non-SA cluster', supported by 78% bootstrap value, includes strains from the WC and GP regions as well as Tanzania. The strain characterised from a Kenyan specimen group outside ‘SA major cluster’ and ‘SA and non SA cluster’ but still within genotype IB.

2.3.3.2 Characterisation of viral strains from water samples

Hepatitis A virus strains from 41 water samples (15 irrigation, 15 wastewater and 11 surface) were genotyped. The VP1 region could be amplified from RNA extracted from 26 water samples (14 irrigation, 6 wastewater and 6 surface)

while the VP1/P2B junction could be amplified from RNA extracted from 33 water samples (11 irrigation, 14 wastewater and 8 surface). Based on phylogenetic analysis of the VP1 genomic region and the VP1/P2B junction, all strains, except one, clustered within genotype IB (Figures 2.4, 2.5). Pairwise analyses indicated that the strains that grouped within genotype IB had a 98.1% and 95.5% nucleotide sequence identity to the type strain HM175 over the VP1 region and the VP1/P2B junction, respectively. Unique SA clusters (SA major cluster) supported by high bootstrap values (Figures 2.4 [100%]; Figure 2.5 [97%], respectively) were observed on both trees.

Hepatitis A virus strains characterised from wastewaters in MP and irrigation water in LP grouped outside the unique SA cluster with 78% bootstrap value (Figure 2.5; [SA minor cluster]), but still within subgenotype IB. The strains characterised from samples collected outside SA also cluster within genotype IB, but outside 'SA major cluster' and 'SA minor cluster'.

The clones from the HAV strain characterised from the sample KZN_20130530 form a cluster (KZN_20130530 cluster) independent from all the HAV genotypes and subgenotypes present on both trees (Figures 2.4; Figure 2.5). The ‘KZN_20130530’ cluster seemed however closely related to genotype V as shown by the 95% and 94% bootstrap values on both trees. Pairwise analysis indicated that the strain has a 81.9% and 79% nucleotide sequence identity to the genotype V type strain AGM27 over the VP1 region and the VP1/P2B junction, respectively.

2.3.3.3 Genetic relatedness of hepatitis A virus from clinical and water sources

Combined phylogenetic analysis clustered all HAV strains (except one irrigation water sample) from clinical and environmental sources within subgenotype IB (Figures 2.6 and 2.7). None of the strains characterised in this study clustered with the strains representing the putative subgenotype IC (Figure 2.6). Within the 'SA major cluster' (Figure 2.6), clinical specimens from the WC group together (WC clusters; strains indicated with ◊) and not with any of the strains characterised from either irrigation water, or surface water, or wastewater

samples collected from other provinces (Figure 2.6a). Hepatitis A virus detection in the sludge samples collected from the WC was unsuccessful. A clinical specimen collected from GP (GP_PT122S; Figure 2.6a) clustered with HAV strains characterised from the effluents of wastewater plants collected in the same province (96% bootstrap value).

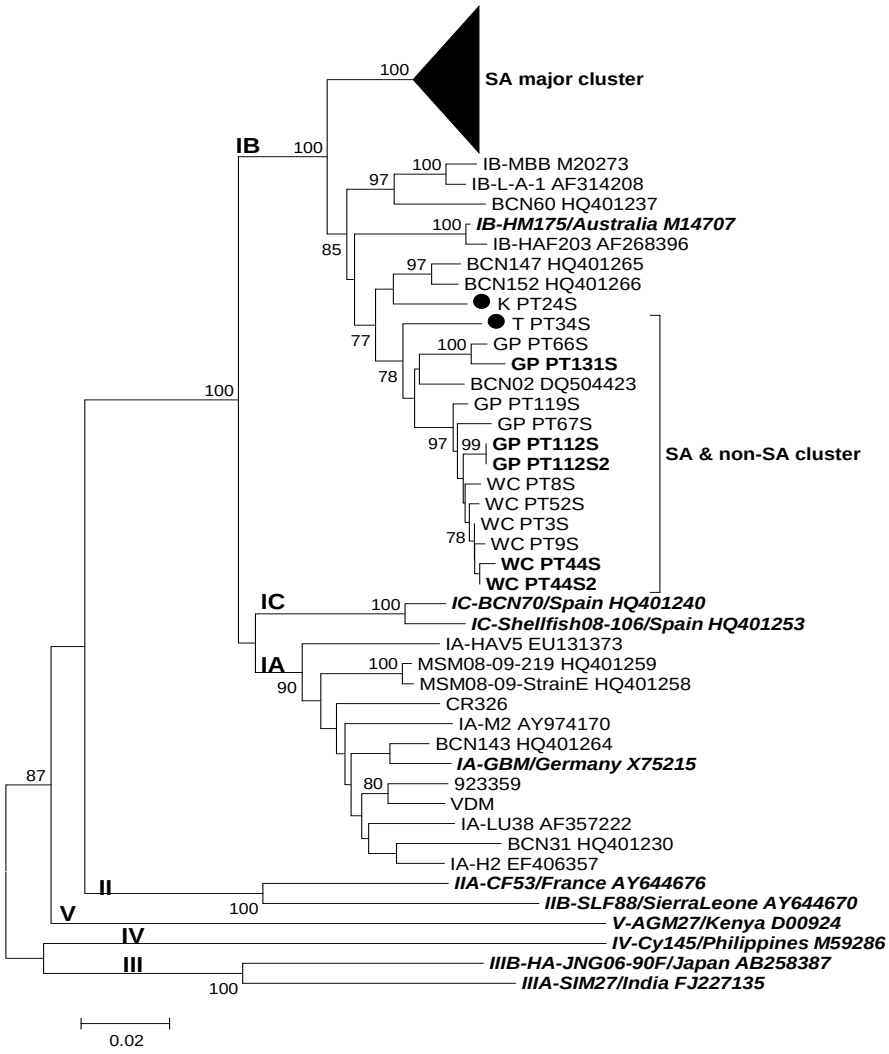

Figure 2.2: Phylogenetic analysis of hepatitis A virus strains characterised from clinical specimens based on the sequence of the VP1 region (900 bp). The tree was constructed using the Kimura two-parameter model of the neighbour-joining method, with 1000 bootstrap replicates; bootstrap values ≥75% are indicated at nodes. Strains indicated with (●) represent non SA strains typed in this study.Gauteng (GP); Western Cape (WC); Kenya (K); Tanzania (T). Reference strains or strains from previous studies that were retrieved from GenBank are present on the tree with the accession number next to the name. Type strains of each genotype or subgenotype are indicated in bold italics.

Figure 2.3: Phylogenetic analysis of hepatitis A virus strains characterised from clinical specimens based on the sequence of the VP1/P2B junction (~350 bp). The tree was constructed using the Kimura two-parameter model of the neighbour-joining method, with 1000 bootstrap replicates; bootstrap values ≥75% are indicated at nodes. Gauteng (GP); Western Cape (WC); Limpopo (LP). Reference strains or strains from previous studies that were retrieved from GenBank are present on the tree with the accession number next to the name. Type strains of each genotype or subgenotype are indicated in bold italics.

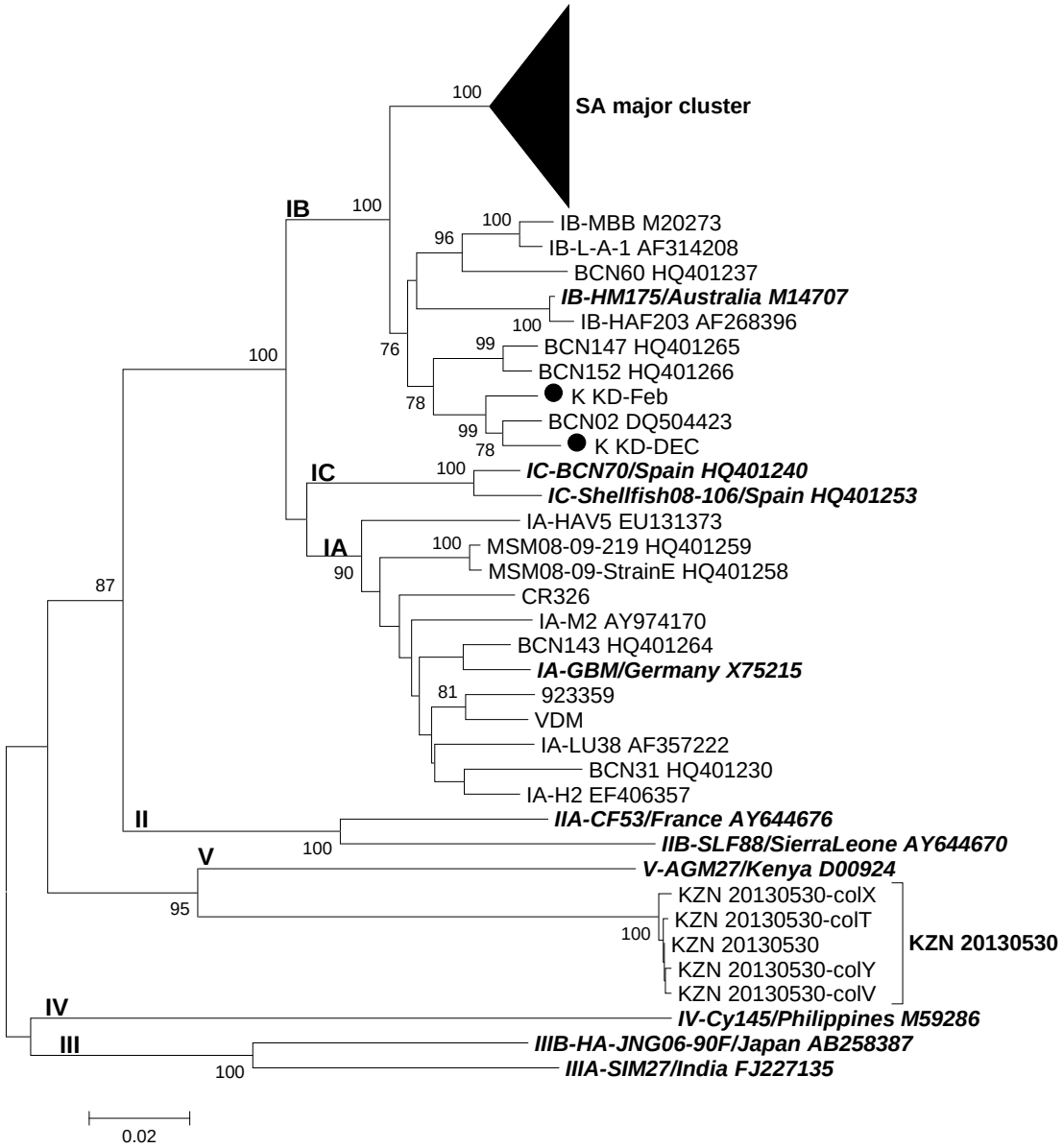

Figure 2.4: Phylogenetic analysis of hepatitis A virus strains characterised from water samples based on the sequence of the VP1 region (900 bp). The tree was constructed using the Kimura two-parameter model of the neighbour-joining method, with 1000 bootstrap replicates; bootstrap values ≥75% are indicated at nodes. Strains indicated with (●) represent non-SA strains typed in this study. KwaZulu-Natal (KZN); Kenya (K). Reference strains or strains from previous studies that were retrieved from GenBank are present on the tree with the accession number next to the name. Type strains of each genotype or subgenotype are indicated in bold italics. Strains with 'col' in their name represent clones.

Figure 2.5: Phylogenetic analysis of hepatitis A virus strains characterised from water samples based on the sequence of the VP1/P2B junction (~350 bp). The tree was constructed using the Kimura two-parameter model of the neighbour-joining method, with 1000 bootstrap replicates; bootstrap values ≥75% are indicated at nodes. . Strains indicated with (●) represent non-SA strains typed in this study. Kenya (K). Reference strains or strains from previous studies that were retrieved from GenBank are present on the tree with the accession number next to the name. Type strains of each genotype or subgenotype are indicated in bold italics. Strains with 'col' in their name represent clones.

The nucleotide sequence of HAV strains previously detected on tomatoes and lettuce (Netshikweta, 2011) and in German tourists (Faber *et al*., 2009) that presumably got infected while visiting SA, where included in the phylogenetic analysis of the VP1/P2B junction (Figure 2.7a). All these strains cluster within 'SA major cluster' (Figures 2.3, 2.5, 2.7). Within the 'SA major cluster' (Figure 2.7), the German HAV strains cluster with clinical strains from the Gauteng region (79% bootstrap value) (Figure 2.7a). In addition, the strain detected on lettuce clustered with the strain detected in a wastewater sample from MP (strain 10/1011) and with clinical strains from the WC (WC_PT49S and WC_PT54S) (87% bootstrap value; Figure 2.7a). Pairwise analysis of nucleotide sequences showed that the HAV strain from lettuce is 100% and 99.7% identical to the strains detected in wastewater from MP and in clinical specimens from the WC, respectively. The remaining clinical strains from the WC group together.

2.4 DISCUSSION

In this chapter, the prevalence and genetic characterisation of HAV in water samples and clinical specimens, collected mainly in SA, but also in neighbouring countries were assessed. Hepatitis A virus and mengovirus were detected in 62% (18/29) and 55% (16/29) of the irrigation water samples, respectively (Table 2.7). A total of 59% (17/29) of the irrigation water samples were faecally contaminated, as indicated by the *E. coli* count ranging from >1 to >1000 cfu/ 100 mL. Mengovirus was used as a process control to monitor the efficiency of recovery and nucleic acid extraction procedures. The process control was detected in seven samples, in the absence of HAV (Table 2.7), indicating that the processing and nucleic acid extraction procedures were adequate and the negative HAV result was therefore a true negative. Nine of the samples from which mengovirus detection failed (13/29), tested positive for HAV still indicating that the method for recovery and detection was appropriate. However, there were four samples that tested negative for both mengovirus and HAV. The absence of mengovirus in these samples could indicate false negative result for HAV. The average recovery efficiency of mengovirus and

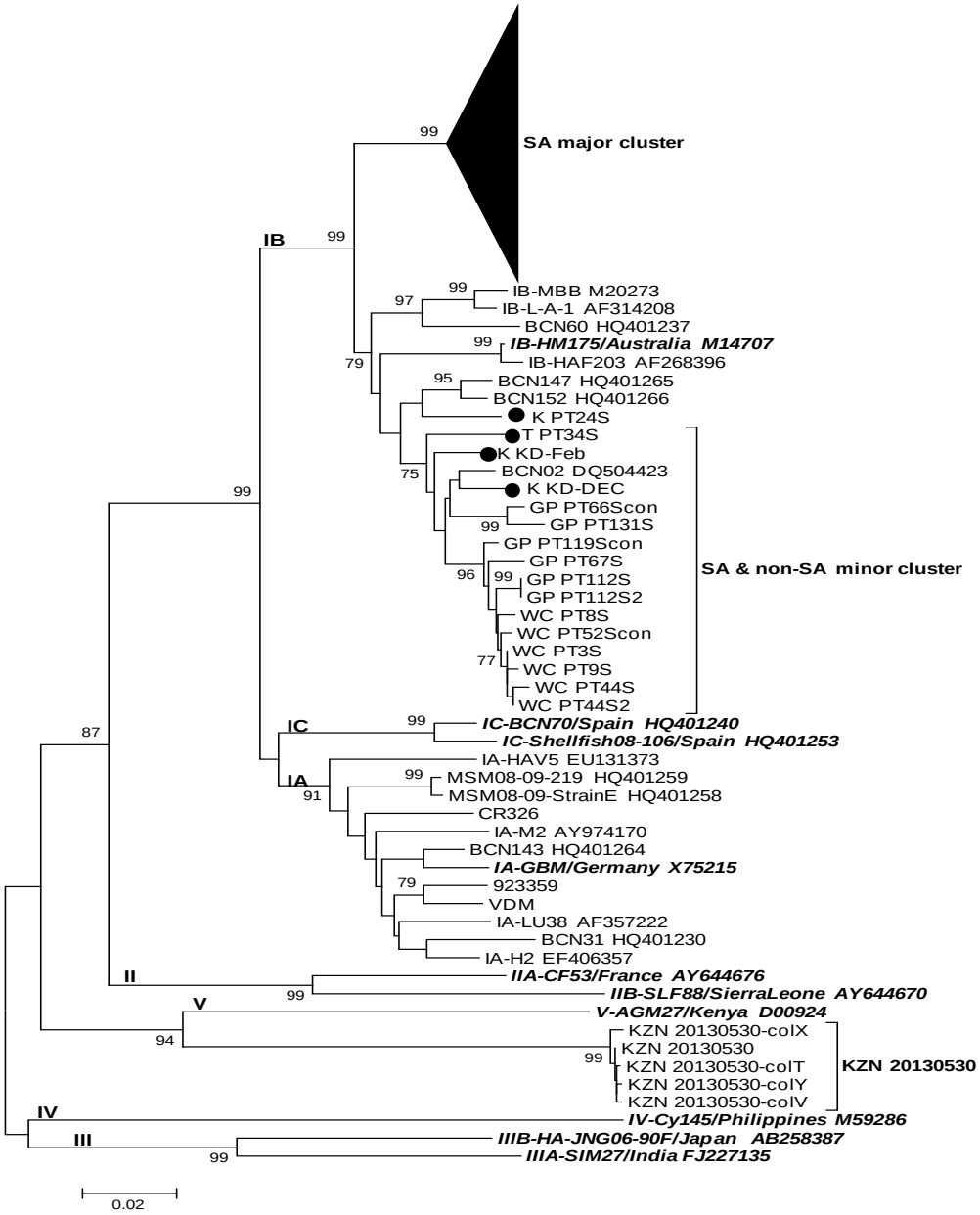

Figure 2.6: Phylogenetic analysis of hepatitis A virus strains characterised from clinical and water sources based on the sequence of the VP1 region (900 bp). The tree was constructed using the Kimura two-parameter model of the neighbour-joining method, with 1000 bootstrap replicates; bootstrap values ≥75% are indicated at nodes. Strains indicated with (●) represent non-SA strains typed in this study. KwaZulu-Natal (KZN); Kenya (K); Tanzania (T). Reference strains or strains from previous studies that were retrieved from GenBank are present on the tree with the accession number next to the name. Type strains of each genotype or subgenotype are indicated in bold italics. Strains with 'col' in their name represent clones.

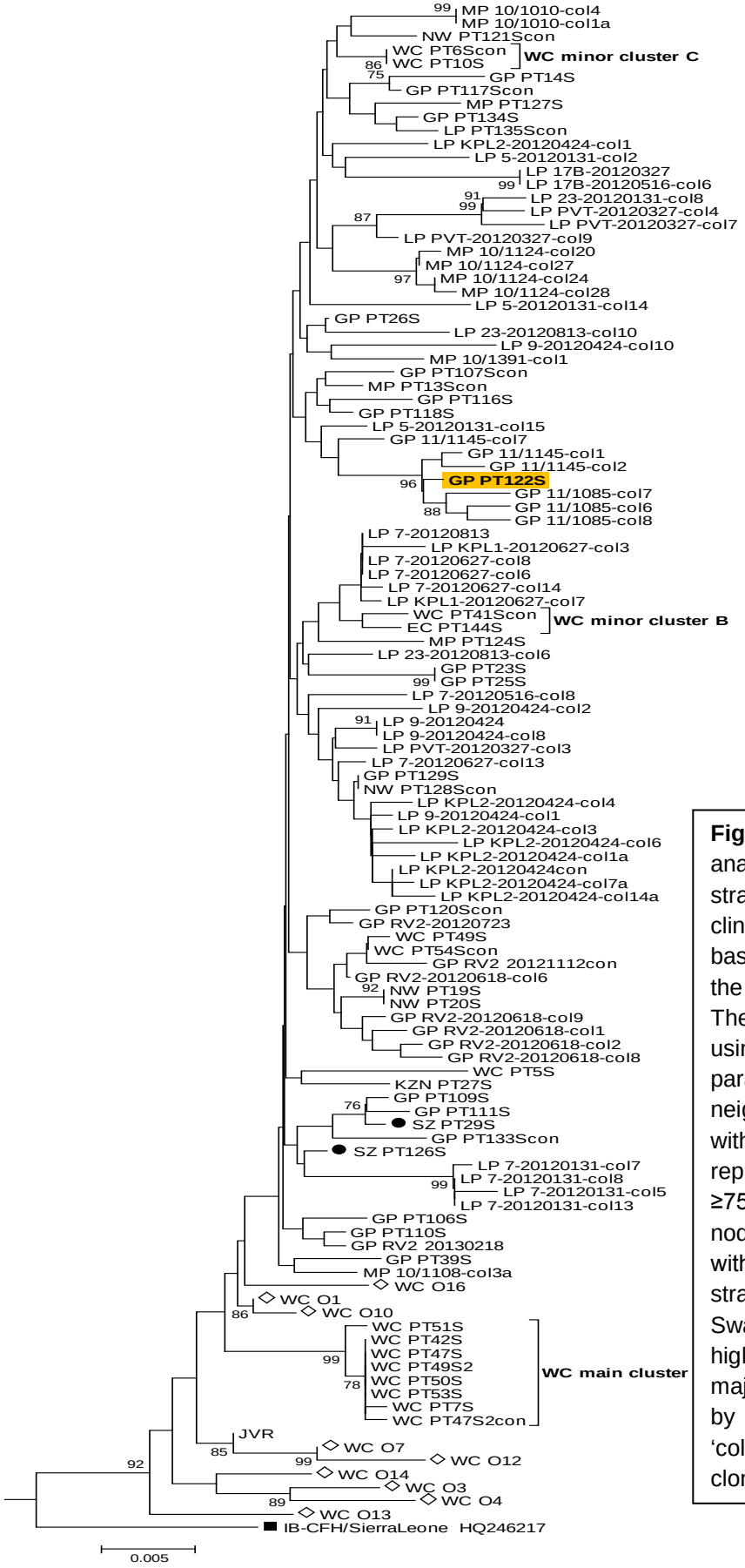

Figure 2.6a: Phylogenetic analysis of hepatitis A virus strains characterised from clinical and water sources based on the sequence of the VP1 region (900 bp). The tree was constructed using the Kimura two-parameter model of the neighbour-joining method, with 1000 bootstrap replicates; bootstrap values ≥75% are indicated at nodes. Strains indicated with (●) represent non-SA strains typed in this study. Swaziland (SZ). The tree highlights strains of 'SA major cluster' as depicted by Figure 2.6. Strains with 'col' in their name represent clones.

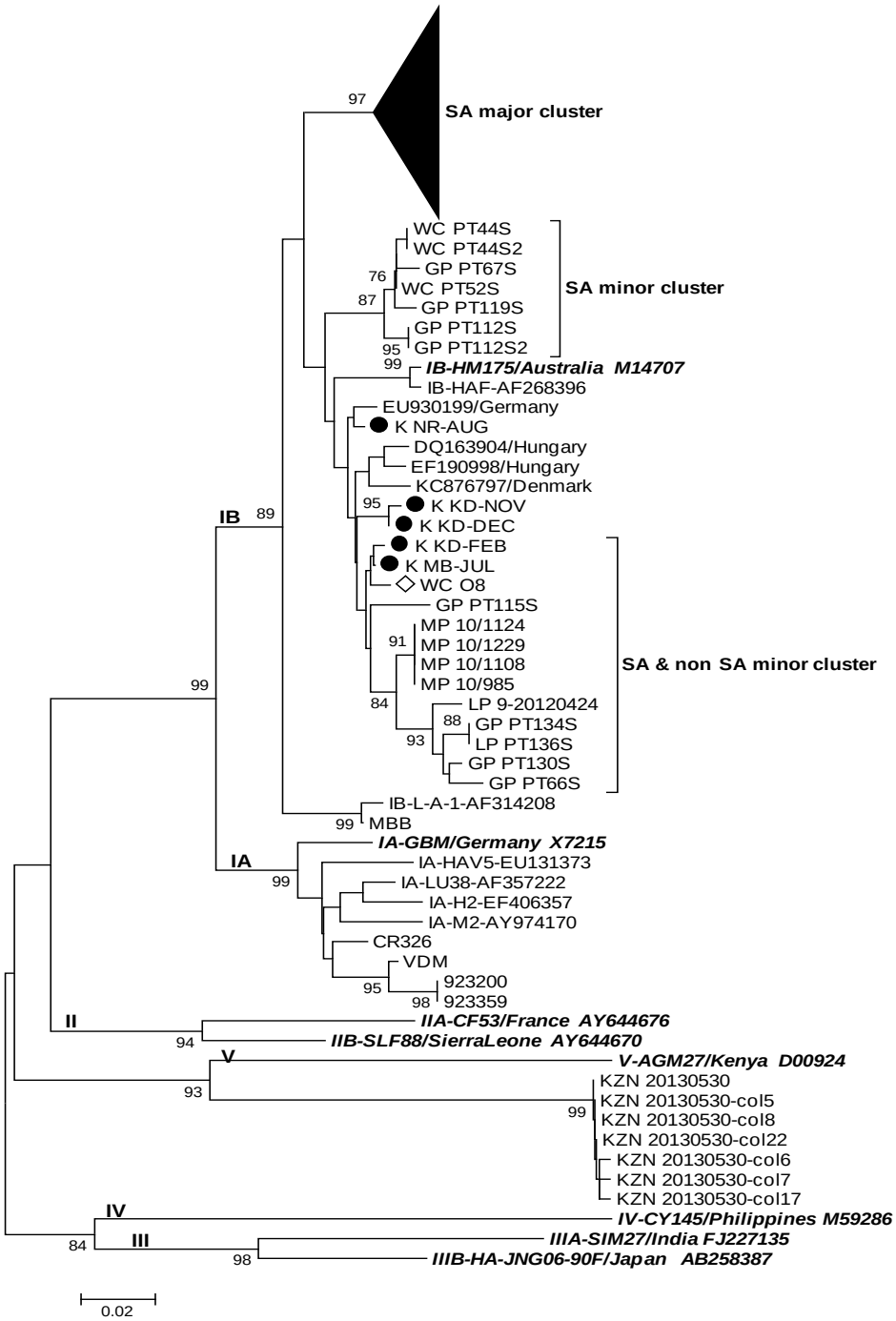

Figure 2.7: Phylogenetic analysis of hepatitis A virus strains characterised from clinical and water sources based on the sequence of the VP1/P2B junction (~350 bp). The tree was constructed using the Kimura two-parameter model of the neighbour-joining method, with 1000 bootstrap replicates; bootstrap values ≥75% are indicated at nodes. Strains indicated with (●) represent non SA strains typed in this study. KwaZulu-Natal (KZN); Kenya (K). Reference strains or strains from previous studies that were retrieved from GenBank are present on the tree with the accession number next to the name. Type strains of each genotype or subgenotype are indicated in bold italics. Strains with 'col' in their name represent clones.

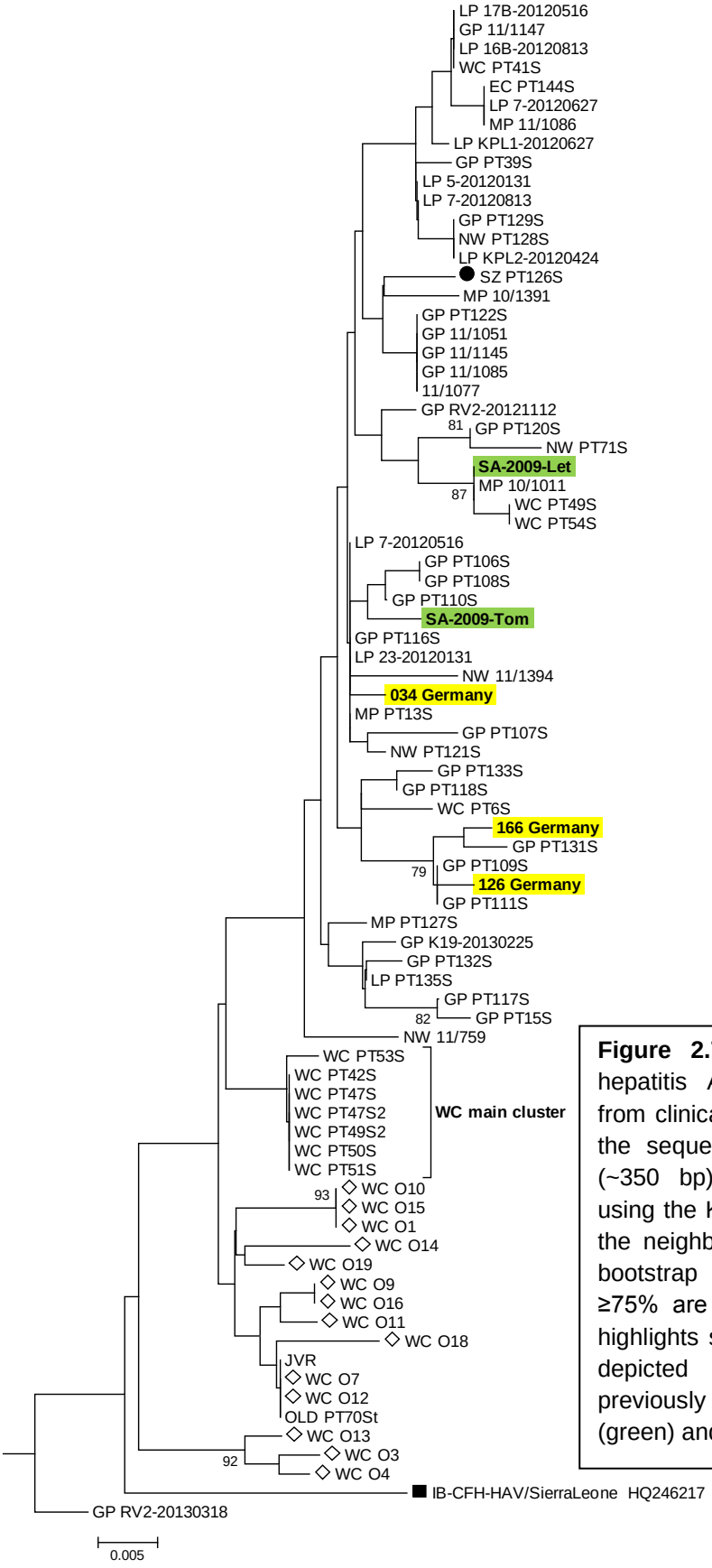

Figure 2.7a: Phylogenetic analysis of hepatitis A virus strains characterised from clinical and water sources based on the sequence of the VP1/P2B junction (~350 bp). The tree was constructed using the Kimura two-parameter model of the neighbour-joining method, with 1000 bootstrap replicates; bootstrap values ≥75% are indicated at nodes. The tree highlights strains of 'SA major cluster' as depicted by Figure 2.7 and strains previously detected on fresh produce (green) and in German tourists (yellow).

HAV from turbid water using the glass wool adsorption-elution method is <1% and 5%, respectively (V. Ruhanya, unpublished data). The detection rate of HAV could therefore be an underestimation of the true prevalence of the virus in these irrigation waters.

The fresh produce samples analysed in this study tested negative for HAV, even though water used for irrigation was contaminated with HAV. Failure to detect viruses in food matrices might be due to the low efficiency of recovery and nucleic acid extraction procedures or the presence of inhibitors to the detection assay (Bosch *et al*., 2011). However, mengovirus was detected in all food samples, indicating that recovery, concentration and nucleic acid extraction procedures were efficient. It is unlikely that inhibitors to the detection assays could have accounted for the negative results since the internal amplification control (IAC) of the RT-PCR reactions were detected. Another possible explanation for negative HAV results on the analysis of fresh produce is that the sample size was very small and therefore not representative. Given that enteric viruses are present in low numbers in the environment (Mattison and Bidawid, 2009; Bosch *et al*., 2011) a small sample size is a limiting factor during analysis of fresh produce. In a previous study 72 fresh produce samples were analysed and HAV was detected in only four of the samples (Netshikweta, 2011). Since HAV was detected in irrigation water more fresh produce should be analysed to confirm the negative results obtained.

Hepatitis A virus was detected in 37% (19/51) of the wastewater samples. A number of factors could account for a low detection rate in wastewater samples as opposed to the irrigation water samples. No prior recovery was performed on the wastewater samples. Inhibitors, which were not completely removed during nucleic acid extraction, could have been co-concentrated during the PEG_{8000}/NaCl precipitation step. Furthermore, the assay used for the wastewater samples lacked an IAC. To circumvent the effect of inhibitors, the assay was repeated with diluted nucleic acid. Noroviruses and sapoviruses were detected in the 19 wastewater samples that tested positive for HAV and in six of the remaining 32 samples (Murray *et al*., 2013). This indicates that the

nucleic acid extraction was efficient and that HAV if present was possibly at levels below the detection limit of the assay.

Hepatitis A virus strain was genotyped from 67 clinical specimens and 33 water samples. All strains characterised from clinical and environmental sources collected in SA form a unique lineage within genotype IB except for strains from the irrigation water sample KZN_20130530. In 1997, it was established that both genotype IA and IB circulated in SA with IB being more prevalent (Robertson *et al*., 1992; Taylor, 1997). When determining the predominance of HAV strains, the presence of more than one genotype or subgenotype could be biased by sampling collection strategies. Predominant genotype or subgenotype might differ from one sample type to another (Vaughan *et al*., 2013a). During the investigation of a hepatitis A outbreak in a day-care centre in Rio de Janeiro, Brazil, both subgenotypes IA and IB were detected, with IA only found in serum specimens and IB in the water sample and saliva of few patients (Amado *et al*., 2011). In a separate study carried out in the same geographical region, it was found that only HAV subgenotype IA is present in raw and treated samples (Prado *et al*., 2012), while in a third study, both IA and IB were equally detected among sewage samples (Villar *et al*., 2007). These findings suggested that different HAV strains might survive differently in the environment and that sampling strategies should include different types of samples. In this study, samples were collected from sera of patients and a wide range of water sources, namely surface water (rivers, dams, and irrigation pivot), groundwater, effluents and sludge of wastewater plants. None, except one, of the SA strains sequenced in this study grouped within a genotype other than IB suggesting that since 1997 genotype IB still predominates in SA. However, additional HAV strains will have to be characterised to ascertain if genotype IA is still present in the country.

There was no apparent grouping of strains according to SA regions except for strains detected in specimens collected from the WC. The WC is a tourist destination that also has many migrants from other African regions. Interaction

between imported and SA HAV strains might have given rise to a unique sublineage of the SA lineage fit to circulate in the WC (Wang *et al*., 2013).

Hepatitis A virus strains were genotyped from clinical specimens collected in Kenya, Tanzania and Swaziland. The strains characterised from Kenya and Tanzania clustered within genotype IB, but outside 'SA major cluster' (Figures 2.2, 2.4, 2.5, 2.6, 2.7). The strains from Swaziland however, clustered within 'SA major cluster' with strains from GP (76% bootstrap value, Figure 2.6a). In 2009, a study was published on the epidemiology of hepatitis A in Germany, with special emphasis on imported and non-imported HAV infections (Faber *et al*., 2009). The study detected and characterised HAV in tourists that presumably got infected while visiting SA. These strains formed a distinct cluster when compared to strains detected in individuals who got infected either in Germany or in other regions of the world. The strains characterised from the tourists were included in this study. Phylogenetic analysis clustered the German strains within 'SA major cluster', more particularly with HAV strains detected in clinical specimens from GP (79% bootstrap value, Figure 2.7a). The presence of a unique SA lineage has great epidemiological significance. This indicates that nucleotide sequence analysis of the HAV genomic regions (VP1 and VP1/P2B) can be used for molecular source tracking purposes, to identify potential infection sources and/or transmission routes of HAV and to distinguish imported from local hepatitis A cases that could possibly be differentiated by region.

Hepatitis A virus is an enteric virus transmitted either directly from person-to-person or indirectly through the ingestion of contaminated food and water (Nainan *et al*., 2006; Pinto *et al*., 2010, 2012; Rodríguez-Lázaro *et al*., 2012). Phylogenetic analysis clustered the majority of clinical and environmental strains genotyped in this study together (Figures 2.6 and 2.7). Thus, HAV strains detected from water sources appear to be clinically relevant. Hepatitis A virus strains were previously detected on tomatoes and lettuce at the point of retail in the Tshwane district, GP (Netshikweta, 2011). The detected strains also clustered within 'SA major cluster' (Figure 2.7a). Of particular importance

is the HAV strain detected on lettuce, which is identical (100% nucleotide identity) to a strain detected in a wastewater sample collected in MP, and closely related to clinical strains from the WC (99.7% nucleotide identity) (Figure 2.7a). Fresh produce can be contaminated at various points throughout the chain of production. Irrigation with surface water contaminated with faeces containing HAV and agrichemicals mixed with contaminated water are possible sources of contamination (Fiore, 2004; Berger *et al*., 2010; Rodríguez-Lázaro *et al*., 2012). The lack of proper washing facilities for food handlers on the farm, at commercial outlets or during preparation are additional sources of contamination (Fiore, 2004; Sun *et al*., 2012). In the present case, the exact source(s) of contamination of the lettuce could not be defined, but the identity to a strain detected in sewage effluent suggests that sewage treatment in MP may not remove all HAV particles. Wastewater treatment in GP may also be inefficient since the HAV strain genotyped from a clinical specimen from GP clustered with HAV strains characterised from the effluents of wastewater plants collected in the same province (96% bootstrap value, Figure2.6a). Two sludge samples were collected from the WC, but no HAV was detected. Unfortunately, no additional water samples were collected from the WC as this could have shed more light on the circulation of HAV in the region.

In summary, HAV strains of subgenotype IB circulate in SA as indicated by the analyses of clinical and environmental sources. The SA HAV strains have unique VP1 and VP1/P2B genomic regions making molecular source tracking possible. The presence of genotype IB in the surface water sources used to irrigate fresh produce indicates human faecal contamination. Therefore, faecally-contaminated water sources are a potential vehicle of transmission of HAV infection and a potential source of contamination of irrigated fresh produce in SA.

CHAPTER 3

GENETIC DIVERSITY OF HEPATITIS A VIRUS IN SOUTH AFRICA

3.1 INTRODUCTION

The genome of HAV has one long ORF of 2 227 amino acids organised into P1, P2 and P3 regions (Wang *et al*., 2013). The capsid proteins VP1 to VP4 are encoded by P1. In the course of HAV infection, antibodies are directed against one conserved immunodominant neutralisation region (Costa-Mattioli *et al*, 2003; Knowles *et al*., 2012) with a number of amino acids from the VP1 and VP3 regions contributing to the immunodominant antigenic site (Pintó *et al*., 2012).

To date, only six genotypes of HAV have been identified. Classification of HAV strains might differ from one genomic region to another (Vaughan *et al*., 2013a). Each of the three genotypes found in humans (genotypes I, II and III) are subdivided into A and B. Candidate strains for a putative subgenotype IC have been isolated from Peru and Spain (Costa-Mattioli *et al*, 2002). Genotypes IV to VI are found in simians (Robertson *et al*., 1992; Costa-Mattioli *et al*, 2003). Hepatitis A virus strains infecting simians have a unique P1 sequence that distinguishes them from human strains (Brown *et al*., 1989; Nainan *et al*., 1991; Vaughan *et al*., 2013a). However, primates can be infected with human strains as owl monkeys infected with human HAV strains have previously been reported, (LeDuc *et al*., 1983; Theamboonlers *et al.*, 2012). It has also been possible to infect the chacma (Cape) baboon (*Papio ursinus*) with a cell culture adapted strain of human HAV (Taylor and Prozesky, 1985).

Structural constraints of the viral capsid and the usage of rare codons seem to prevent the emergence of new antigenic variants of HAV (Sánchez *et al*., 2003b; Aragonès *et al*., 2008). Despite a low antigenic variability, HAV replicates as a swarm of mutants known as quasispecies (Sánchez *et al*., 2003a; Pintó *et al*., 2010), the majority of which are negatively selected (Aragonès *et al*., 2008, 2010; Pintó *et al*., 2012; Vaughan *et al*., 2013a). Evidence of quasispecies within clinical isolates has been reported (Sánchez *et al*., 2003a; Sulbaran *et al*., 2010; Wang *et al*., 2013). Naturally occurring antigenic escape mutants, with amino acid change at the immunodominant site, have also been documented (Pérez-Sautu *et al*., 2011a,b).

The aim of this aspect of the study was to assess the genetic diversity among HAV strains detected in clinical and environmental sources (Chapter 2). The characterised strains will be screened for the presence of novel and/or previously described antigenic variants, as well as quasispecies.

3.2 MATERIALS AND METHODS

3.2.1 Nucleotide sequence analyses

Pairwise analysis performed previously were repeated (section 2.2.6.5). Pairwise comparisons were performed on the amino acid sequences by using the p-distance method in MEGA 5.10. Nucleic acid sequences of clones from the sample KZN_20130530 were compared to reference strains from other genotypes and subgenotypes of HAV by pairwise analysis at both nucleotide and amino acid levels, using the two-parameter model of Kimura in MEGA 5.10.

3.2.2 Detection of novel or predicted mutations

Sequence alignments performed in MAFFT, for both the VP1 and VP1/P2B regions (section 2.2.6.5), were used to check for novel or predicted mutations at nucleotide and amino acid levels with the software BioEdit Sequence Alignment Editor (v6.0.5). Positions at which nucleotide mutations cause amino acid

change potentially resulting in vaccine escape mutants (Pérez-Sautu *et al*., 2011a; Pintó *et al*., 2012) were carefully analysed. The amino acid sequence of strains for which amino acid substitution was observed at or near the published neutralising sites was submitted to PHYRE^2 (Kelley and Sternberg, 2009) together with the sequence of HM175 (M14707) and published naturally occurring antigenic mutants (Table 2.5) for 2D protein structure prediction. Predicted protein structures were compared to each other and to HM175. Changes observed on the 2D prediction model were entered into LabViewer and a 3D model of the VP1 protein was computed.

3.2.3 Phylogenetic analysis

Phylogenetic analyses were performed to analyse strains that group differently over the VP1 and VP1/P2B genomic regions as described previously (section 2.2.6.6).

3.2.4 Cloning of clinical strains and quasispecies analysis

Hepatitis A virus strains characterised from five clinical specimens were cloned. Five clones were analysed per specimen by means of nucleotide sequence analysis (section 2.2.6.4). Nucleic acid sequence editing, verification and analysis were performed as previously described (section 2.2.6.5). Multiple alignments of the verified nucleic acid sequence, together with HM175 were created using MAFFT version 7.110 (http://mafft.cbrc.jp/alignment/server/). Phylogenetic trees were constructed from the nucleic acid sequence alignments by using the Kimura two-parameter model from the neighbour-joining method. The trees were assessed by bootstrap analysis (1000 pseudoreplicates).

3.3 RESULTS

3.3.1 Pairwise analysis

Genetic divergence, at the nucleotide and amino acid level, of all characterised strains was higher over the VP1/P2B junction (4.2% and 1.4% for clinical strains; 4.5% and 1.4% for environmental strains, respectively) when compared to the VP1 region (3.5% and 0.3% for clinical strains; 1.9 and 0.6% for environmental strains) (Table 3.1). No mixed infections were recorded during molecular analysis of clinical strains over the two genomic regions.

Table 3.1: Pairwise analysis (percentage of identities and divergence) results of HAV strains characterised from clinical specimens and water samples.

	VP1 region		VP1/P2B junction	
	Nucleotide	**Amino acid**	**Nucleotide**	**Amino acid**
Clinical specimens				
genetic identity	96.5%	99.7%	95.8%	98.6%
genetic divergence	3.5%	0.3%	4.2%	1.4%
Water samples				
genetic identity	98.1%	99.4%	95.5%	98.6%
genetic devergence	1.9%	0.6%	4.5%	1.4%

First row in each cell corresponds to genetic identities and second rows to genetic divergence.

When comparing KZN_20130530 to representatives of HAV genotypes and subgenotypes, the highest percentage of nucleotide and amino acid identity was seen with genotype V over both the VP1 (81.9% and 97.8%, respectively) and VP1/P2B (79% and 96.2%, respectively) regions (Table 3.2). Genetic identities were lowest with genotype IIB over the VP1 region (74.5% and 81.1%) and with genotype IIIA (70.4% and 84.6%) over the VP1/P2B region.

Table 3.2: Genetic distances (percentage of identities and divergence) between type strains of HAV genotypes I to VI and KZN_20130530.

	VP1 region		VP1/P2B junction	
	Nucleotide	**Amino acid**	**Nucleotide**	**Amino acid**
IB (HM175 – M14707)				
genetic identity	78%	94.5%	74.2%	89.1%
genetic divergence	22%	5.5%	25.8%	10.9%
IA (GBM – X75215)				
genetic identity	76.6%	93.1%	73.8%	87.3%
genetic divergence	23.4%	6.9%	26.2%	12.7%
IC[a] (BCN70 – Q401240)				
genetic identity	77.3%	94.5%		
genetic divergence	22.7%	5.5%	-	-
IIA (CF53/Berne – AY644676)				
genetic identity	75.5%	93.5%	73.2%	87.3%
genetic divergence	24.5%	6.5%	26.8%	12.7%
IIB (SLF88 – AY644670)				
genetic identity	74.5%	81.1%	72.2%	87.3%
genetic divergence	25.5%	18.9%	27.8%	12.7%
IIIA (Sim27 – FJ227135)				
genetic identity	74.6%	92.8%	70.4%	84.6%
genetic divergence	25.4%	7.2%	29.6%	15.4%
IIIB (HA-JNG06-90F - AB258387)				
genetic identity	77.8%	93.5%	74.6%	88.2%
genetic divergence	22.2%	6.5%	25.4%	11.8%
IV (Cy145 – M59286)				
genetic identity	76.2%	95.4%	66.4%	89.1%
genetic divergence	23.8%	4.6%	33.6%	10.9%
V (AGM27 – D00924)				
genetic identity	81.9%	97.8%	79%	96.2%
genetic divergence	18.1%	2.2%	21%	3.8%
VI[b] (JM55 – L07731)				
genetic identity			80.4%	95%
genetic divergence	-	-	19.6%	5%

a: nucleotide sequence of VP1/P2B junction not available

b: nucleotide sequence of VP1 region not available; only VP1/P2A sequence was available from GenBank.

3.3.2 Phylogenetic analysis

3.3.2.1 Phylogenetic analysis of HAV strains from clinical specimens

Phylogenetic analysis, of both the VP1 and VP1/P2B regions, was performed on the 67 HAV positive specimens (Figures 3.1, 3.3). Hepatitis A virus was characterised from two specimens collected from the same patient on four occasions: three in the Western Cape and once in Gauteng. Strains WC_PT47S and WC_PT47S2 (Figures 3.2, 3.4) clustered within 'WC main cluster' with minor differences between the two strains over the VP1 region only (WC_PT47S2 branch was longer than that for WC_PT47S) (Figure 3.2). Strain WC_PT49S2 also grouped within 'WC main cluster' but WC_PT49S collected from the same patient clustered with WC_PT54S within 'WC minor cluster B' (85% and bootsrap value) instead (Figures 3.2). Strains characterised from two specimens in Swaziland form the 'SZ cluster' within 'SA major cluster' (Figure 3.2). The SZ cluster also includes HAV strains from the GP region. The HAV strain that led to the unique SA lineage is closely related to a strain characterised from a chimpanzee in Sierra Leone (CFH-HAV indicated with ■; Figure 3.2) as this strain, together with the SA strains, form the 'SA major cluster' supported by 100% bootstrap value.

Strains WC_PT44S and WC_PT44S2 (Figures 3.1, 3.3), GP_PT112S and GP_PT112S2 (Figures 3.1, 3.3) are the two other instances where two specimens were collected from the same patient. The strains GP_PT112S and GP_PT112S2 are identical, but there seemed to be minor differences between WC_PT44S and WC_PT44S2 over the VP1 region (Figure 3.1).

Two additional clusters (SA minor clusters A and B) were seen on Figure 3.3. Strains WC_PT44S and WC_PT44S2, GP_PT112S and GP_PT112S2 clustered within 'SA minor clusters A' along with WC_PT52S, GP_PT67S and GP_PT119S as previously seen in 'SA and non-SA cluster' (Figure 3.1). Strains GP_PT131S and GP_PT134S (Figures 3.1, 3.3) showed different grouping from Figures 3.1, 3.3. The strain GP_PT131S clustered with the strain GP_PT66S (100% bootstrap value) outside 'SA major cluster' (Figure 3.1) while it clustered

with the German strain 166 within 'SA major cluster' (Figure 3.4). The opposite was seen with the strain GP_PT134S that clustered with LP_PT135S (79% bootstrap value) within 'SA major cluster' (Figure 3.2), while it clustered with LP_PT136S (92% bootstrap value) outside 'SA major cluster' (Figure 3.3). Strain LP_PT135S remains within 'SA major cluster'.

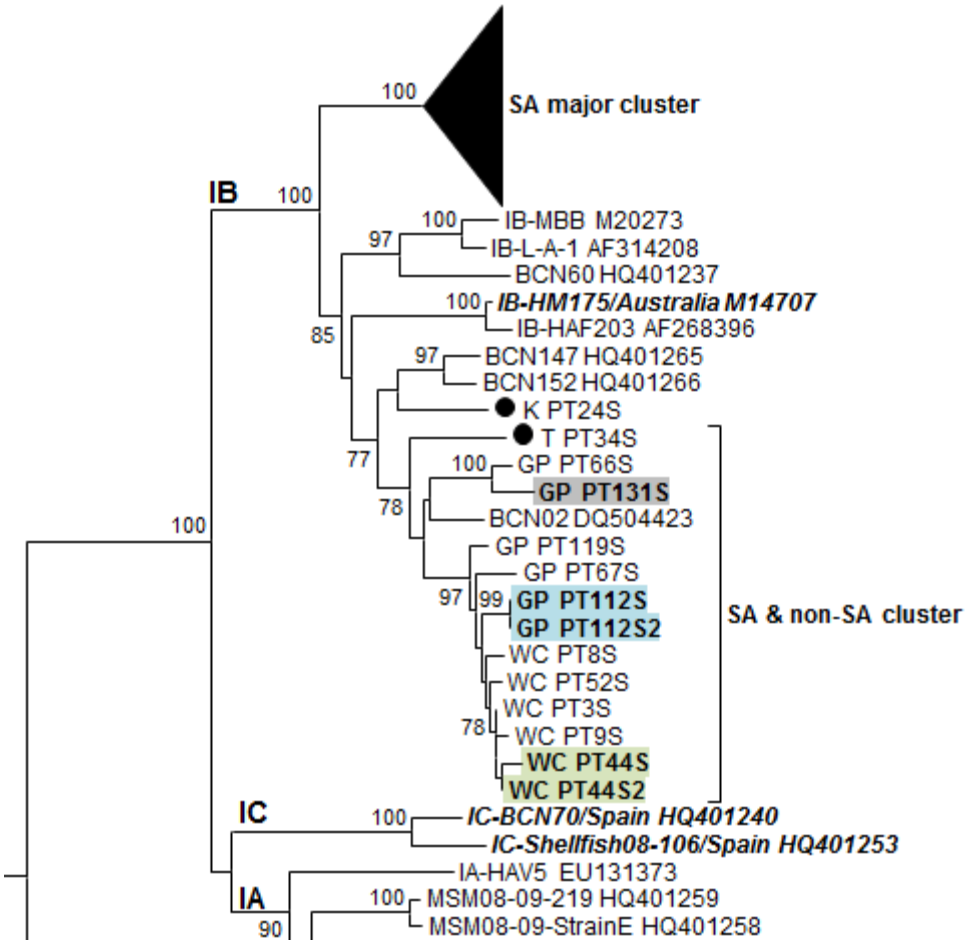

Figure 3.1: Phylogenetic analysis of HAV strains characterised from clinical specimens based on the sequence of the VP1 region (900 bp). The tree was constructed using the Kimura two-parameter model of the neighbour-joining method, with 1000 bootstrap replicates; bootstrap values ≥75% are indicated at nodes. Strains indicated with (●) represent non SA strains typed in this study. Kenya (K). The tree highlights the strains GP_PT131s (light grey), GP_PT112S and GP_PT112S2 (light blue), WC_PT44S and WC_PT44S2 (light green) of 'SA and non-SA cluster' as depicted by Figure 2.2.

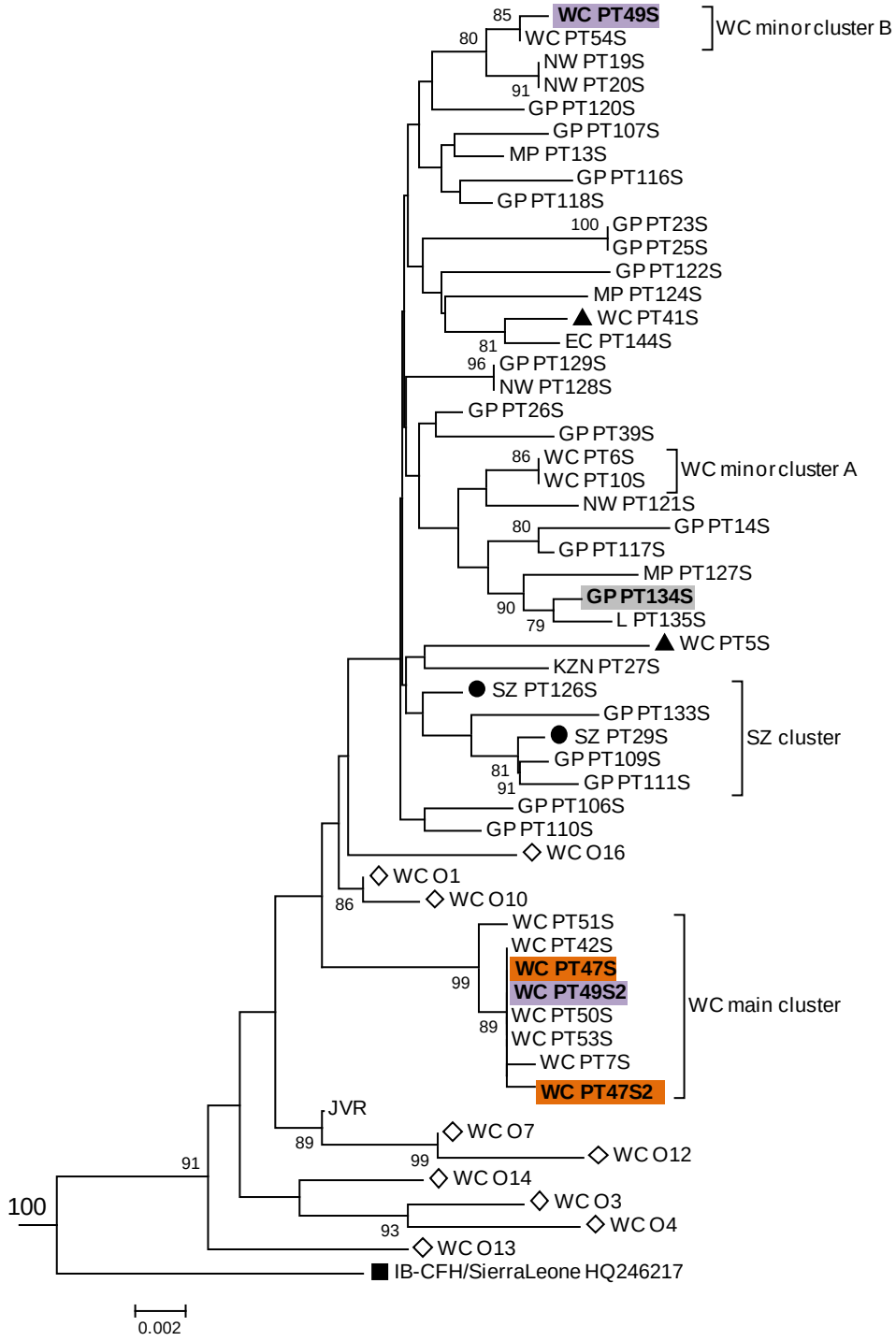

Figure 3.2: Phylogenetic analysis of HAV strains characterised from clinical specimens based on the sequence of the VP1 region (900 bp). The tree was constructed using the Kimura two-parameter model of the neighbour-joining method, with 1000 bootstrap replicates; bootstrap values ≥75% are indicated at nodes. Strains indicated with (●) represent non-SA strains typed in this study. Swaziland (SZ). The tree highlights the strains GP_PT134S (light grey), WC_PT47S and WC_PT47S2 (Heavy orange), WC_PT49S and WC_PT49S2 (Light purple) of 'SA major cluster' as depicted by Figure 3.1.

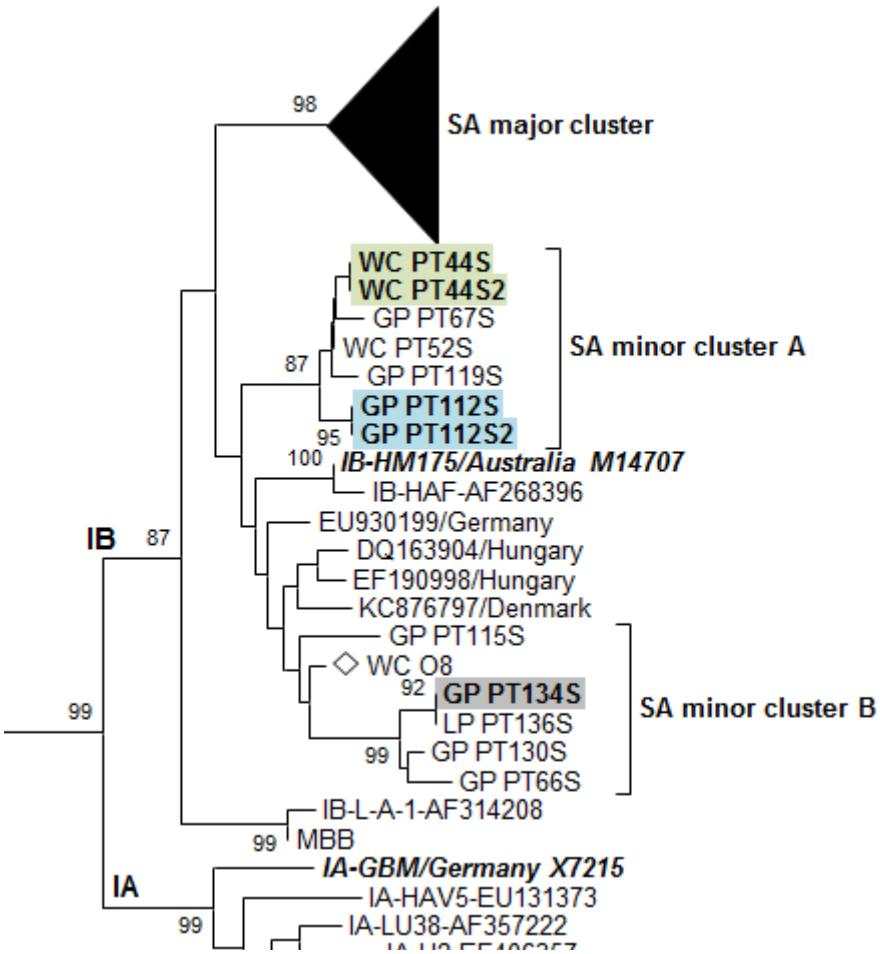

Figure 3.3: Phylogenetic analysis of HAV strains characterised from clinical specimens based on the sequence of the VP1/P2B junction (~350 bp). The tree was constructed using the Kimura two-parameter model of the neighbour-joining method, with 1000 bootstrap replicates; bootstrap values ≥75% are indicated at nodes. The tree highlights the strains GP_PT134S (light grey), GP_PT112S and GP_PT112S2 (light blue), WC_PT44S and WC_PT44S2 (light green) of 'SA minor clusters A and B' as depicted by Figure 2.3.

3.3.2.2 Phylogenetic analysis of HAV strains from water samples

Unusual clustering was not detected for HAV strains characterised from water samples. All strains detected and characterised from the water samples clustered within 'SA major cluster' with the exception of HAV strains from wastewater samples in MP and LP that formed 'SA minor cluster' on Figure 2.5.

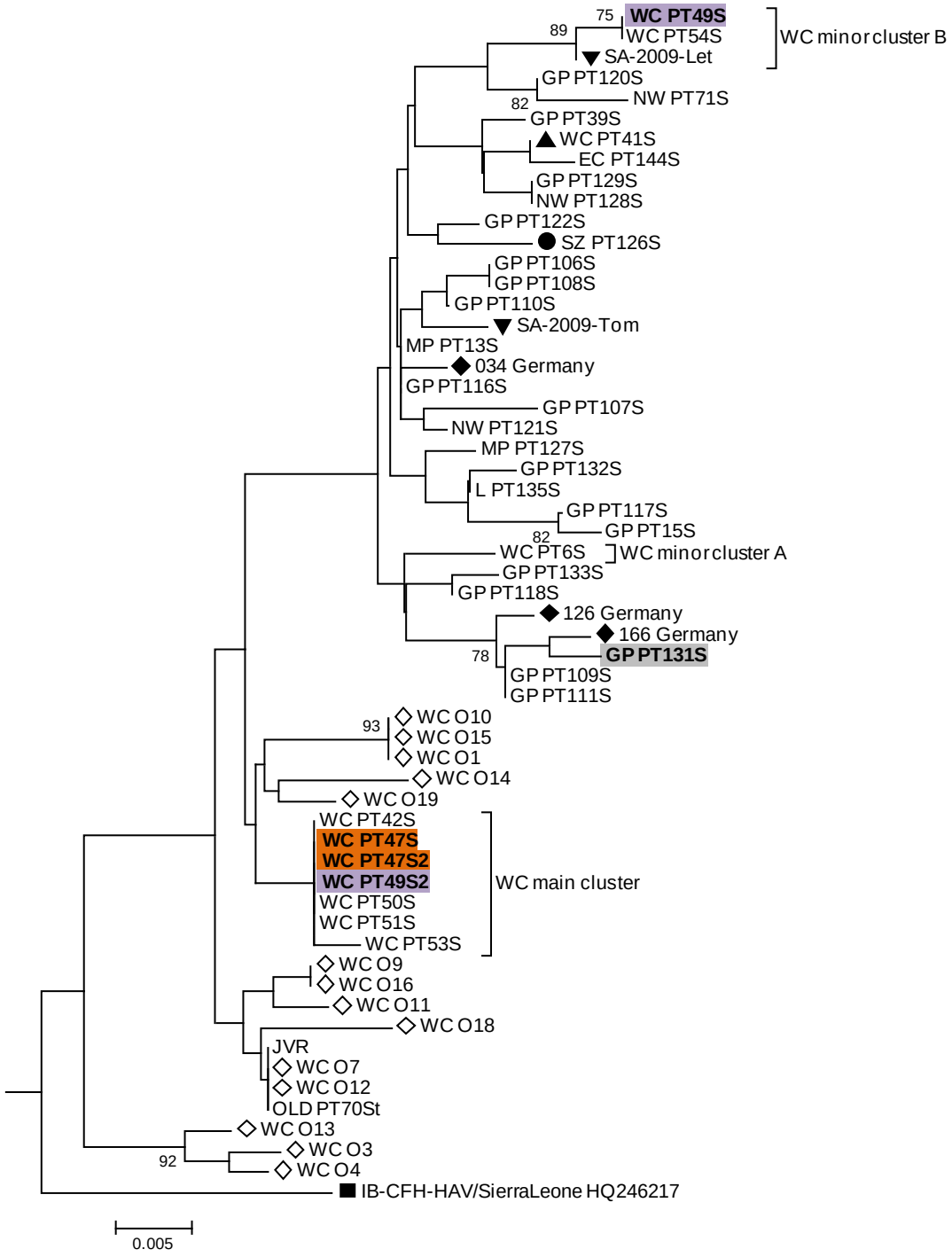

Figure 3.4: Phylogenetic analysis of HAV strains characterised from clinical specimens based on the sequence of the VP1/P2B region (~350 bp). The tree was constructed using the Kimura two-parameter model of the neighbour-joining method, with 1000 bootstrap replicates; bootstrap values ≥75% are indicated at nodes. Strains indicated with (●) represent non-SA strains typed in this study. Swaziland (SZ). The tree highlights the strains GP_PT131S (light grey), WC_PT47S and WC_PT47S2 (Heavy orange), WC_PT49S and WC_PT49S2 (Light purple) of 'SA major cluster' as depicted on Figure 3.3.

3.3.3 Quasispecies analysis

The strains WC_PT44S and WC_PT44S2 were detected in two sera specimens collected from the same patient, as well as strains WC_PT49S, WC_PT49S2. These strains showed minor differences at nucleotide level during phylogenetic analysis (Figures 3.1, 3.2, 3.3, 3.4). Strain GP_PT131S clustered differently over the VP1 and VP1/P2B regions. The five strains were cloned for quasispecies analysis. Phylogenetic analysis indicates genetic diversity among clones (Figure 3.5).

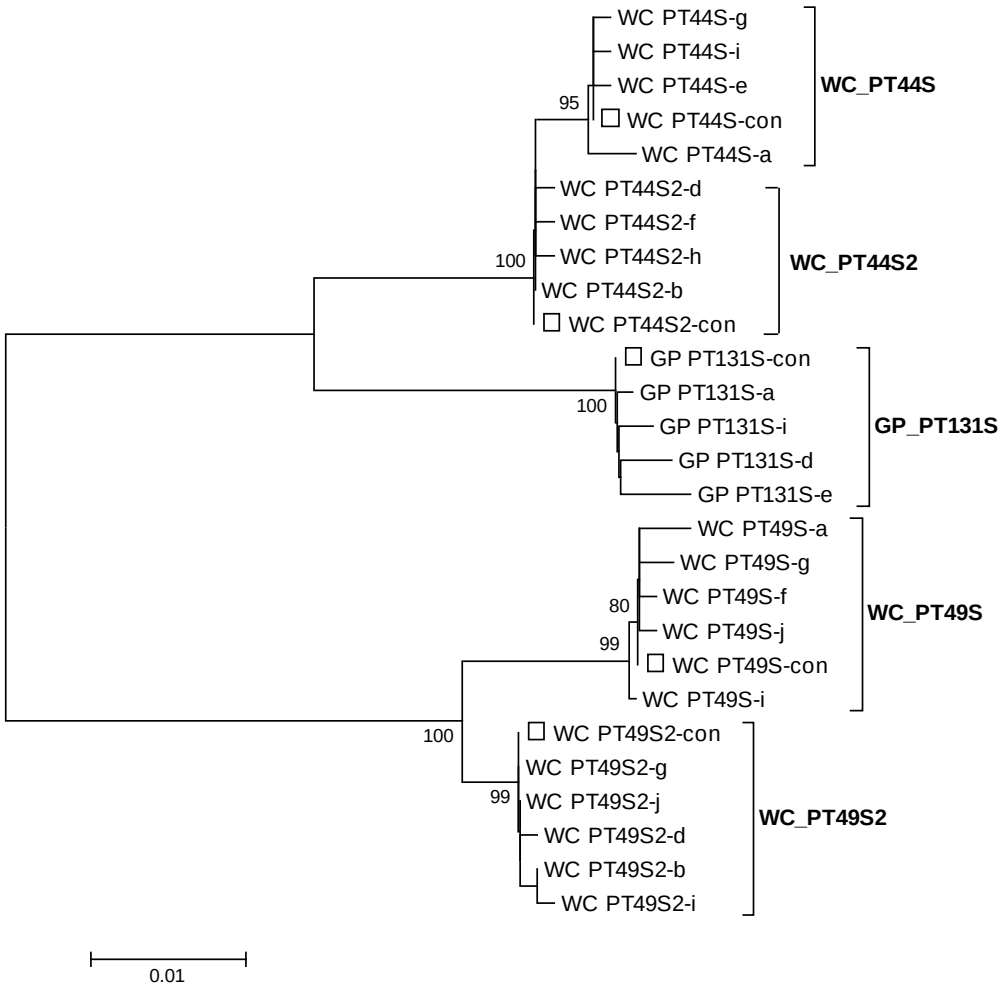

Figure 3.5: Quasispecies analysis of hepatitis A virus strains from five clinical specimens. The tree was constructed using the Kimura two-parameter model of the neighbour-joining method, with 1000 bootstrap replicates; bootstrap values ≥75% are indicated at nodes. Strains indicated with (□) represent consensus sequences. Strains with lowercase alphabetical letter at the end of their name represent clones characterised during the analysis.

3.3.4 Amino acid sequence analysis

Non-synonymous and synonymous substitutions were recorded among characterised strains. The recorded amino acid changes occurred at positions near neutralisation sites except for two, LP_9-20120424-col10 and LP_PVT-20120327 (Table 3.3). A mutation at residue 218 of the VP1 region (LP_KPL1-20120627-col3, LP_KPL2-20120424-col7a, LP_KPL2-20120424-col14a), that is located next to the glycophorin A binding site (G217), was also recorded. The mutations detected were not consistent among strains except for the R298K change over the VP1 region that corresponds to R63K over the VP1/P2B junction (positions relative to that of type strain HM175). The VP1 sequence of all SA HAV strains characterised from water and 35 clinical strains had the amino acid change R298K. The same amino acid substitution was seen in the sequences of HAV strains previously detected on tomato and lettuce (Netshikweta, 2011) as well as German tourists that presumably got infected while visiting SA (Faber *et al*., 2009). Among the non-SA strains only those from Swaziland had the R298K mutation. During quasispecies analysis, a number of amino acid changes were also recorded (Table 3.3). These mutations were not recorded after amplification and direct sequencing of the VP1 region.

Interesting amino acid changes were recorded over the VP1/P2B junction characterised from clinical specimens (Table 3.4). In addition to the R63K change, a R71S change is present on almost all sequences (21 clinical specimens and 29 strains from water samples) of strains characterised in this study as well as HAV strains previously detected on tomato and lettuce, and German tourists. The strain WC_PT49S has both the R63K and R71S mutations, that were absent on the sequence of the strain WC_PT49S2 detected from the same patient. Sequences for which the mutations C70S and E91D were recorded did not have the R63K and R71S mutations. The mutation M104I, recorded for five clinical strains from GP and WC, was also present on genotype IIIA and IIIB VP1/P2B sequences (Table 3.4).

Table 3.3: Amino acid substitutions observed at or near the neutralisation sites in the VP1 sequences characterised in this study.

Neutralising site[a,b,c]	Clinical specimens		Water samples	
	Strains	**Mutations**	**Strains**	**Mutations**
S102	WC_PT5S	S178T	LP_5-20120131-col14	C96Y, T116A, Y258H
N104	WC_PT44S	S246N, S293L	LP_7-20120131-col5	P291S
K105	WC_PT44S-a	I200T, S246N, S293L	LP_7-20120131-col7	N265K
S114	WC_PT44S-e	L113S, S246N, S293L	LP_9-20120424-col1	R292S
V166	WC_PT44S-g	L249S, S246N, S293L	LP_9-20120424-col2	T148I
W170	WC_PT44S-i	S246N, S293L	LP_9-20120424-col8	R292S
V171	WC_PT44S2	S293L	LP_9-20120424-col10	P69L, A176V, N197S, S266P, L284W
A176	WC_PT49S-a	L141S, D240G	LP_17B-20120327	T123S
Y181	WC_PT49S-g	E253G	LP_17B-20120327-col6	T123S
R189	WC_PT49S-j	R278G	LP_23-20120813-col6	V64I, L249S
G217	WC_PT49S2-b	S212P	LP_PVT-20120327-col4	S114F, V188G
K221	WC_PT49S2-i	S212P, R278K	LP_KPL1-20120627-col3	L270F
Q232	GP_PT111S	E174G	LP_KPL2-20120424	L218S
A280	GP_PT131S-d	T160I	LP_KPL2-20120424-col1	C96Y, L143S
	GP_PT131S-e	K182E	LP_KPL2-20120424-col6	F130L, I179T
	GP_PT131S-i	H238R	LP_KPL2-20120424-col7a	L218S
			LP_KPL2-20120424-col14a	N132D, L218S
			GP_RV2-20121112	S239F
			GP_11/1085-col8	S206F, D290V
			GP_11/1145-col2	T108A, S293P
			MP_10/1124-col20, 24, 27, 28	V229I
			MP_10/1391	N265D, S266P

a: Aragonès *et al.,* 2008
b: Pérez-Sautu *et al*., 2011a
c: Pintó *et al*., 2012

Table 3.4: Amino acid substitutions observed in the VP1/P2B sequences

Clinical specimens		**Water samples**	
Strains	**Mutations**	**Strains**	**Mutations**
GP_PT66S, GP_PT130S, GP_PT134S, LP_PT136S	C70S	GP_RV2-20121112	S4F
WC_PT44S, WC_PT44S2, GP_PT67S, WC_PT52S, GP_PT112S, GP_PT112S2, GP_PT119S	E91D	K_KD-Nov	K62R, L107P
WC_O9, WC_O11, WC_O16, GP_PT66S, GP_PT130S	M104I	K_KD-Dec	K62R

The mutations R63K and R71S have been consistently recorded on sequences of the VP1/P2B junction characterised from water samples. Hepatitis A virus strains characterised from water samples collected in Kenya have the R71S change but not the R63K mutation. Other amino acid changes have also been recorded (Table 3.4).

Amino acid sequence analysis of the sequence of clones from the HAV strain characterised from the sample KZN_20130530 indicated mutations unique to genotype IV and/or V over the VP1 and VP1/P2B regions (Table 3.5). Amino acid substitutions were recorded at the residues 102, 105 and 221 that have previously been reported to be neutralising sites (Pérez-Sautu *et al*., 2011a; Pintó *et al*., 2012).

Table 3.5: Similar amino acid substitutions recorded for the KZN_20130530 strain when compared to other hepatitis A virus genotypes.

	VP1 region	**VP1/P2B junction**
Unique mutation	V64I, K221S	104F,
Genotype IV	A281G	46G, 111S
Genotype V	M28R, S271V, E294A	70Q, 71G, 80M,
Genotypes IV and V	V1T, S102A, K105R, E174Q, S178T, V188I, L270M, T272S, M276L, S277D	97I
Genotype IIIA and IIIB	L270M, T272S, S277D	-
Genotype IIIB	M276L	-

3.3.5 Search for antigenic escape variants

Amino acid sequences of strains for which substitutions occurred at or near residues of the VP1 region contributing to the immunodominant antigenic site were entered into PHYRE[2]. The predicted 2D protein structures were compared to that of HM175. On the predicted VP1 protein structure of one strain, L_KPL2-20120424-col6, the Beta strand was interrupted by I179T. The same gap was observed for BCN60 that had a Y181S mutation. Another strain, GP_PT131S-e, detected during quasispecies analysis (Figure 3.5) had a K182E mutation, but there was no change on the predicted 2D protein structure. A 3D model structure of the VP1 protein with the recorded I179T change showed that the mutation was not exposed on the surface of the protein (Figure 3.6).

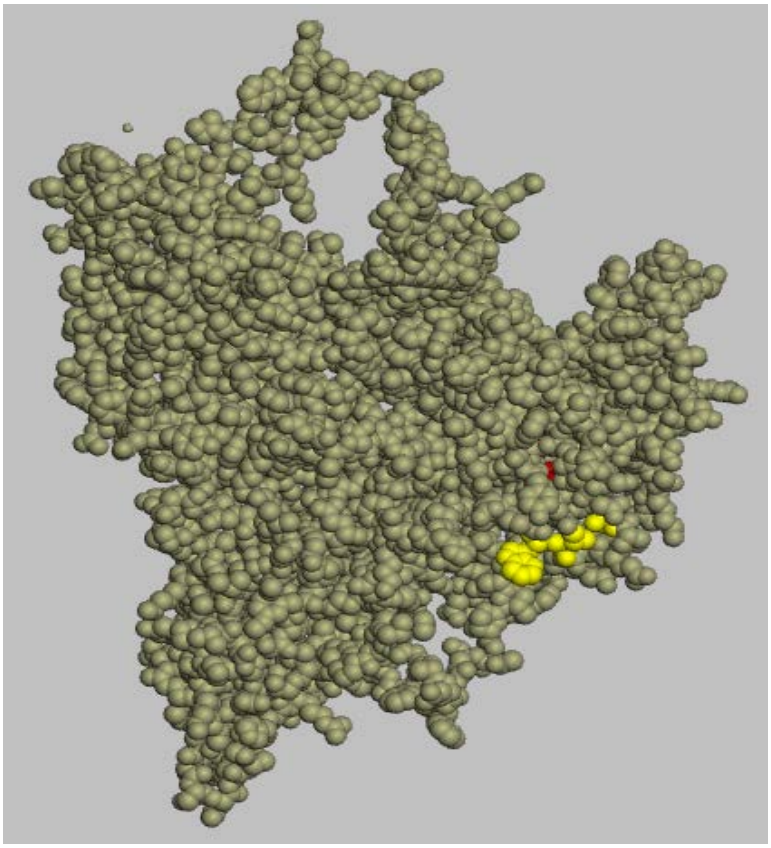

Figure 3.6: 3D model of VP1 structure highlighting neutralisation site at residue 181 (yellow) and recorded I179T mutation (red).

3.4 DISCUSSION

In this chapter, the genetic diversity of HAV detected in water samples and clinical specimens were assessed by pairwise analysis, phylogenetic analysis and quasispecies analysis. Phylogenetic analysis showed that the HAV strain ancestral to the unique SA lineage is closely related to the strain CFH-HAV (Figures 3.2, 3.4). Strain CFH-HAV was isolated from a chimpanzee that was transported from Sierra Leone to Japan, where it succumbed to fulminant hepatitis A (Abe and Shikata, 1982; Theamboonlers *et al.*, 2012). Hepatitis A virus has two biotypes, one of which is made of humans and chimpanzees (Knowles *et al*., 2012). It is therefore possible that the animal was infected by a handler who got infected by a HAV strain endemic to the southern African region.

The pairwise analysis indicated that genetic divergence of all characterised strains was higher over the VP1/P2B junction when compared to the VP1 region. This is contrary to previous reports indicating that VP1 is the most variable region of the HAV genome, as is seen with other picornaviruses (Costa-Mattioli *et al*., 2002; Pérez-Sautu *et al*., 2011b). During amplification of the VP1 region from HAV strains detected in water samples, multiple strains were characterised from a single sample (as identified in a number of clones) (Figure 2.6a). In contrast, no more than a single strain was detected during amplification of the VP1/P2B junction (Figure 2.7a). Mixed infections were not recorded while characterising clinical strains using either genomic region, even though two strains (GP_PT131S and GP_PT134S) grouped differently during phylogenetic analysis based on VP1 region and VP1/P2B junction (Figures 3.1, 3.2, 3.3, 3.4). A published molecular epidemiological study on foodborne hepatitis A outbreaks found that implicated food items were possibly contaminated with more than two HAV strains of which only one had established infection in patients (Vaughan *et al*., 2013b). In addition, water is a composite sample derived from a number of infected individuals, which could explain the presence of multiple strains in water samples and not in clinical specimens, but it does not clarify genetic divergence observed from pairwise analysis.

In chapter two, clones from the HAV strain characterised from the sample KZN_20130530, formed a cluster independent from all HAV genotypes and subgenotypes (Figures 2.4, 2.5, 2.6, 2.7). Based on the definitions of Robertson et al (1992) and Costa-Mattioli et al (2002), phylogenetic analysis together with pairwise analyses suggest that the strain KZN_20130530 belongs to a new HAV genotype. Genetic identities indicate that it could be a new simian strain, since highest genetic identities were obtained with a genotype V strain. The low nucleotide identity percentage seen with genotype IV was not considered for analysis because the VP1/P2B sequence downloaded from GenBank was incomplete at the carboxyl terminus. The high nucleotide identity seen with genotype VI was also not considered because the sequence available in GenBank is the VP1/P2A junction that is shorter than the VP1/P2B region.

Hepatitis A virus genotype V was isolated from an African green monkey (*Cercopithecus aethiops*) that was taken from Kenya to the Union of Soviet Socialist Republics (USSR) (Robertson *et al*., 1992). The sample, characterised in this study, was collected from a dam used to irrigate cabbage on a commercial farm in KZN. Vervet monkeys (*Cercopithecus pygerythrus*) are often seen on the farm where the sample was collected, as well as other farms in the province. The Samango monkey (*Cercopithecus mitus* subspecies *erythrarchus* and *Cercopithecus mitus* subspecies *Labiatus*) and the chacma Baboon are also found in KZN, but they were not recorded on the farm. Hepatitis A virus seroprevalence studies were previously carried out on the Chacma baboon and vervet monkeys in SA (Taylor and Prozesky, 1985; Drewe *et al*., 2012). The study found a high rate of seropositivity in the animals, but could not establish whether it was due to natural infection or contact with human excreta or water or food contaminated with human faeces. Additionally, the nucleic acid sequences of the clones from the strain KZN_20130530 had amino acid substitutions unique to genotype V and/or genotype IV, as well as genotypes IIIA and/or IIIB (Table 3.5), but the sequences lacked the R298K seen in the remaining SA strains. Three of these amino acid changes occurred at neutralising sites, suggesting that immune response to the KZN_20130530 strain might be different than immune responses to previously isolated HAV strains. It is therefore possible that unique HAV strains are circulating in the SA simian population. Additional sampling of simian's excreta and sludge of the sampled dam in the KZN region could help clarify the recorded observations.

Hepatitis A virus has six genotypes with distinct geographical distribution (Robertson *et al*., 1992; Nainan *et al*., 2006; Vaughan *et al*., 2013a). Hepatitis A virus strains characterised from clinical specimens and water samples collected from SA had a R298K mutation at the amino acid level. That mutation was also present on the amino acid sequence of German tourists that presumably were infected in SA (Faber *et al*., 2009) and of fresh produce previously analysed (Netshikweta, 2011). However, the mutation was absent on the sequences obtained from non-SA strains. In addition, six clinical strains from WC and six clinical strains from GP did not have the R63K and R71S

mutations recorded for other SA strains, suggesting that these specific strains may have been imported. This further accentuates the importance of molecular characterisation of HAV strains by means of nucleotide sequence analysis for source tracking purposes.

Hepatitis A virus has a low degree of antigenic variability as demonstrated by the presence of a single serotype (Jacobsen and Wiersma, 2010; Knowles *et al*., 2012). Several amino acid substitutions were recorded during analysis of the VP1 sequences. Changes were recorded at/or near residues contributing to the immunodominant site and the glycophorin A binding site. Genetic diversity was further highlighted during quasispecies analysis. Phylogenetic analysis of clones from clinical strains showed that HAV replicates as quasispecies in infected individuals in SA (Figure 3.5). The amino acid sequences of the clones harboured mutations that were not detected on the consensus sequence obtained after VP1 amplification and direct sequencing. Quasispecies in HAV infected individuals have previously been reported (Sánchez *et al*., 2003a; Sulbaran *et al*., 2010; Wang *et al*., 2013), but it seems like these mutations are negatively selected as they were not commonly observed. Previous studies have reported that structural constraint on the HAV capsid and usage of rare codons might prevent the emergence of new variants of the virus (Aragonès *et al*., 2008, 2010; Pintó *et al*., 2012; Vaughan *et al*., 2013a). However, specific immune pressure conditions such as partial immunity to HAV coupled with exposure to multiple strains of the virus might lead to the emergence of new antigenic escape mutants. In Barcelona, Spain, a study isolated antigenic escape mutants among active homosexual men who had not completed the vaccination schedule and were exposed to multiple strains of HAV (Pérez-Sautu *et al*., 2011a,b). In SA, the presence of large numbers of immunocompromised individuals due to HIV/AIDS, approximately 5.26 million people (Statistics SA, 2013), might allow the mutants that we detected in the study to emerge and establish themselves in the population.

A study published in 2011 reported a naturally occurring antigenic escape mutant (BCN60) with a Y181S mutation (Pérez-Sautu *et al*., 2011a). In the

present study, the amino acid substitutions I179T and K182E were observed on the sequence of LP_20120424-col6 and GP_PT131S-e, which were clones obtained during genotyping and quasispecies analysis, respectively. The 2D protein model structure of LP_20120424-col6 indicated a gap on the Beta strand as was seen with BCN60. No change was recorded on the 2D protein model of GP_PT131S-e. The 3D model structure of the VP1 protein with I179T change showed that the mutation was not exposed at the surface of the protein (Figure 3.6). Therefore, the conformation of the immunodominant antigenic site may be conserved. However, the mutation replaced an amino acid containing an aliphatic side chain with one containing an aromatic side chain. Hence, interaction with the amino acid at position 181 might impact on the antibody binding sites.

In summary, a genetically diverse HAV population is circulating in SA with the possibility of new variants. The HAV strain that gave rise to the SA lineage seemed to have gain fitness advantage in the geographical region through R298K mutation. This study provides the first evidence of quasispecies among HAV strains infecting individuals in SA. In addition, data suggests that a previously unidentified HAV strain is possibly circulating in the SA simian population. The presence of HAV variants with mutations close to the neutralisation site and potential new genotypes stresses the need to implement a proper surveillance system to monitor the circulation of the virus within human and simian SA population.

CHAPTER 4

GENERAL DISCUSSION

Hepatitis A virus infection is the most common cause of acute viral hepatitis infection worldwide (Wang *et al*., 2013). Hepatitis A virus is an antigenically stable picornavirus with a modest degree of genetic diversity (Nainan *et al*., 2006; Knowles *et al*., 2012; Pintó *et al*., 2012; Wang *et al*., 2013). Initially, nucleotide sequence analysis of the VP1/P2A region was used to genotype HAV strains (Robertson *et al*., 1992). Current typing methods use larger genomic regions like the VP1 region (Costa-Mattioli *et al*., 2002) or the VP1/P2B junction (Wheeler *et al*., 2005; Nainan *et al*., 2006; Bialek *et al*., 2007; Shieh *et al*., 2007; Vaughan *et al*., 2013b). Molecular epidemiological data on the types of HAV circulating in SA has not been published since 1997 when it was established that genotype IB predominates in SA (Taylor, 1997). Since then, HAV has been detected in surface water sources (Taylor *et al*., 2001; Venter, 2004) and on fresh produce at the point of retail in the country (Netshikweta, 2011) suggesting that faecally-contaminated food and water is a possible source of HAV in SA. In addition, recent publications have reported the detection of “escape” mutants which could result in vaccination failure (Pérez-Sautu *et al*., 2011a; Pintó *et al*., 2012). For these reasons, the aim of the study was to establish which genotypes, and possibly variants thereof, of HAV are circulating the community compared to the genotypes present in water and food sources in SA.

The first objective of the investigation was to recover and detect HAV in environmental samples i.e. fresh produce and faecally-contaminated surface water. Fresh produce that are minimally processed such as tomatoes, cabbages and lettuce were collected at pre- and post-harvest points and analysed. Viruses were recovered from the fresh produce samples by elution

with TGBE and concentrated by PEG_{8000}/NaCl precipitation (Croci *et al*., 2008; Netshikweta, 2011). None of the 20 fresh produce samples analysed tested positive for HAV. The small limited sample size could have accounted for the negative result as presence of inhibitors was ruled out by the positive detection of the process control.

Water samples (irrigation, wastewater and surface) were collected from six SA regions (GP, KZN, LP, MP, NW and WC) as well as Kenya. Viruses were recovered and concentrated from the water samples using the glass wool adsorption elution method (Vilaginès *et al.,* 1993; Venter, 2004; Mans *et al*., 2013) with PEG_{8000}/NaCl precipitation for secondary concentration (ISO/TS 15216-2:2013[E]). Thereafter, HAV was detected by real-time RT-PCR analysis. Hepatitis A virus was detected in 62% of the irrigation water samples and 37% of the wastewater samples. All the surface water analysed tested positive for HAV. Discrepancies were noted during the analysis of water samples. The detection of HAV in wastewater samples was approximately half of the detection rate in irrigation water samples. Hepatitis A virus is an enteric virus and high HAV detection rates are expected to be recorded for wastewaters. The presence of inhibitors in the wastewater samples could have accounted for the low detection rate. The sample RNA for which the detection assay failed were diluted and retested for HAV, but the detection rate was not substantially improved. Although other enteric viruses, namely noroviruses and sapoviruses were detected in the wastewaters analysed (Murray *et al*., 2013) the low detection rate indicates that the prevalence of HAV reported for these samples may have been an underestimate.

The second objective of this investigation was to detect HAV by real-time RT-PCR in clinical specimens (stools and serum). A total of 117 specimens were collected from SA and non-SA regions. Hepatitis A virus was detected in 81% of the SA specimens and in 45% of the non-SA specimens. The low detection rate recorded for non-SA specimens could be due to transport and storage conditions or to specimens collected long after the onset of symptoms (Pérez-Sautu *et al*., 2011b).

The third and fourth objectives were to genotype HAV strains detected in clinical and environmental sources and to determine genetic relatedness between strains from different sources. Hepatitis A virus strains could be genotyped from 67 HAV positive specimens and 33 water samples, by nucleotide sequence analysis of the VP1 and VP1/P2B genomic regions, pairwise comparison with reference sequence from GenBank and phylogenetic analysis. Except for one strain from an irrigation water sample, all the HAV strains (from both clinical and water sources) characterised in this study clustered within genotype IB. The SA strains formed a unique cluster (SA major cluster) together with strains previously detected on fresh produce (tomatoes and lettuce) (Netshikweta, 2011) and in German tourists that presumably were infected in SA (Faber *et al*., 2009) within genotype IB with >95% nucleotide identity to the type strain HM175 (M14707). A number of SA strains grouped outside the 'SA major cluster' but still within genotype IB, some of which demonstrated different clustering during phylogenetic analysis using the VP1 or VP1/P2B regions. The strains characterised from non-SA clinical and water sources clustered within genotype IB but outside 'SA major cluster'.

Pairwise analysis performed at the amino acid levels revealed that all SA strains characterised from water samples and 35 SA strains characterised from clinical specimens in this study had a unique R298K mutation over the VP1 region. In addition a R71S substitution was recorded for 50 SA VP1/P2B sequences. The amino acid mutations were also present on the sequences of strains previously characterised from fresh produce (Netshikweta, 2011) and German tourists (Faber *et al*., 2009) but not on the nucleic acid sequences of non-SA HAV strains. The results from pairwise analysis and phylogenetic analysis further confirm the distinct geographical distribution of HAV strains. This is important, as it indicates that the data obtained can be used as reference data for molecular source tracking in the event of an outbreak to distinguish locally acquired infection from imported ones and possibly identify sources of contamination and/or infections.

The fifth and last objective was to analyse the nucleotide sequences obtained in the study for the presence of novel mutations. The analysis of nucleotide sequences of strains characterised in this study revealed that a genetically diverse HAV population circulates in SA. Amino acid substitutions that have not been previously reported were recorded at/or near neutralisation site. Interesting amino acid substitutions were also recorded at the VP1/P2B junction. Even if residues within the VP1/2B junction may not contribute to the immunodominant antigenic site, the presence of mutations together with a high nucleotide divergence recorded and different grouping of strains indicate the importance of VP1/P2B during genetic characterisation of HAV strains. A possible antigenic "escape" mutant (LP_KPL2-20120424-col6) was detected in an irrigation water sample from LP. The strain had a I179T mutation in the VP1 region that gave rise to a 2D protein structure similar to the mutant BCN60 which had a Y181S mutation (Pérez-Sautu *et al*., 2011a). In addition, this study provided the first evidence of quasispecies among HAV strains infecting individuals in SA and a candidate for a new HAV genotype infecting simian.

Analysing water sources, which provides valuable information with regard to viruses in the surrounding communities, and analysing clinical specimens indicated that unique HAV genotype IB strains predominate in the SA community with the possible emergence of antigenic escape mutants. The unique VP1 and VP1/P2B sequences of the genomes of SA HAV strains make molecular source tracking in the event of an outbreak possible. Due to the presence of genotype IB in surface water sources used to irrigate fresh produce it can be hypothesised that faecally-contaminated water sources are a potential vehicle of transmission of HAV infection and a potential source of contamination of irrigated fresh produce in SA. The identification of a novel strain unrelated to previously described HAV suggest that a new genotype, possibly VII, circulate in the SA simian population. The presence of HAV variants with mutations close to the neutralisation site and potential new genotypes stresses the need to implement a proper surveillance system to monitor the circulation of the virus within human and simian SA population.

CHAPTER 5

REFERENCES

Abe K, Shikata T. Fulminant type A viral hepatitis in a chimpanzee. Acta Pathol Jpn 1982;32:143–148.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997;25:3389–3402.

Amado LA, Villar LM, de Paula VS, Gaspar AMC. Comparison between serum and saliva for the detection of hepatitis A virus RNA. J Virol Methods 2008;148:74-80.

Amado LA, Villar LM, de Paula VS, Pintó MA, Gaspar AM. Exposure to multiple subgenotypes of hepatitis A virus during an outbreak using matched serum and saliva specimens. J Med Virol 2011;83:768-775.

Aragonès L, Bosch A, Pintó RM. Hepatitis A virus mutant spectra under the selective pressure of monoclonal antibodies: codon usage constraints limit capsid variability. J Virol 2008;82:1688-1700.

Aragonès L, Guix S, Ribes E, Bosch, A, Pintó RM. Fine-Tuning Translation Kinetics Selection as the Driving Force of Codon Usage Bias in the Hepatitis A Virus Capsid. PLoS Path 2010;doi: 10.1371/journal.ppat.1000797.

Arora D, Jindal D, Shukla RK, Bansal R. Water borne hepatitis A and hepatitis E in Malwa region of Punjab, India. JCDR 2013;7: 2163-2166.

Beard MR, Lemon SM. Hepatitis A Virus (*Picornaviridae*). In: Granoff A, Webster RG, editors. Encyclopedia of Virology, 2nd ed. Norwich: Academic Press; 1999. p. 631-639.

Berger CN, Sodha SV, Shaw RK, Griffin PM, Pink D, Hand P, Frankel G. Fresh fruit and vegetables as vehicles for the transmission of human pathogens. Environ Microbiol 2010;12:2385-2397.

Bialek SR, George PA, Xia G-L, Glatzer MB, Motes ML, Veazey JE, Hammond RM, Jones T, Shieh YC, Wamnes J, Vaughan G, Khudyakov Y, Fiore AE. Use of molecular epidemiology to confirm a multistate outbreak of hepatitis A caused by consumption of oysters. Clin Infect Dis 2007;44:838-840.

Bidawid S, Faber JM, Sattar SA. Survival of hepatitis A virus on modified atmosphere-packaged (MAP) lettuce. Food Microbiol 2001;18:95-102.

Boom R, Sol CJ, Salimans MM, Jansen CL, van Dillen PMW, van der Noordaa J. Rapid and simple method for purification of nucleic acids. J. Clin. Microbiol 1990;28:495-503.

Bosch A, Sánchez G, Abbaszadegan M, Carducci A, Guix S, Le Guyader FS, Netshikweta R, Pintó RM, van der Poel WHM, Rutjes S, Sano D, Taylor MB, van Zyl WB, Rodríguez-Lázaro D, Kovač K, Sellwood J. Analytical methods for virus detection in water and food. Food Anal Methods 2011;4:4-12.

Bosch A, Sanchez G, Le Guyader F, Vanaclocha H, Haugarreau L, Pintó RM. Human enteric viruses in Coquina clams associated with a large hepatitis A outbreak. Water Sci. Technol 2001;43:61–65.

Bovier PA, Bock J, Ebengo TF, Frösner G, Glaus J, Herzog C, Loutan L. Predicted 30-year protection after vaccination with aluminium-free virosomal hepatitis a vaccine. J Med Virol 2010;82:1629-1634.

Bowen GS, McCarthy MA. Hepatitis A associated with a hardware store water fountain and a contaminated well in Lancaster County, Pennsylvania, 1980. Am J Epidemiol 1983;117:695-705.

Bower WA, Nainan OV, Han X, Margolis HS. Duration of viremia in hepatitis A virus infection. J Infect Dis 2000;182:12-17.

Brown EA, Jansen RW, Lemon SM. Characterization of a simian hepatitis A virus (HAV): antigenic and genetic comparison with human HAV. J Virol 1989;63:4932-4937.

Casas N, Suñén N. Detection of enteroviruses, hepatitis A virus and rotaviruses in sewage by means of an immunomagnetic capture reverse transcription-PCR assay. Microbiol Res 2002;157:169-176.

Centers for Disease Control and Prevention. 12 October 2013. Foodborne outbreak online database. CDC, Bethesda, MD. Available from URL http://wwwn.cdc.gov/foodborneoutbreaks/Default.aspx.

Collier L, Oxford J. Introduction to the hepatitis viruses. In: Collier L, Oxford J, Pipkin J, editors. Human Virology. A text for students of medicine, dentistry, and microbiology. New York: Oxford University Press Inc; 2006. Chapter 21. p.159-160.

Costafreda MI, Bosch A, Pintó RM. Development, evaluation, and standardization of a real-time Taqman reverse transcription-PCR assay for the quantification of hepatitis A virus in clinical and shellfish samples. Appl Environ Microbiol 2006;72:3846-3855.

Costafreda MI, Ribes E, Franch A, Bosch A, Pintó RM. A single mutation in the glycophorin A binding site of hepatitis A virus enhances virus clearance from the blood and results in a lower fitness variant. J Virol 2012;86:7887-7895.

Costa-Mattioli M, Cristina J, Romero H, Pérez-Bercof R, Casane D, Colina R, Garcia L, Vega I, Glikman G, Romanowsky V, Castello A, Nicand E, Gassin M, Billaudel S, Ferré V. Molecular evolution of hepatitis A virus: a new classification based on the complete VP1 protein. J Virol 2002;76:9516-9525.

Costa-Mattioli M, Di Napoli A, Ferré V, Billaudel S, Pérez-Bercoff R, Cristina J. Genetic variability of hepatitis A virus. J Gen Virol 2003;84:3191-3201.

Craun GF, Brunkard JM, Yoder JS, Roberts VA, Carpenter J, Wade T, Calderon LR, Roberts JM, Beach MJ, Roy SL. Causes of outbreaks associated with drinking water in the United States from 1971 to 2006. Clin Microbiol Rev 2010;23:507-28.

Cristina J, Costa-Mattioli M. Genetic variability and molecular evolution of hepatitis A virus. Virus Res 2007;127:151-157

Croci L, De Medici D, Scalfaro C, Fiore A, Toti L. The survival of hepatitis A virus in fresh produce. Int J Food Microbiol 2002;73:29-34.

Croci L, Dubois E, Cook N, de Medici D, Schultz AC, China B, Rutjes SA, Hoorfar J, van der Poel WHM. Current methods for extraction and concentration of enteric viruses from fresh fruit and vegetables; Towards international standard. Food Anal Methods 2008;1:73-84.

Cromeans TL, Sobsey MD. Development of molecular methods for detection of infectious viruses in treated wastewater. WERF (99-HHE-5UR), London, IWA Publishing, 2004.

Cuthbert JA. Hepatitis A: old and new. Clin Microbiol Rev 2001;14:38-58.

De Jong G. Guidelines for the control of hepatitis A in South Africa. 2007.p.1-15. Available from URL: http://nicd.ac.za/assets/files/NICD%20Guidelines%20for%20the%20Control%22of%20Hepatitis%20A%20in%20SA.PDF.

Deboosere N, Pinon A, Caudrelier Y, Delobel A, Merle G, Perelle S, Temmam S, Loutreul J, Morin T, Estienney M, Belliot G, Pothier P, Gantzer C, Vialette M. Adhesion of human pathogenic enteric viruses and surrogate viruses to inert and vegetal food surfaces. Food Microbiol 2012;32:48-56.

Deng MY, Day SP, Cliver DO. Detection of hepatitis A virus in environmental samples by antigen-capture PCR. Appl Environ Microbiol. 1994;60:1927–1933.

Dentinger CM, Bower WA, Nainan OV, Cotter SM, Myers G, Dubusky LM, Fowler S, Salehi EDP, Bell BP. An outbreak of hepatitis A associated with green onions. J Infect Dis 2001;183:1273-1276.

Desbois D, Couturier E, Mackiewicz V, Graube A, Letort MJ, Dussaix E, Roque-Afonso AM. Epidemiology and genetic characterization of hepatitis A virus genotype IIA. J Clin Microbiol 2010;4:3306-3315.

Desenclos JC, Klontz KC, Wilder MH, Nainan OV, Margolis HS, Gunn RA. A multistate outbreak of hepatitis A caused by the consumption of raw oysters. Am J Public Health 1991;81:1268-72.

Donnan EJ, Fielding JE, Gregory JE, Lalor K, Rowe S, Goldsmith P, Antoniou M, Fullerton KE, Knope K, Copland JG, Bowden DS, Tracy, SL, Hogg GG, Tan A, Adamopoulos J, Gaston J, Vally H. A multistate outbreak of hepatitis A associated with semidried tomatoes in Australia, 2009. Clin Infect Dis 2012;54:775-781,doi: 10.1093/cid/cir949

Dotzauer A. Hepatitis A Virus. In: Mahy BWJ, van Regenmortel MV, editors. Encyclopedia of Virology, 3rd ed. Slovenia: Academic Press; 2008. p. 343-350.

Drewe JA, O'Riain MJ, Beamish E, Currie H, Parsons S. Survey of infections transmissible between baboons and humans, Cape Town, South Africa. Emerg Infect Dis 2012;18:298-301.

Dubois E, Hennechart C, Merle G, Burger C, Hmila N, Ruelle S, Perelle S, Ferré V. Detection and quantification by real-time RT-PCR of hepatitis A virus from inoculated tap waters, salad vegetables, and soft fruits: characterization of the method performances. Int J Food Microbiol 2007;117:141-149.

El-Sayed R, El-Karaksy H. Acute pancreatitis complicating acute hepatitis A virus infection. Arab J Gastroenterol 2012;13:184-185.

Endo K, Takahashi M, Masuko K, Inoue K, Akahane Y, Okamoto H. Full-length sequences of subgenotype IIIA and IIIB hepatitis A virus isolates: characterization of genotype III HAV genomes. Virus Res 2007;126:116-127.

Faber MS, Stark K, Behnke SC, Schreier E, Frank C. Epidemiology of hepatitis A infection, Germany, 2007-2008. Emerg Infect Dis 2009;15:1760-1768.

Feinstone SM, Kapikian AZ, Purceli RH. Hepatitis A: detection by immune electron microscopy of a viruslike antigen associated with acute illness. Science 1973;182:1026-1028.

Feng Z, Hensley L, McKnight KL, Hu F, Madden V, Ping L, Jeong SH, Walker C, Lanford RE, Lemon SM. A pathogenic picornavirus acquires an envelope by hijacking cellular membranes. Nature 2013;496:367-371.

Fiore AE. Hepatitis A transmitted by food. Clin Infect Dis 2004;38:705-715.

FitzSimons D, Hendrickx G, Vorsters A, Van Damme P. Hepatitis A and E: update on prevention and epidemiology. Vaccine 2010;28:583-588

Forbi JC, Agwale SM, Ndip LM, Esona MD. Genetic analysis of hepatitis A virus variants circulating among children presenting with acute diarrhea in Cameroon. J Med Virol 2012;84:728-732.

Forbi JC, Esona MD, Agwale SM. Molecular characterization of hepatitis A virus isolates from Nigeria. Intervirol 2013;56:22-26.

Franco E, Meleleo C, Serino L, Sorbara D, Zaratti L. Hepatitis A: epidemiology and prevention in developing countries. World J Hepatol 2012;4:68-73.

Gallot C, Grout L, Roque-Afonso A-M, Couturier E, Carillo-Santisteve P, Pouey J, Letort M-J, Hoppe S, Capdepon P, Saint-Martin S, De Valk H, Vaillant V. Hepatitis A associated with semidried tomatoes, France, 2010. Emerg Infect Dis 2011;17:566-567.

Gharbi-Khelifi H, Abid NBS, Beji A, Bhiri L, Harrath R, Sdiri K, Billaudel S, Ferre V, Aouni M. Seroprevalence and molecular characterisation of human hepatitis A virus in serum samples of Tunisian patients with clinical symptoms of viral hepatitis. Indian J. Virol 2012;23:29–35.

Gharbi-Khelifi H, Ferre V, Sdiri K, Berthome M, Fki L, Harrath R, Billaudel S, Aouni M. Hepatitis A in Tunisia: phylogenetic analysis of hepatitis A virus from 2001 to 2004. J Virol Methods 2006;138:109-116.

Halliday ML, Kang LY, Zhou TK, Hu MD, Pan QC, Fu TY, Huang YS, Hu SL. An epidemic of hepatitis A attributable to the ingestion of raw clams in Shanghai, China. J Infect Dis 1991;164:852-859.

Hollinger FB, Emerson SU. Hepatitis A virus. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA, Roizman B, Straus SE, editors. Fields Virology, 5th ed. Philadelphia Lippincott Williams and Wilkins; 2007. p. 913-947.

Hutin YJF, Pool V, Cramer EH, Nainan OV, Weth J, Williams IT, Goldstein ST, Gensheimer KF, Bell BP, Shapiro CN, Alter MJ, Margolis HS. A multistate, foodborne outbreak of hepatitis A. N Engl J Med 1999;340:595–602.

Ikner LA, Gerba CP, Bright KR. Concentration and recovery of viruses from water: a comprehensive review. Food Environ Virol 2012;4:41-67.

ISO/TS 15216-2:2013(E). Microbiology of food and animal feed – horizontal method for determination of hepatitis A virus and norovirus in food using real-time RT-PCR – Part 2: Methods for qualitative detection.

Jacobsen KH, Wiersma ST. Hepatitis A virus seroprevalence by age and world region, 1990 and 2005. Vaccine 2010;28:6653-6657.

Jacobsen KH. The global prevalence of hepatitis A virus infection and susceptibility: a systematic review. Geneva: WHO. 2009.p.1-431.

Jean J, Blais B, Darveau A, Fliss I. Detection of hepatitis A virus by the nucleic acid sequence-based amplification technique and comparison with reverse transcription-PCR. Appl Environ Microbiol 2001;67:5593–5600.

Jiang YJ, Liao GY, Zhao W, Sun MB, Qian Y, Bain CX, Jiang SD. Detection of infectious hepatitis A virus by integrated cell culture/strand-specific reverse transcriptase-polymerase chain reaction. J Appl Microbiol 2004;97:1105-1112.

Jothikumar N, Cliver DO, Mariam TW. Immunomagnetic Capture PCR for Rapid Concentration and Detection of Hepatitis A Virus from Environmental Samples. Appl Environ Microbiol 1998;64:504–508.

Kedda MA, Kew MC, Cohn MJ, Field SP, Schwyzer R, Song E, Fernandes-Costa F. An outbreak of hepatitis A among South African patients with hemophilia: evidence implicating contaminated factor VIII concentrate as the source. Hepatol 1995;22:1363-1367.

Kelley LA, Sternberg MJE. Protein structure prediction on the Web: a case study using the Phyre server. Nat Protoc 2009;4:363-371.

Knowles NJ, Hovi T, Hyypiä T, King AMQ, Lindberg AM, Pallansch MA, Palmenberg AC, Simmonds P, Skern T, Stanway G, Yamashita T, Zell R. Family Picornaviridae. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ. Editors, Virus Taxonomy. Ninth Report of the International Committee on Taxonomy of Viruses. San Diego: Academic Press; 2012. p. 855-880.

Koopmans M, Duizer E. Foodborne viruses: an emerging problem. Int J Food Microbiol 2004;90:23-41.

Kumar V, Das S, Jameel S. The biology and pathogenesis of hepatitis viruses. Curr Sci 2010;98:312-325.

La Rosa G, Fratini M, della Libera S, Iaconelli M, Muscillo M. Emerging and potentially emerging viruses in water environments. Ann Ist Super Sanità 2012;48:397-406.

Lassen SG, Soborg B, Midgley SE, Steens A, Vold L, Stene-Johansen K, Rimhanen-Finne R, Kontio M, Löfdahl M, Sundqvist L, Edelstein M, Jensen T, Verstergaard HT, Fischer TK, Mølbak K, Ethelberg S. Ongoing multi-strain food-borne hepatitis A outbreak with frozen berries as suspected vehicle: four Nordic countries affected, October 2012 to April 2013. Euro Surveill 2013;18:pii=20467. Available online: http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20467.

Lavanchy D. Viral hepatitis: global goals for vaccination. J Clin Virol 2012;55:296-302.

LeDuc JW, Lemon SM, Keenan CM, Graham RR, Marchwicki RH, Binn LN: Experimental infection of the New World owl monkey (*Aotus trivirgatus*) with hepatitis A virus. Infect Immunol 1983;40:766–772.

Li J, Bi Y, Chen C, Yang L, Ding C, Liu W. Genetic characterization of duck hepatitis A viruses isolated in China. Virus Res 2013;178:211-216.

Lodder WJ, van den Berg HHJL, Rutjes SA, de Roda Husman AM. Presence of enteric viruses in source waters for drinking water production in the Netherlands. Appl Environ Microbiol 2010;76:5965-5971.

Lu L, Ching KZ, de Paula VS, Nakano T, Siegl G, Weitz M, Robertson BH. Characterization of the complete genomic sequence of genotype II hepatitis A virus (CF53/Berne isolate). J Gen Virol 2004;85:2943-2952.

Mahoney FJ, Farley TA, Kelso KY, Wilson SA, Horan JM, McFarland LM. An outbreak of hepatitis A associated with swimming in a public pool. J Infect Dis 1992;165:613-8.

Mans J, Netshikweta R, Magwalivha M, van Zyl WB, Taylor MB. Diverse norovirus genotypes identified in sewage-polluted river water in South Africa. Epidemiol Infect 2013;141:303-313

Martin A, Lemon SM. Hepatitis A virus: from discovery to vaccines. Hepatology 2006;43:S164–S172.

Mattison K, Bidawid S. Analytical methods for food and environmental viruses. Food Environ Virol 2009;1:107-122.

Mbithi JN, Spingthorpe VS, Boulet JR, Sattar SA. Survival of hepatitis A virus on human hands and its transfer on contact with animate and inanimate surfaces. J Clin Virol 1992;30:757-763.

Mihigo R, Nshimirimana D, Hall A, Kew M, Wiersma S, Clements CJ. Control of viral hepatitis infection in Africa: are we dreaming? Vaccine 2013;31:341-346.

Miyamura T. Hepatitis E virus infection in developed countries. Virus Res 2011;161:40-46.

Murray TY, Mans J, Taylor MB. Human calicivirus diversity in wastewater in South Africa. J Appl Microbiol 2013;114:1843-1853.

Nainan O, Xia G, Vaughan G, Margolis HS. Diagnosis of hepatitis A virus infection: a molecular approach. Clin Microbiol Rev 2006;19:63-79.

Nainan OV, Margolis HS, Robertson BH, Balayan M, Brinton MA. Sequence analysis of a new hepatitis A virus naturally infecting cynomolgus macaques (*Macaca fascicularis*). J Gen Virol 1991;72:1685-1689.

Nejati A, Makvandi M, Samarbafzadeh A, Neisi N, Moradzadegan H. Molecular epidemiology of hepatitis A virus in patients in the Ahwaz region of Iran. J Med Virol 2012;84:582-586.

Netshikweta R. Optimisation and assessment of real-time PCR techniques for the detection of selected food- and waterborne viruses. [MSc dissertation]. Pretoria: University of Pretoria; 2011.

Normann A, Badur S, Onel D, Kilic A, Sidal, M, Larouze B, Massari V, Muller J, Flehmig B. Acute hepatitis A virus infection in Turkey. J Med Virol 2008;80:785-790.

Oba IT, Spina AMM, Saraceni CP, Lemos MF, Senhoras RCFA, Moreira RG, Granato CFH. Detection of hepatitis A antibodies by ELISA using saliva as clinical samples. Rev. Inst. Med. Trop. São Paulo 2000;42:1–9.

Okoh AI, Sibanda T, Gusha SS. Inadequately treated wastewater as a source of human enteric viruses in the environment. Int J Environ Res Public Health 2010;7:2620-2637.

Pan M, Yang X, Zhou L, Ge X, Guo X, Liu J, Zhang D, Yang H. Duck hepatitis A virus possesses a distinct type IV internal ribosome entry site element of picornavirus. J Virol 2012;86:1129-1144.

Pérez-Sautu U, Costafreda MI, Caylà J, Tortajada C, Lite J, Bosch A, Pintó RM. Hepatitis A virus vaccine escape variants and potential new serotype emergence. Emerg Infect Dis 2011a;17:734-737.

Pérez-Sautu U, Costafreda MI, Lite J, Sala R, Barrabeig I, Bosch A. Molecular epidemiology of hepatitis A virus infections in Catalonia, Spain, 2005-2009: circulation of newly emerging strains. J Clin Virol 2011b;52:98-102.

Pintó RM, Saiz JC. Enteric hepatitis viruses. In: Bosch A, editors. Human Viruses in Water. Oxford, UK: Elsevier; 2007.p.39-57.

Pintó RM, Costafreda I, Bosch A. Risk assessment in shelfish-borne outbreaks of hepatitis A. Appl Environ Microbiol 2009;75:7350-7355.

Pintó RM, Costafreda I, Pérez-Rodriguez FJ, D'Andrea L, Bosch A. Hepatitis A virus: State of the art. Food Environ Virol 2010;2:127-135.

Pintó RM, D'Andrea L, Pérez-Rodriguez FJ, Costafreda MI, Ribes E, Guix S, Bosch A. Hepatitis A virus evolution and the potential emergence of new variants escaping the presently available vaccines. Future Microbiol 2012;7:1-15

Prado T, Fumian TM, Miagostovich MP, Gaspar AM. Monitoring the hepatitis A virus in urban wastewater from Rio de Janeiro, Brazil. Trans R Soc Trop Med Hyg 2012;106:104-109.

Prata C, Ribeiro A, Cunha A, Gomes NCM, Almeida A. Ultracentrifugation as a direct method to concentrate viruses in environmental waters: virus-like particle enumeration as a new approach to determine the efficiency of recovery. J Environ Monitor 2012;14:64-70.

Provost PJ, Hilleman MR. Propagation of human hepatitis A virus in cell culture in vitro. Proc Soc Exp Biol Med 1979;160:213-21.

Racaniello VR. *Picornaviridae*: the viruses and their replication. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA, Roizman B, Straus SE, editors. Fields Virology, 5th ed. Philadelphia: Lippincott Williams and Wilkins; 2007. p. 795-805.

Richards GP, Watson MA, Kingsley DH. A SYBR green, real-time RT-PCR method to detect and quantitate Norwalk virus in stools. J Virol Methods 2004;116:63-70.

Robertson BH, Jansen RW, Khanna B, Totsuka A, Nainan OV, Siegl G, Widell A, Margolis SH, Isomura S, lto K, lshizu T, Moritsugu Y, Lemon SM. Genetic relatedness of hepatitis A virus strains recovered from different geographical regions. J Gen Virol 1992;73:1365-1377.

Rodríguez-Lázaro D, Cook N, Ruggeri FM, Sellwood J, Nasser A, Nascimento MSJ, D'Agostino M, Santos R, Saiz JC, Rzeżutka A, Bosch A, Gironés R, Carducci A, Muscillo M, Kovač K, Diez-Valcarce M, Vantarakis A, Bonsdorff C-H, Husman AMR, Hernández M, van der Poel WHM. Virus hazards from food, water and other contaminated environments. FEMS Microbiol Rev 2012;36:786-814.

Rosenblum LS, Mirkin IR, Allen DT, Stafford S, Hadler SC. A multistate outbreak of hepatitis A traced to commercially distributed lettuce. Am J Public Health 1990;80:1075–80.

Sánchez G, Pintó RM, Vanaclocha H, Bosch A. 2002. Molecular characterization of hepatitis A virus isolates from a transcontinental shellfish-borne outbreak. J. Clin. Microbiol 2002;40:4148–4155.

Sánchez G, Bosch A, Gomez-Mariano G, Domingo E, Pintó RM. Evidence for quasispecies distributions in the human hepatitis A virus genome. Virology 2003a;315:34-42.

Sánchez G, Bosch A, Pintó RM. Genome variability and capsid structural constraints of hepatitis A virus. J Virol 2003b;77:452-459.

Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson M-A, Roy SL, Jones JL, Griffin PM. Foodborne illness acquired in the United States-major pathogens. Emerg Infect Dis 2011;17:7-15.

Shieh YC, Khudyakov YE, Xia G, Ganova-raeva LM, Khambaty FM, Woods JW, Veazey JE, Motes ML, Glatzer MB, Bialek SR, Fiore AE. Molecular confirmation of oysters as the vector for hepatitis A in a 2005 multistate outbreak. J Food Prot 2007;70:145-150.

Shouval D. The immunological basis for immunization series: module 18: hepatitis A. Geneva: WHO. 2011.p.1-50. Available from URL: http://whqlibdoc.who.int/publications/2011/9789241501422_eng.pdf.

Sinclair RG, Jones EL, Gerba CP. Viruses in recreational water-borne disease outbreaks: a review. J Appl Microbiol 2009;107:1769-80.

Spradling PR, Martin A, Feinstone S. Hepatitis A Virus. In: Richman DD, Whitley RJ, Hayden FG, editors. Clinical Virology. Washington DC: ASM Press; 2009. p.1083-1108.

Standard Methods for the Examination of Water and Wastewater (2005). 21st edn, American Public Health Association/American Water Works Association/Water Environment Federation, Washington DC, USA.

Statistics South Africa. Mid-year population estimates. 2013;p.1-16. Available from URL http://www.statssa.gov.za/publications/P0302/P03022013.pdf.

Stene-Johansen K, Tjon G, Schreier E, Bremer V, Bruisten L, Ngui SL, King M, Pintó RM, Aragonès L, Mazick A, Corbet S, Sundqvist L, Blystad H, Norder H, Skaug K. Molecular epidemiological studies show that hepatitis A virus is endemic among active homosexual men in Europe. J Med Virol 2007;79:356-365.

Sulbaran Y, Gutierrez CR, Marquez B, Rojas D, Sanchez D, Navas J, Rovallo E, Pujol FH. Hepatitis A virus genetic diversity in Venezuela: exclusive circulation of subgenotype IA and evidence of quasispecies distribution in the isolates. J Med Virol 2010;82:1829-1834.

Sun Y, Laird DT, Shieh YC. Temperature-dependent survival of hepatitis A virus during storage of contaminated onions. Appl Environ Microbiol 2012;78:4976-4983.

Tang JW. Viral hepatitis. In: Shetty N, Tang JW, Andrews J. editors. Infectious Disease: Pathogenesis, Prevention, and Case Studies. West Sussex: Wiley Blackwell; 2009. Chapter 19. p.491-494.

Tapparel C, Siegrist F, Petty TJ, Kaiser L. Picornavirus and enterovirus diversity with associated human diseases. Infect Genet Evol 2013;14:282-293.

Taylor MB. Molecular epidemiology of South African strains of hepatitis A virus:1982-1996. J Med Virol 1997;51:273-279.

Taylor MB, Prozesky OW. Experimental infection of the Cape baboon (*Papio ursinus*) with hepatitis A virus. South African Journal of Science 1985;81:597-600.

Taylor MB, Cox N, Vrey MA, Grabow WOK. The occurrence of hepatitis A and astroviruses in selected river and dam waters in South Africa. Water Res 2001;35:2653-2660.

Theamboonlers A, Abe K, Thongmee C, Poovorawan Y. Complete coding sequence and molecular analysis of hepatitis A virus from a chimpanzee with fulminant hepatitis. J Med Primatol 2012;41:11–17.

Ticehurst JR, Racaniello VR, Baroudy BM, Baltimore D, Purcell RH, Feinstone SM. Molecular cloning and characterization of hepatitis A virus cDNA. Proc Natl Acad Sci USA 1983;80:5885-5889.

Tortajada C, de Olalla PG, Pintó RM, Bosch A, Cayla J. Outbreak of hepatitis A among men who have sex with men in Barcelona, Spain, September 2008-March 2009. Euro Surveill 2009;14.

Tortajada C, de Olalla PG, Diez E, Pintó RM, Bosch A, Pérez U, Sanz M, Caylà JA. Hepatitis A among men who have sex with men in Barcelona, 1989-2010: insufficient control and need for new approaches. BMC Infectious Diseases 2012;12:11.

Van Heerden J, Ehlers MM, Heim A, Grabow WOK. Prevalence, quantification and typing of adenoviruses detected in river and treated drinking water in South Africa. J Appl Microbiol 2005;99:234-242.

Vandamme P, Banatvala J, Fay O, Iwarson S, McMahon B, van Herck K, Shouval D, Bonanni P, Connor B, Cooksley G, Leroux-Roels G, Von Sonnenburg F. Hepatitis A booster vaccination: is there a need? Lancet 2003;362:1065-1071.

Vantarakis A, Papapetropoulou M. Detection of enteroviruses, adenoviruses and hepatitis A viruses in raw sewage and treated effluents by nested-PCR. Water Air Soil Pollut 1999;114:85-93.

Vaughan G, Rossi LMG, Forbi JC, de Paula VS, Purdy MA, Xia G, Khudyakov YE. Hepatitis A virus: host interactions, molecular epidemiology and evolution. Infect Genet Evol 2013a; doi: http://dx.doi.org/10.1016/j.meegid.2013.10.023.

Vaughan G, Xia G, Forbi JC, Purdy MA, Rossi LMG, Spradling PR, Khudyakov YE. Genetic relatedness among hepatitis A virus strains associated with food-borne outbreaks. PLoS ONE 2013b;8: e74546. doi:10.1371/journal.pone.0074546.

Venter JME. The incidence of hepatitis A virus in selected water sources and associated risk of infection in South Africa. [MSc dissertation]. Pretoria: University of Pretoria; 2004.

Venter JME, van Heerden J, Vivier JC, Grabow WOK, Taylor MB. Hepatitis A virus in surface water in South Africa: what are the risks? J Water Health 2007;5:229-240.

Vilaginès P, Sarrette B, Husson G, Vilaginès R. Glass wool for virus concentration at ambient water pH level. Water Sci Technol 1993;27: 229–306.

Villar LM, de Paula VS, Diniz-Mendes L, Guimaraes FR, Ferreira FF, Shubo TC, Miagostovich MP, Lampe E, Gaspar AM. Molecular detection of hepatitis A virus in urban sewage in Rio de Janeiro, Brazil. Lett Appl Microbiol 2007;45:168-173.

Vivier JC, Ehlers MM, Grabow WOK. Detection of enteroviruses in treated drinking water. Water Res 2004;38:2699-2705.

Wang H, Zheng H, Cao J, Zhou W, Yi Y, Jia Z, Bi S. Genetic diversity of hepatitis A virus in China: VP3-VP1-2A genes and evidence of quasispecies distribution in the isolates. PLoS ONE 2013;8:e74752. doi:10.1371/journal.pone.0074752.

Wheeler C, Vogt TM, Armstrong GL, Vaughan G, Weltman A, Nainan OV, Dato V, Xia G, Waller K, Amon J, Lee TM, Highbaugh-Battle A, Hembree C, Evenson S, Ruta MA, Williams IT, Fiore AE, Bell BP. An outbreak of hepatitis A associated with green onions. N Engl J Med 2005;353:890-897.

World Health Organization. Foodborne disease outbreaks: guidelines for investigation and control. 2008.p.1-162. Available from URL: http://www.who.int/foodsafety/publications/foodborne_disease/outbreak_guidelines.pdf.

Wyn-Jones P. The detection of waterborne viruses. In: Bosch A, editor. Human Viruses in Water. Oxford, UK: Elsevier; 2007.p.177-197.

Xu ZY, Wang XY. Live attenuated hepatitis A vaccines developed in China. Hum Vaccin Immunother 2013;10: PMID: 24280971; doi.org/10.4161/hv.27124.

APPENDIX A

BUFFERS AND REAGENTS

A.1 1M hydrochloric acid (HCl)

98 mL HCl (32%) (Merck)

The acid was added to 902 mL molecular grade water to make up a 1M solution.

A.2 1M sodium hydroxide

40 g NaOH (Merck)

The solid was dissolved in 1 L molecular grade water to make up a 1M solution.

A.3 Glycine 0.5% beef-extract buffer (pH 9.5)

3.75 g Glycine (Merck)

5 g Beef extract (BBL™ Becton Dickinson and Co., Sparks, MD)

The solids were dissolved in 800 mL molecular grade water. The pH was adjusted to pH 9.5 and the volume was made up to 1000 mL. The buffer was autoclaved and stored at 4°C.

A.4 Polyethylene glycol/sodium chloride precipitation solution

500 g PEG 8000 (Merck)

87 g NaCl (Merck)

The solids were dissolved in 450 mL molecular grade water on a shaking incubator at 50°C and at a speed of 135 rpm. After dissolution, the volume was made up to 1000 mL with molecular grade water to obtain a 5 x PEG/NaCl solution. The solution was autoclaved and stored at room temperature.

A.5 Phosphate buffered saline (pH 7.2)

The phosphate buffered saline was prepared using phosphate buffered saline tablets (Sigma-Aldrich, Co.) as per manufacturer's instructions. The solution was autoclaved and stored at 4°C.

A.6 Tris-glycine 1% beef extract buffer (pH 9.5)

12.1 g Tris Base (Roche)

3.8 g Glycine (Merck)

10 g Beef extract (BBL™)

The solids were dissolved in 800 mL molecular grade water. After dissolution, the volume was made up to 1000 mL and the pH adjusted to pH 9.5. The buffer was autoclaved and stored at 4°C.

APPENDIX B

B.1: Saïd R, Wolfaardt M, Taylor MB Molecular characterisation of hepatitis A virus strains from water sources in South Africa [Presentation]. Faculty Day, Faculty of Health Sciences, University of Pretoria 28-29 August 2012: HW Snyman Building North, Pretoria.

Introduction: Hepatitis A virus (HAV) can be transmitted through consumption of faecally-contaminated food and water. There are six genotypes of HAV with distribution being region specific. Genotype IA is the most prevalent genotype worldwide, except in South Africa (SA), where genotype IB predominates. Recently, however, no molecular epidemiological studies have been done to establish whether new or different genotypes are circulating in SA. The analysis of water samples provides a passive surveillance system as to what is circulating in the surrounding communities. In addition, it provides valuable baseline data for molecular source tracking in the event of an outbreak.
Aim: To characterise HAV strains found in selected South African surface waters.
Methods: Fourteen faecally-contaminated surface water samples (10L each) used for irrigation were collected in Limpopo. Eight of the samples were collected from rivers and the remaining six were borehole water. Viruses were recovered from the samples using the glasswool adsorption-elution method. The viruses were further concentrated using $PEG_{8000}/NaCl_2$ precipitation. Nucleic acid was extracted from each sample and tested for HAV by real-time reverse-transcriptase polymerase chain reaction (PCR). Complementary DNA (cDNA) was synthesised from recovered viruses using random primers. The full VP1 gene (900bp) and the VP1/P2B junction (390bp) were amplified by conventional PCR. The positive amplicons were sequenced directly using the ABI Prism BigDye® Terminator v3.1 Cycle sequencing Kit. The sequences were analysed and edited with Sequencher™ v4.10.1 and BioEdit v6.0.5, then aligned in MAFFT with reference sequences retrieved from Genbank. Phylogenetic analyses were performed in MEGA 5 using the neighbor-joining method. The tree generated was assessed by bootstrap analysis (1000 pseudoreplicates).
Results: Twelve out of 14 water samples tested positive for HAV. To date, seven HAV strains have been characterised. Sequence analyses from both the full VP1 and VP1/P2B regions revealed that all seven strains were subgenotype IB, with six out of seven strains showing >87% identity to subgenotype IB type strain, HM175.
Discussion/Conclusion: Phylogenetic analysis revealed that the HAV strains from the water samples grouped within genotype IB. Hence in addition to direct person-to-person transmission, faecally-contaminated water may be a source of HAV infection in SA.

B.2: Saïd R, Wolfaardt M, Taylor MB. Molecular characterisation of hepatitis A strains from water sources in South Africa [Presentation]. Fountain of Knowledge UP Water Institute Members' Forum Conference 9 November 2012 Law Building, Hatfield Campus, University of Pretoria

Hepatitis A virus (HAV) can be transmitted through the consumption of faecally-contaminated food and water. The virus has six genotypes with distribution being region specific. The analysis of water samples provides a passive surveillance system as to what genotype is circulating in the community and a valuable baseline data for molecular source tracking during outbreaks. The study aimed to characterise HAV strains found in selected South African surface waters. Fourteen faecally-contaminated irrigation water samples were collected from Polokwane and processed using a glasswool adsorption-elution method. Hepatitis A virus was detected by real-time reverse transcriptase-PCR and characterised by nucleotide sequence analyses. Twelve water samples tested positive for HAV of which seven characterised thus far, were >87% identical to subgenotype IB type strain, HM175. Based on these data, HAV genotype IB predominates in the water samples from Polokwane which may potentially contaminate irrigated fresh produce.

B.3: Saïd R**,** Wolfaardt M, Taylor MB. Molecular characterisation of hepatitis A virus strains from water sources in South Africa [Presentation]. 3rd Regional Conference of the Southern African Young Water Professionals 16-18 July 2013: Konservatorium, University of Stellenbosch, Stellenbosch, South Africa.

Hepatitis A virus (HAV) is the leading cause of acute hepatitis in the world and is frequently acquired through the ingestion of faecally-contaminated food and water. There are six genotypes of HAV infecting human (I, II and III) and non-human primates (IV, V and VI), with reports of new emerging variants. These genotypes have a distinct geographical distribution with genotype IA the most prevalent genotype worldwide, except in South Africa (SA), where genotype IB predominates. For the past 15 years, HAV was detected in various water sources and on fresh produce in SA. However, no molecular epidemiological studies have been done to establish whether new or different genotypes are circulating in SA. With the possible emergence of variants that escape the effective vaccine against HAV, analysing water samples provides a passive surveillance system as to what HAV types are circulating in the surrounding communities. In addition, it provides valuable baseline data for molecular source tracking in the event of an outbreak. The aim of this study was to characterise HAV strains found in selected SA surface waters. Twenty six surface water samples (10 L), used for irrigation of fresh produce or domestic purposes collected from four provinces, and 51 samples (500 mL) obtained from the inflow and outflow of wastewater plants were analysed. Viruses were recovered from the samples using the glasswool adsorption-elution method and then further concentrated using polyethylene glycol / sodium cholride precipitation. After automated nucleic acid extraction, samples were tested for HAV by real-time reverse-transcriptase polymerase chain reaction (PCR). Hepatitis A virus strains were characterised by nucleotide sequence analysis of the capsid gene VP1 and the VP1/P2B junction. The junction was used for rapid genotyping and the capsid gene for identification of possible vaccine escape mutants. The sequences were compared to reference sequences retrieved from Genbank and phylogenetic analyses were performed. Hepatitis A viruses were detected in 69% (18/26) of the surface water samples and in 37% (19/51) of the samples from the wastewater plants. Of the 37 strains detected, 29 were characterised. Phylogenetic analyses from both the full VP1 and VP1/P2B regions revealed that all 29 strains were subgenotype IB. The presence of genotype IB in the water sources confirms human faecal contamination. Hence in addition to direct person-to-person transmission, these faecally-contaminated water sources may be a potential transmission route of HAV infection and potential source of contamination of irrigated fresh produce in SA.

APPENDIX C

Saïd R, Wolfaardt M, Taylor MB Molecular characterisation of hepatitis A virus strains from water sources in South Africa. Water Science and Technology 2014;69:923-933.

Molecular characterisation of hepatitis A virus strains from water sources in South Africa

R. Saïd, M. Wolfaardt and M. B. Taylor

ABSTRACT

Hepatitis A virus (HAV) strains found in selected South African (SA) surface waters were characterised to establish what HAV types are circulating in the environment, thus reflecting circulation in the surrounding communities. Surface water samples used for irrigation or domestic purposes, and water samples from the outflow of wastewater plants were collected from six provinces. Viruses were recovered from the samples using a glass wool adsorption-elution method and then further concentrated using polyethylene glycol/sodium chloride precipitation. After automated nucleic acid extraction, samples were analysed for HAV by real-time reverse-transcriptase polymerase chain reaction. HAV strains were genotyped by nucleotide sequence analysis of the capsid gene VP1 and the VP1/P2B junction. HAVs were detected in 76% (16/21) of the surface water samples and in 37% (19/51) of the samples from the wastewater plants. Strains were characterised from 32 of the 35 samples and classified within genotype IB. The presence of genotype IB in the water sources confirms human faecal contamination. Hence, these faecally-contaminated water sources may be a potential transmission route of HAV infection and a potential source of contamination of irrigated fresh produce in SA.

Key words | genotypes, hepatitis A virus, irrigation water, South Africa

R. Saïd (corresponding author)
M. B. Taylor
M. Wolfaardt
Department of Medical Virology,
Faculty of Health Sciences,
University of Pretoria,
Private Bag X323,
Arcadia 0007 Pretoria,
South Africa
E-mail: *sadrachida19@yahoo.fr*

M. B. Taylor
National Health Laboratory Service,
Tshwane Academic Division,
Pretoria,
South Africa

INTRODUCTION

Hepatitis A virus (HAV) is the leading cause of acute viral hepatitis worldwide (Costafreda *et al.* 2006; Pintó *et al.* 2010). Infection with HAV can proceed with or without symptoms depending on the age of the individual at the time of the infection and the presence or not of an underlying chronic liver disease (Collier & Oxford 2006; Nainan *et al.* 2006; Pintó *et al.* 2010). Children younger than five years of age rarely present with symptoms, while in older children and adults the infection can be severe (Dotzauer 2008; Spradling *et al.* 2009). HAV is an enteric virus transmitted via the faecal-oral route either directly from person-to-person or indirectly through the ingestion of faecally contaminated food and water (Nainan *et al.* 2006; Pintó *et al.* 2010, 2012; Rodríguez-Lázaro *et al.* 2012). The virus may less commonly be transmitted through blood transfusion and blood products, the sharing of needles by drug abusers and sexual contact – particularly between male homosexuals (Nainan *et al.* 2006; Pérez-Sautu *et al.* 2011a). HAV has a long incubation period ranging from 15 to 50 days, during which time the virus is actively replicating in the liver and found in the blood and faeces (Dotzauer 2008; Pintó *et al.* 2012). An infected individual may shed up to 10^{11} HAV particles per gram of faeces just before the onset of symptoms and it is at that point that the individual is most infectious (Pintó *et al.* 2010). A high and long-lasting viraemia was also identified for HAV occurring from before the beginning of symptoms to an average of 6 weeks after symptoms have started (Bower *et al.* 2000; Pintó *et al.* 2012).

HAV is the sole member of the genus *Hepatovirus* within the family *Picornaviridae* (Costa-Mattioli *et al.* 2002; Spradling *et al.* 2009; Knowles *et al.* 2012). The infectious viral particle of HAV consists of a non-enveloped icosahedral capsid that encloses a 7.5 kb positive-sense single-stranded RNA genome (Costa-Mattioli *et al.* 2003; Dotzauer 2008; Spradling *et al.* 2009; Knowles *et al.* 2012). HAV is antigenically stable with a single serotype (Nainan *et al.* 2006; Pintó *et al.* 2012). Infection with the virus confers lifelong immunity and an effective vaccine was developed to protect individuals that have not been previously exposed (Nainan *et al.* 2006; Pérez-Sautu *et al.* 2011a; Pintó *et al.* 2012).

doi: 10.2166/wst.2013.799

However, recent publications have reported the detection of 'escape' mutants which could result in vaccination failure (Pérez-Sautu *et al.* 2011b; Pintó *et al.* 2012). The capsid gene VP1 and the junction VP1/P2B are used to genotype HAV strains (Nainan *et al.* 2006). The junction is used for rapid genotyping and the capsid gene for identification of possible vaccine escape mutants. Based on the nucleic acid sequence of the entire VP1 region, six genotypes, with their subdivisions (A and B), have been defined: I, II and III grouping HAVs associated with humans, and IV, V and VI associated with non-human primates (Costa-Mattioli *et al.* 2002). These genotypes have a distinct geographical distribution: genotype IA prevails worldwide except in South Africa (SA) and the rest of the African continent where genotype IB predominates (Robertson *et al.* 1992; Taylor 1997).

The incidence of HAV depends on socio-economic standards like sanitary conditions or access to clean water sources. In much of the developing world, where such conditions are relatively poor, infection occurs early in childhood and almost all adults are immune to the disease. In industrialised countries, HAV has a low to moderate endemicity. As a result, most adults who have not been infected during childhood are susceptible to the infection, leading to the occurrence of hepatitis A outbreaks (Pintó *et al.* 2010). Hepatitis A is hyperendemic in SA with the characteristics of both developing and industrialised countries (Robertson *et al.* 1992; Taylor 1997; Venter *et al.* 2007).

HAV is also one of the leading causes of viral foodborne infection (Bosch *et al.* 2011; Rodríguez-Lázaro *et al.* 2012). Food products can be contaminated directly or indirectly through the use of water contaminated with faeces containing HAV. This is mainly due to the fact that sewage treatment is unable to completely remove or inactivate viral particles, which in turn may contaminate surface waters that can be used to irrigate or wash food products (Pintó *et al.* 2010). The faecal contamination of water in shellfish-growing beds and irrigation water used for fresh produce has been implicated as the source of virus in foodborne outbreaks of hepatitis A. To date, the largest outbreak of hepatitis A recorded involved the consumption of sewage-contaminated clams and approximately 300,000 people became ill (Halliday *et al.* 1991; Nainan *et al.* 2006). The use of contaminated water to mix agrichemicals or inadequate sanitary facilities for food handlers at the point of harvest, at commercial outlets or at preparation facilities also poses a risk of spreading the virus.

HAV has been detected in surface water sources in SA (Taylor *et al.* 2001) and on fresh produce at the point of retail (Netshikweta 2011). The risk of infection following ingestion of water for the different communities has also been quantified (Venter *et al.* 2007). Since the studies of Taylor (1997) there have been no molecular epidemiological investigations in SA to establish whether or not different genotypes or mutant strains are emerging or circulating in the country. Analysing water samples provides a passive surveillance system as to what HAV types are circulating in the environment as well as in the surrounding communities. In addition, it provides valuable baseline data for molecular source tracking in the event of an outbreak. The aim of this study was therefore to characterise HAV strains detected in selected SA surface waters.

METHODS

Water samples

Surface water sample collection

From January 2012 to August 2012, 21 water samples (10 L) were collected monthly (except in February and July) from six sites near Polokwane, Limpopo (L) province, as part of the SA Water Research Commission (WRC) project K5/1875//4 (Figure 1). These samples originate from rivers and dams used to irrigate fresh produce on a large commercial

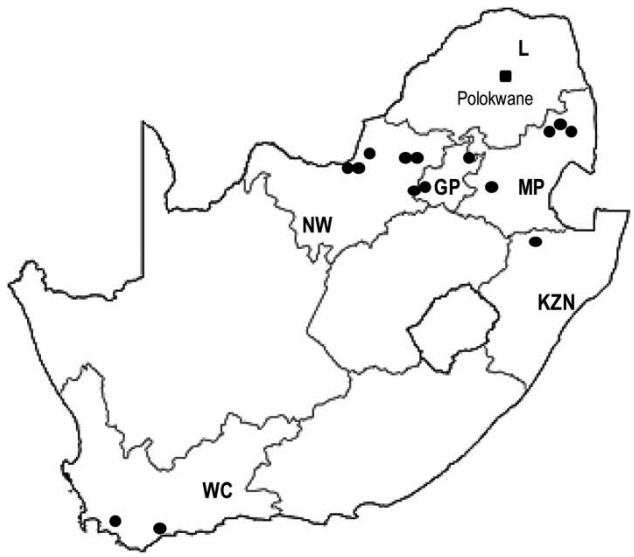

Figure 1 | Distribution of surface water (■) and wastewater (•) sampling sites among six provinces of South Africa. GP, Gauteng province; KZN, Kwazulu-Natal province; L, Limpopo province; MP, Mpumalanga province; NW, North West province; WC, Western Cape province (adapted from Murray *et al.* (2013)).

farm. The samples were transported in cooler bags with cold packs to the Department of Medical Virology, Faculty of Health Sciences, University of Pretoria. The temperature and pH of each sample were recorded upon arrival. The samples were either processed immediately or stored at 4 °C before being processed within 24 h of receipt. The sites from which samples were collected were: Bloed River (site 1) (two samples); River water drawn from below the Polokwane sewage works before the Sandrivier (site 5) (two samples); River water on the fresh produce farm (site 7) (five samples); Dam A at the pump site on the fresh produce farm (site 16B) (two samples); Dam B at the pump site on the fresh produce farm (site 17B) (five samples); River water from below the Seshego sewage works in the township (site 23) (two samples). Water was also collected from an irrigation pivot at the farm, at two points (three samples). Each sample was given a unique code, comprising site number and the date on which it was taken.

Water sampling at wastewater treatment works (WWTW)

From August 2010 to December 2011, a retrospective analysis was performed on 51 WWTW samples (500 mL) that were collected from the outflow of WWTW in three SA provinces: Gauteng (GP, nine samples), Mpumalanga (MP, 20 samples) and North West (NW, 16 samples) (Figure 1). In addition, sludge was collected in a treatment plant in KwaZulu-Natal (KZN, two samples) and in the Western Cape (WC, two samples) (Figure 1), two samples were of unknown provincial origin. The 51 WWTW samples were referred to the Department of Medical Virology, Faculty of Health Sciences, University of Pretoria for microbial indicator analysis. These samples were selected because they had (thermotolerant) faecal coliform counts greater than 10^6 colony forming units (cfu) per 100 mL (Murray *et al.* 2013).

Microbial indicator organisms count

The *Escherichia coli* and (thermotolerant) faecal coliform counts, expressed as cfu/100 mL were determined for the surface water samples and the WWTW samples, respectively. The choice of method for the different matrices was specified in the protocols of each sample type. The membrane filtration technique coupled with m-ColiBlue24® broth (Millipore Corp., Billerica, MA, USA) or M-FC agar (*Standard Methods for the Examination of Water and Wastewater* 2005) was used to determine the *E. coli* and (thermotolerant) faecal coliform counts, respectively.

Virus recovery and concentration

Surface water samples

Prior to recovery, 100 µL of mengovirus (5×10^4 copies/10 µL) was added to each sample as a process control to monitor the efficiency of viral recovery and downstream nucleic acid extraction. Viruses were recovered from the 10 L water samples using a glass wool adsorption-elution method based on the method of Vilaginès *et al.* (1993) as described by Mans *et al.* (2013). After elution, viruses were further concentrated to a volume of 10 mL in phosphate-buffered saline pH 7.4 (PBS; Sigma-Aldrich Co., St Louis, MO, USA), using polyethylene glycol (Amresco, Solon, OH, USA) and sodium chloride (Merck Schuchardt OHC, Hohenbrunn, Germany) (PEG_{8000}/NaCl) precipitation. The recovered virus concentrates were stored at −20 °C.

Wastewater treatment works samples

Viruses were concentrated from 75 or 100 mL WWTW samples to a volume of 2 mL in PBS pH 7.4, using the PEG_{8000}/NaCl precipitation. No prior viral recovery was performed as the sample volumes were too small. Virus concentrates were also stored at −20 °C.

Nucleic acid extraction

Total nucleic acid was extracted from 1 mL of virus concentrates using the automated platform of the MagNA Pure LC instrument (Roche Diagnostics GmbH, Mannheim, Germany) and the semi-automated platform of the NucliSENS® EasyMAG® instrument (BioMérieux, Marcy l'Etoile, France). The MagNA Pure LC Total Nucleic Acid Isolation Kit (Large volume) was used for viral extraction on the MagNA Pure LC System (Roche Diagnostics). Extraction on the MagNA Pure LC System (Roche Diagnostics) was done for virus concentrates of the WWTW samples, while the NucliSENS® EasyMAG® instrument (BioMérieux) was used for the surface water samples. In both methods, extracted nucleic acid was eluted in 100 µL of elution buffer and stored at −70 °C.

Mengovirus detection

Nucleic acid extracted from surface water samples was tested for mengovirus using a one-step real-time reverse transcriptase-polymerase chain reaction (RT-PCR) assay. The detection was performed using mengo@ceeramTools™

Kit (Ceeram s.a.s, La Chappelle-Sur-Erdre, France). The detection was performed on 5 μL of RNA extracts and the conditions of the assay were as follows: 45 °C for 10 min, 95 °C for 10 min and 45 cycles of 95 °C for 15 s, 60 °C for 45 s. The primers (Table 1) included in the kit amplify a 99 bp fragment within the 5′noncoding region (5′NCR) of the mengovirus genome (Pintó *et al*. 2009).

HAV detection

Nucleic acid (5 μL) from each sample (surface water and WWTW) was tested for HAV using a one-step real-time RT-PCR assay. Detection of HAV in surface water samples was performed using a hepatitisA@ceeramTools™ Detection Kit (Ceeram s.a.s) while detection in WWTW samples was performed using an RNA Ultrasense™ One-step qRT-PCR system kit (Invitrogen, Carlsbad, CA). Both assays were carried out in a LightCycler 2.0 (Roche Diagnostics) and used primers that amplified a 172 bp fragment within the most conserved region of the HAV genome, the 5′NCR (Table 1). The hepatitisA@ceeramTools™ Detection Kit (Ceeram s.a.s) includes an internal control to monitor the amplification process. The RNA Ultrasense™ One-step qRT-PCR system kit (Invitrogen) does not include an internal control, but a positive control, the cell culture adapted HAV strain HM175 43c was added to each run to monitor the amplification process. The conditions for the assay were different for each kit. For the hepatitisA@ceeramTools™ Detection Kit (Ceeram s.a.s) the conditions of the assay were as follows: 50 °C for 1 h, for the reverse transcription reaction, followed by a hot start of 95 °C for 10 min and 45 cycles of 95 °C for 15 s for denaturation, 60 °C for 1 min for annealing and 70 °C for 1 min for extension (Costafreda *et al*. 2006). The conditions of the assay for the RNA Ultrasense™ One-step qRT-PCR system kit (Invitrogen) were as follows: 50 °C for 45 min, for the reverse transcription reaction, followed by a hot start of 95 °C for 15 min and 50 cycles of 95 °C for 15 s for denaturation, 60 °C for 1 min for annealing and 65 °C for 1 min for extension (Netshikweta 2011).

HAV genotyping

Complementary DNA (cDNA) synthesis

Using random hexamers primers (Roche Diagnostics), cDNA (20 μL) was synthesised from 10 μL of RNA extracts from samples that tested positive for HAV. The synthesis was performed using a Thermo Scientific RevertAid™ Premium Reverse Transcriptase kit (Thermo Scientific, Waltham, MA). The protocol used was a slight modification of that of the manufacturer's instructions, as 50 units of the reverse transcriptase was used instead of 200 units.

Table 1 | Sequences of primers and probes used in this study

PCR	Region	Primer/Probe name	Nucleotide Sequence (5′ → 3′)	Position[a]	Reference
Real time – Mengo		Mengo110-F	GCG GGT CCT GCC GAA AGT	–	Pintó *et al*. (2009)
		Mengo209-R	GAA GTA ACA TAT AGA CAG ACG CAC AC	–	Pintó *et al*. (2009)
		Mengo147-probe	FAM-ATC ACA TTA CTG GCC GAA GC-TAMRA	–	Pintó *et al*. (2009)
Real time – HAV	5′NCR	HAV68-F	TCA CCG CCG TTT GCG TAG	68–85	Costafreda *et al*. (2006)
		HAV240-R	GGA GAG CCC TGG AAG AAA G	240–222	
		HAV150-probe	FAM-TTA ATT CCT GCA GGT TCA GG-TAMRA	150–169	
First round PCR	VP1	HAV1-F	GTT TTG CTC CTC TTT ATC ATG CTA TG	2167–2192	Costa-Mattioli *et al*. (2002)
		HAV2-R	AGT CAC ACC TCT CCA GGA AAA CTT	3308–3285	
	VP1/P2B	2870P-F	GAC AGA TTC TAC ATT TGG ATT GGT	2870–2893	Nainan *et al*. (2006)
		3381N-R	CCA TTT CAA GAG TCC ACA CAC T	3381–3360	
Second round PCR	VP1	2172P-F	GCT CCT CTT TAT CAT GCT ATG GAT	2172–2195	Costa Mattioli *et al*. (2002); Nainan *et al*. (2006)
		3125N-R	CCT GCA TTC TAT ATG ACT CT	3125–3106	
	VP1/P2B	2896P-F	CTA TTC AGA TTG CAA ATA CAA T	2,896–2,918	Nainan *et al*. (2006)
		3289N-R	AAC TTC ATT ATT TCA TGC TCC T	3,289–3,268	

[a]Positions are relative to the genome of the HAV strain HM175 (accession number M14707).

Amplification of the VP1 region

The full length VP1 region (900 bp) was amplified from cDNA (5 or 1 μL) by conventional PCR, using primers published by Costa-Mattioli *et al.* (2002) (Table 1).

The total volume of the reaction mix was 25 μL and contained the following: 5 × KAPATaq™ HotStart buffer (Kapa Biosystems, Cape Town, South Africa), 10 mM dNTPs, and 10 μM each of the forward and reverse primers, respectively, 25 mM $MgCl_2$, and 1.25 units of KAPATaq™ HotStart DNA polymerase (Kapa Biosystems). The conditions under which the VP1 region was amplified were: pre-denaturation at 94 °C for 3 min, followed by 45 cycles of 94 °C for 30 s, 50 °C for 30 s, 72 °C for 1 min 30 s, and, final extension at 72 °C for 5 min. If no positive amplicons were obtained after the first round of PCR, a second PCR was performed and the VP1 gene was amplified using internal primers (Costa-Mattioli *et al.* 2002; Nainan *et al.* 2006) (Table 1) and 0.5 μL of the first PCR reaction as template. The assay conditions for the second round of PCR are essentially the same, except for the annealing temperature, which decreased from 50 to 48 °C and the number of cycles that went from 45 to 35 cycles.

Amplification of the VP1/P2B junction

The VP1/P2B junction (~350 bp), was amplified from the same cDNA (5 μL) used to amplify the VP1 region using primers published by Nainan *et al.* (2006) (Table 1). The total volume of the reaction mix was 25 μL and contained the following: 5 × KAPATaq™ HotStart buffer (Kapa Biosystems), 10 mM dNTPs, and 10 μM each of the forward and reverse primers respectively, 25 mM $MgCl_2$, and 1.25 units of KAPATaq™ HotStart DNA polymerase (Kapa Biosystems). The conditions under which the VP1/P2B junction was amplified were: pre-denaturation at 94 °C for 3 min, followed by 35 cycles of 94 °C for 30 s, 45 °C for 45 s, 72 °C for 1 min, and, final extension at 72 °C for 5 min. If no PCR products were visible after the first round of PCR, a nested PCR was performed using internal primers (Nainan *et al.* 2006) (Table 1) and 0.5 μL of the first PCR reaction as template. The conditions for the second round of VP1/P2B amplification are similar to the ones used for the first round PCR, except for the annealing temperature that increased from 45 to 48 °C.

Sequencing and cloning

Products from the PCR were run through a 2% agarose gel by electrophoresis. The gel was stained with ethidium bromide for visualisation. Positive amplicons were purified directly or from the gel using the Zymogen DNA Clean & Concentrator-25™ Kit or the Zymoclean™ Gel DNA Recovery Kit (Zymo Research, Irvine, CA, USA), respectively. The primers used to amplify the different regions were used to sequence positive amplicons directly after amplification, in both the forward and reverse direction, using the ABI Prism BigDye® Terminator v3.1 Cycle Sequencing Kit on an ABI 3130 automated analyser (Applied Biosystems), as per the manufacturer's instructions. In cases where more than one strain was detected in a sample, PCR products were cloned using the cloneJET™ PCR cloning kit (Fermentas, Canada). After colony PCR, at least 10 clones were randomly selected for each sample and sequenced using the pJet 1.2/blunt specific primers.

Nucleotide sequence analyses

The sequences obtained were analysed and edited with Sequencher™ v4.10.1 and BioEdit v6.0.5. A BLAST search (http://0-blast.ncbi.nlm.nih.gov.innopac.up.ac.za/) was performed to verify the identity of the output sequences. Thereafter, the edited sequences were compared to each other and to reference sequences retrieved from GenBank by pairwise comparison. The sequences obtained, along with reference sequences for HAV and closely matched sequences from the output BLAST search, were aligned in MAFFT version 6 (http://0-mafft.cbrc.jp.innopac.up.ac.za/alignment/server/index.html).

Detection of novel or predicted mutation

After alignment, the sequences obtained were compared to reference sequences using the software BioEdit v6.0.5, to check for novel or predicted mutations at nucleotide and protein levels. Positions at which nucleotide mutations cause amino acid change potentially resulting in vaccine escape mutants (Pérez-Sautu *et al.* 2011b; Pintó *et al.* 2012) were carefully analysed.

Phylogenetic analysis

Phylogenetic analyses were performed in MEGA 5 using the two-parameter model of Kimura from the neighbour-joining method. Two trees were generated (one for VP1 and one for VP1/P2B) and assessed by bootstrap analysis (1000 pseudoreplicates). The genotype of each sequenced strain was determined by how well it clusters with reference sequences on the phylogenetic tree.

RESULTS AND DISCUSSION

HAV detection

HAV was detected in 76% (16/21) of the surface water samples (Table 2). A number of the surface water samples were faecally contaminated as the *E. coli* counts were between 1 and 1,000 cfu/100 mL (data not shown). Mengovirus was detected in 43% (9/21) of the surface water samples where HAV was detected. The process control was detected in samples where no HAV was detected (site 16B and 17B on 2012-01-31) indicating that the processing and nucleic acid extraction procedures were adequate and the negative HAV result was therefore a true negative. The process control was not detected in samples where HAV was detected (site 5 on 2012-08-13; site 7 from 2012-05-16 to 2012-08-13; site 16B on 2012-08-13; site 17B on 2012-05-16; site 23 on 2012-08-13; site Pivit on 2012-05-16; site KPL1 on 2012-06-27) still indicating that the method for recovery and detection was appropriate. However, the process control was not detected in samples where HAV was not detected (site 1 on 2012-08-13; site 17B from 2012-06-27 to 2012-08-13) (Table 2). The average recovery efficiency of mengovirus and HAV from turbid water using the glass wool adsorption-elution method is <1% and 5%, respectively (V. Ruhanya, unpublished data). The detection rate of HAV could therefore be an underestimation of the true prevalence of the virus in these surface waters.

HAV was detected in 37% (19/51) of the WWTW samples. A number of factors could account for a low detection rate in WWTW samples as opposed to surface water samples. No prior viral recovery was performed on the WWTW samples. Inhibitors, which were not completely removed during nucleic acid extraction, could have been co-concentrated during the PEG_{8000}/NaCl precipitation step. As the assay used for the WWTW samples lacked an internal amplification control, the assay was repeated with diluted nucleic acid, to circumvent the possible effect of inhibitors. Noroviruses and sapoviruses were detected in the 19 WWTW samples that tested positive for HAV and in some of the remaining 31 samples (Murray *et al.* 2013). This indicates that the nucleic acid extraction was efficient and HAV was possibly present at levels below the detection limit of the assay.

Table 2 | Summary of results of the analysis of surface water samples from the Polokwane sites for HAV

				Real-time RT-PCR analysis	
Site no.	**Date**	**Sample type**	**pH/ temp**	**Mengovirus**	**HAV**
1	2012-01-31	Bloed river	7/22	+	+
	2012-08-13		7.5/13.5	−	−
5	2012-01-31	River water below the Polokwane sewage works	?/24	+	+
	2012-08-13		7.5/13.5	−	+
7	2012-01-31	River water on fresh produce farm	7/22	+	+
	2012-03-27		7.5/16	+	+
	2012-05-16		7.5/11	−	+
	2012-06-27		7.5/11	−	+
	2012-08-13		7.5/10.5	−	+
16B	2012-01-31	Dam A at pump site on fresh produce farm	7/22	+	−
	2012-08-13		7.5/12	−	+
17B	2012-01-31	Dam B at pump site on fresh produce farm	7/22	+	−
	2012-03-27		7.5/16	+	+
	2012-05-16		7.5/11.5	−	+
	2012-06-27		7.5/11.5	−	−
	2012-08-13		8/12	−	−
23	2012-01-31	River water below Seshego sewage works in the township	7/23	+	+
	2012-08-13		7.5/19	−	+
Pivit	2012-03-27	Irrigation pivot	7.5/16	+	+
	2012-05-16		7.5/12	−	+
KPL1	2012-06-27	Irrigation pivot	7.5/13	−	+

Phylogenetic analysis

HAV was detected in 35 samples in total. Of the 35 samples, HAV strains could be genotyped from 32 samples: 15 surface water samples and 17 WWTW samples. The junction VP1/P2B could be amplified from 29 of the samples (13 surface water samples and 16 WWTW samples), while the VP1 region could only be amplified from 17 samples (11 surface water samples and six WWTW samples). Based on the phylogenetic analysis of both the VP1 genomic region and the VP1/P2B junction all the strains typed in this study clustered within genotype IB (Figure 2 and 3, respectively). In 1997, it was established that genotype IB predominated in SA (Taylor 1997). None of the strains sequenced in this study grouped within a genotype other than IB, suggesting that since 1997 genotype IB still predominates in SA. However, additional HAV strains will have to be characterised before a conclusive statement in this regard can be made.

The strains sequenced in this study form a unique SA cluster within genotype IB (Figure 2: South African cluster; Figure 3: South African cluster A). The clusters are supported by high bootstrap values (99% on Figure 2 and 97% on Figure 3). However, few strains grouped outside the unique cluster on Figure 3 (South African cluster B; sample 10/1010 indicated with ■) but were still within genotype IB. All the samples that grouped outside the unique cluster were either collected from the effluents of sewage treatment plants or from rivers into which sewage effluents are discharged.

When analysing samples from the environment, multiple strains are expected to be characterised from a single sample. During amplification and sequencing of the VP1/P2B junction, no more than one strain was detected in each sample except for one. On the other hand, multiple strains were detected in 15 out of the 17 samples from which the VP1 region could be amplified. Although HAV is antigenically stable, some regions of its genome are more variable than others and therefore more informative. It is mainly for that reason that genotyping of the different HAV strains using nucleotide sequence analyses of the full VP1 region, instead of other genomic regions commonly used, was proposed (Costa-Mattioli *et al.* 2002). Furthermore, some amino acids encoded from the VP1 region contribute to the immunodominant antigenic site targeted by most antibodies that are directed against HAV (Costa-Mattioli *et al.* 2003; Knowles *et al.* 2012) and these residues are prone to mutation (Pérez-Sautu *et al.* 2011b; Pintó *et al.* 2012). The predicted amino acid sequence of the strains characterised in this study showed great similarity (>98% in amino acid identity) to that of the type strain HM175 (accession number M14707). However, minor amino acid changes were detected. These changes did not occur on the residues that contribute to the immunodominant antigenic site. In addition, the mutations detected among the strains were not consistent, except one: a R298 K (position is relative to the genome of the HAV strain HM175). Even though the mutation occurred at the end of the protein sequence, it was consistent in all strains for which the VP1 region could be amplified. An arginine (*R*) to lysine (*K*) mutation in the amino acid sequence has little effect on the final conformation of the VP1 protein. Despite a small structural difference on the side chain, both amino acids are positively charged. Therefore the strains bearing this mutation will still be antigenically stable. Although no mutants were detected in this study, the presence of multiple strains of HAV that differ in the VP1 region suggests the importance of implementing a surveillance system in order to monitor the possible emergence of new variants. This would not only help in molecular source tracking in the event of an outbreak but in implementing vaccination policies if necessary.

In 2009, a study was published on the epidemiology of hepatitis A in Germany, with special emphasis on imported and non-imported HAV infections (Faber *et al.* 2009). The study detected and characterised HAV in tourists who were infected while visiting SA. These strains formed a distinct cluster when compared to strains detected in individuals who were infected either in Germany or in other regions of the world. In Figure 3, the strains isolated from the tourists were included (indicated with Δ on the tree). Phylogenetic analysis clustered the German strains within SA cluster A. The high bootstrap value at the node leading to the cluster indicates a close relationship with strains characterised in this study. This further confirms the distinct geographical distribution of HAV, a fact that can be exploited when investigating the source of an outbreak.

HAV strains were previously detected on tomatoes and lettuce at the point of retail in the Tshwane district (Netshikweta 2011). The detected strains (indicated with ▲ on Figure 3) also group within SA cluster A. The high bootstrap value at the node indicates that strains isolated from the water samples are similar to those isolated from the fresh produce. Furthermore, the strain isolated from the lettuce is identical to a strain detected in a WWTW sample (bootstrap value of 97%). The sample in question was collected from the Mpumalanga region (data not shown). Fresh produce can be contaminated at various points throughout the chain of production. Irrigation with surface water contaminated with faeces containing HAV and agrichemicals mixed with

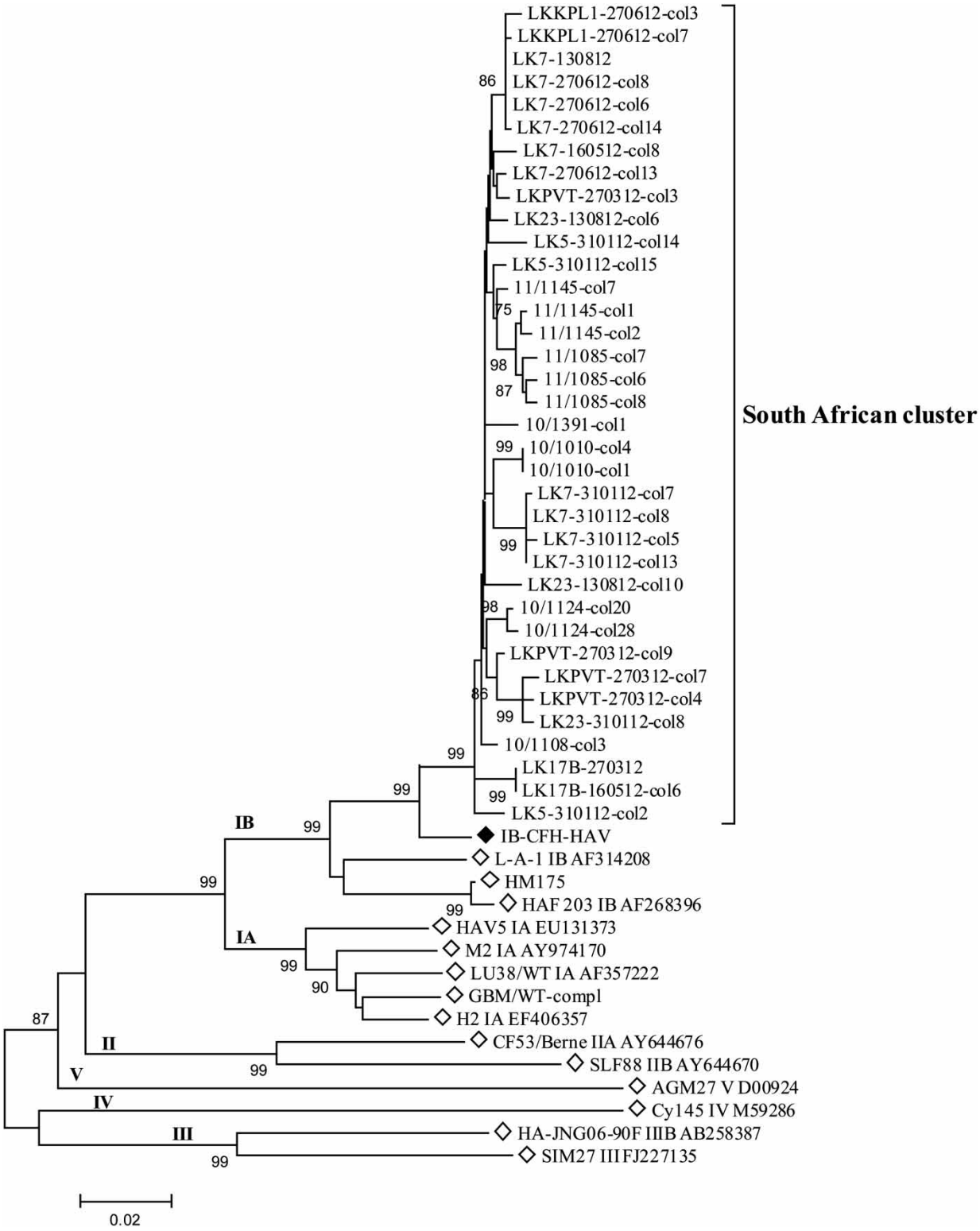

Figure 2 | Phylogenetic tree based on the sequence of the full VP1 region (900 bp) of strains characterised in this study. The tree was constructed using the Kimura two-parameter model of the neighbour-joining method, with 1000 bootstrap replicates; bootstrap values ≥75% are indicated. Strains indicated with ◊ or ♦ are reference strains or strains from previous studies that were retrieved from GenBank. Strains with 'col' and a number at the end represent clones from strains characterised in this study.

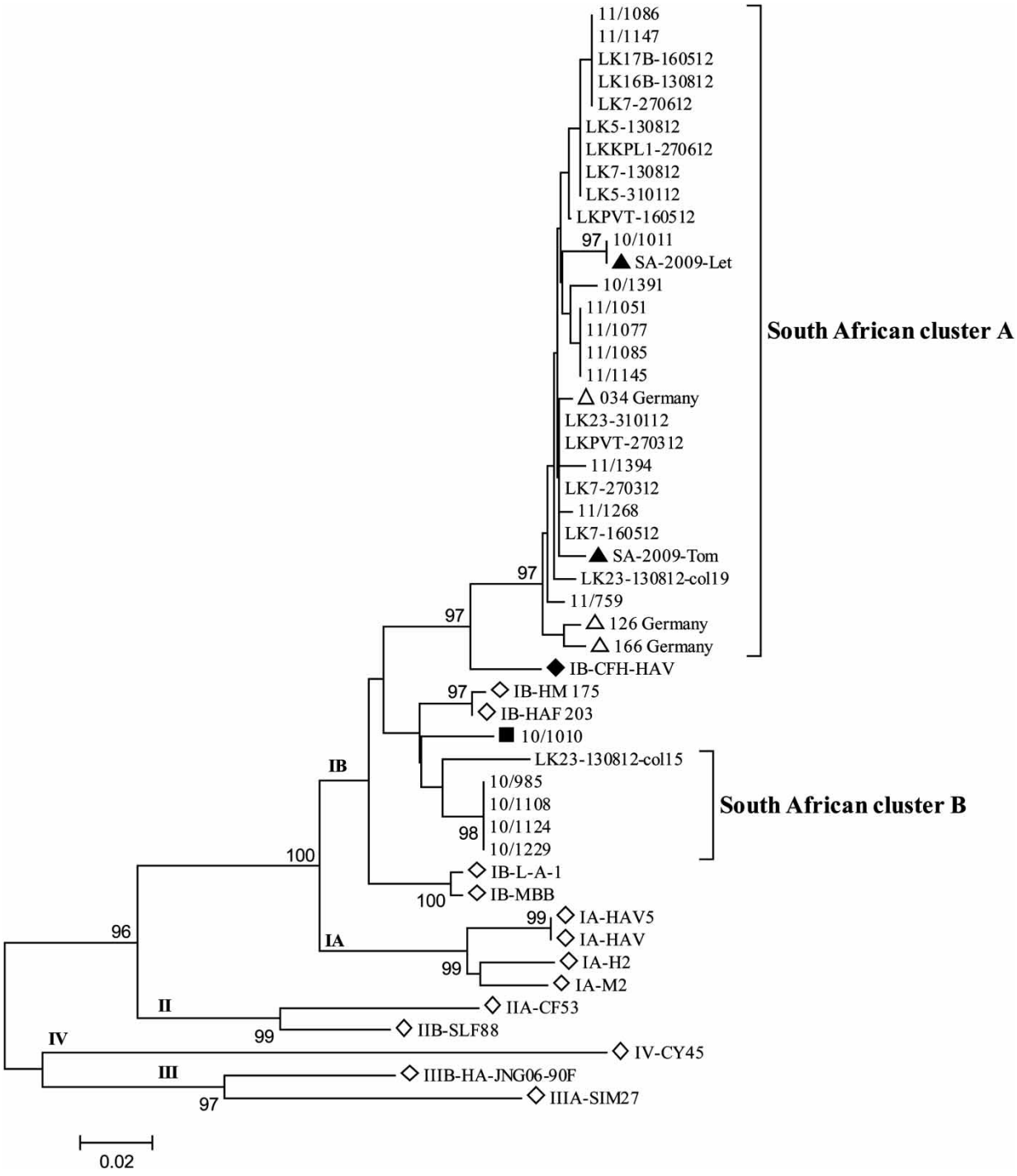

Figure 3 | Phylogenetic tree based on the sequence of the VP1/P2B junction (~ 350 bp) of strains characterised in this study. The tree was constructed using the Kimura two-parameter model of the neighbour-joining method, with 1000 bootstrap replicates; bootstrap values ≥75% are indicated. Strains indicated with ◊ or ◆, and, ▲ or Δ are reference strains or strains from previous studies that were retrieved from GenBank. Strains with 'col' and a number at the end represent clones from strains characterised in this study.

contaminated water are possible sources of contamination. The lack of proper washing facilities for food handlers on the farm, at commercial outlets or during preparation are additional sources of contamination. In the present case the exact source of contamination of the produce could not be defined, but the similarity with a strain detected in sewage effluent suggests that sewage treatment in the region of the WWTW sample does not remove all HAV particles.

CONCLUSIONS

HAV strains of genotype IB are still prevalent in SA as no other genotypes were detected in the water samples. The presence of genotype IB in the surface water sources used to irrigate fresh produce indicates human faecal contamination. Therefore, faecally-contaminated water sources are a potential vehicle of transmission of HAV infection and a potential source of contamination of irrigated fresh produce in SA. However, it would be essential to characterise HAV from infected individuals in order to assess whether or not the strains detected in the environment are clinically relevant.

ACKNOWLEDGEMENTS

The authors thank Dr Erika Du Plessis, from the Department of Microbiology and Plant Pathology, University of Pretoria, for the collection of the surface water samples. Rachida Saïd was supported by a Poliomyelitis Research Foundation post-graduate bursary. This study was funded, in part, by the National Research Foundation (NRF), and in part by an on-going solicited research project (K5/1875//4) funded by the WRC and co-funded by the Department of Agriculture, Forestry and Fisheries, South Africa.

REFERENCES

Bosch, A., Sánchez, G., Abbaszadegan, M., Carducci, A., Guix, S., Le Guyader, F. S., Netshikweta, R., Pintó, R. M., van der Poel, W. H. M., Rutjes, S., Sano, D., Taylor, M. B., van Zyl, W. B., Rodríguez-Lázaro, D., Kovač, K. & Sellwood, J. 2011 Analytical methods for virus detection in water and food. *Food Analytical Methods* **4**, 4–12.

Bower, W. A., Nainan, O. V., Han, X. & Margolis, H. S. 2000 Duration of viremia in hepatitis A virus infection. *The Journal of Infectious Diseases* **182**, 12–17.

Collier, L. & Oxford, J. 2006 Introduction to the hepatitis viruses. In: *Human Virology. A Text for Students of Medicine, Dentistry, and Microbiology* (L. Collier, J. Oxford & J. Pipkin, eds). Oxford University Press Inc., New York, pp. 159–160.

Costafreda, M. I., Bosch, A. & Pintó, R. M. 2006 Development, evaluation, and standardization of a real-time Taqman reverse transcription-PCR assay for the quantification of hepatitis A virus in clinical and shellfish samples. *Applied and Environmental Microbiology* **72**, 3846–3855.

Costa-Mattioli, M., Cristina, J., Romero, H., Perez-Bercof, R., Casane, D., Colina, R., Garcia, L., Vega, I., Glikman, G., Romanowsky, V., Castello, A., Nicand, E., Gassin, M., Billaudel, S. & Ferré, V. 2002 Molecular evolution of hepatitis A virus: a new classification based on the complete VP1 protein. *Journal of Virology* **76** (18), 9516–9525.

Costa-Mattioli, M., Di Napoli, A., Ferré, V., Billaudel, S., Perez-Bercoff, R. & Cristina, J. 2003 Genetic variability of hepatitis A virus. *Journal of General Virology* **84**, 3191–3201.

Dotzauer, A. 2008 Hepatitis A virus. In: *Encyclopedia of Virology* (B. W. J. Mahy & M. V. van Regenmortel, eds). 3rd edn, Academic Press, Slovenia, pp. 343–350.

Faber, M. S., Stark, K., Behnke, S. C., Schreier, E. & Frank, C. 2009 Epidemiology of hepatitis A infection, Germany, 2007–2008. *Emerging Infectious Diseases* **15**, 1760–1768.

Halliday, M. L., Kang, L. Y., Zhou, T. K., Hu, M. D., Pan, Q. C., Fu, T. Y., Huang, Y. S. & Hu, S. L. 1991 An epidemic of hepatitis A attributable to the ingestion of raw clams in Shanghai, China. *The Journal of Infectious Diseases* **164**, 852–859.

Knowles, N. J., Hovi, T., Hyypiä, T., King, A. M. Q., Lindberg, A. M., Pallansch, M. A., Palmenberg, A. C., Simmonds, P., Skern, T., Stanway, G., Yamashita, T. & Zell, R. 2012 Family Picornaviridae. In: *Virus Taxonomy. Ninth Report of the International Committee on Taxonomy of Viruses* (A. M. Q. King, M. J. Adams, E. B. Carstens & E. J. Lefkowitz, eds). Elsevier Academic Press, San Diego, pp. 855–880.

Mans, J., Netshikweta, R., Magwalivha, M., Van Zyl, W. B. & Taylor, M. B. 2013 Diverse norovirus genotypes identified in sewage-polluted river water in South Africa. *Epidemiology and Infection* **141**, 303–313.

Murray, T. Y., Mans, J. & Taylor, M. B. 2013 Human calicivirus diversity in wastewater in South Africa. *Journal of Applied Microbiology* **114**, 1843–1853.

Nainan, O., Xia, G., Vaughan, G. & Margolis, H. S. 2006 Diagnosis of hepatitis A virus infection: a molecular approach. *Clinical Microbiology Reviews* **19**, 63–79.

Netshikweta, R. 2011 *Optimisation and Assessment of Real-time PCR Techniques for the Detection of Selected Food- and Waterborne Viruses*. MSc dissertation, University of Pretoria, Pretoria, viewed 130428 http://upetd.up.ac.za/thesis/available/etd-05222012-143527.

Pérez-Sautu, U., Costafreda, M. I., Lite, J., Sala, R., Barrabeig, I., Bosch, A. & Pintó, R. M. 2011a Molecular epidemiology of hepatitis A virus infections in Catalonia, Spain, 2005–2009: circulation of newly emerging strains. *Journal of Clinical Virology* **52**, 98–102.

Pérez-Sautu, U., Costafreda, M. I., Caylà, J., Tortajada, C., Lite, J., Bosch, A. & Pintó, R. M. 2011b Hepatitis A virus vaccine escape variants and potential new serotype emergence. *Emerging Infectious Diseases* **17** (4), 734–737.

Pintó, R. M., Costafreda, I. & Bosch, A. 2009 Risk assessment in shellfish-borne outbreaks of hepatitis A. *Applied and Environmental Microbiology* **75**, 7350–7355.

Pintó, R. M., Costafreda, I., Pérez-Rodriguez, F. J., D'Andrea, L. & Bosch, A. 2010 Hepatitis A virus: State of the art. *Food and Environmental Virology* **2**, 127–135.

Pintó, R. M., D'Andrea, L., Pérez-Rodriguez, F. J., Costafreda, M. I., Ribes, E., Guix, S. & Bosch, A. 2012 Hepatitis A virus evolution and the potential emergence of new variants escaping the presently available vaccines. *Future Microbiology* **7**, 1–15.

Robertson, B. H., Jansen, R. W., Khanna, B., Totsuka, A., Nainan, O. V., Siegl, G., Widell, A., Margolis, H. S., Isomura, S., Ito, K., Ishizu, T., Moritsugu, Y. & Lemon, S. M. 1992 Genetic relatedness of hepatitis A virus strains recovered from different geographical regions. *Journal of General Virology* **73**, 1365–1377.

Rodríguez-Lázaro, D., Cook, N., Ruggeri, F. M., Sellwood, J., Nasser, A., Nascimento, M. S. J., D'Agostino, M., Santos, R., Saiz, J. C., Rzeżutka, A., Bosch, A., Gironés, R., Carducci, A., Muscillo, M., Kovač, K., Diez-Valcarce, M., Vantarakis, A., von Bonsdorff, C.-H., de Roda Husman, A. M., Hernández, M. & van der Poel, W. H. M. 2012 Virus hazards from food, water and other contaminated environments. *FEMS Microbiology Reviews* **36**, 786–814.

Spradling, P. R., Martin, A. & Feinstone, S. 2009 Hepatitis A Virus. In: *Clinical Virology* (D. D. Richman, R. J. Whitley & F. G. Hayden, eds). ASM Press, Washington, DC, pp. 1083–1108.

Standard Methods for the Examination of Water and Wastewater 2005 21st edn, American Public Health Association/ American Water Works Association/Water Environment Federation, Washington, DC, USA.

Taylor, M. B. 1997 Molecular epidemiology of South African strains of hepatitis A virus: 1982–1996. *Journal of Medical Virology* **51**, 273–279.

Taylor, M. B., Cox, N., Vrey, M. A. & Grabow, W. O. K. 2001 The occurrence of hepatitis A and astroviruses in selected river and dam waters in South Africa. *Water Research* **35**, 2653–2660.

Venter, J. M., van Heerden, J., Vivier, J. C., Grabow, W. O. K. & Taylor, M. B. 2007 Hepatitis A virus in surface water in South Africa: what are the risks? *Journal of Water and Health* **5**, 229–240.

Vilaginès, P., Sarrette, B., Husson, G. & Vilaginès, R. 1993 Glass wool for virus concentration at ambient water pH level. *Water Science and Technology* **27**, 299–306.

First received 16 August 2013; accepted in revised form 6 December 2013. Available online 20 December 2013

APPENDIX D

The Research Ethics Committee, Faculty Health Sciences, University of Pretoria complies with ICH-GCP guidelines and has US Federal wide Assurance.

* **FWA** 00002567, Approved dd 22 May 2002 and Expires 20 Oct 2016.
* **IRB** 0000 2235 IORG0001762 Approved dd 13/04/2011 and Expires 13/04/2014.

Universiteit van Pretoria
University of Pretoria

Faculty of Health Sciences Research Ethics Committee
Fakulteit Gesondheidswetenskappe Navorsingsetiekkomitee

DATE: 22/11/2012

NUMBER	**211/2012**
TITLE OF THE PROTOCOL	Molecular characterisation of hepatitis A virus strains from clinical and environmental sources in South Africa
PRINCIPAL INVESTIGATOR	**Student Name & Surname:** Rachida Saïd **Dept:** Medical Virology; University of Pretoria. **Cell:** 0764789760 **E-Mail:** sadrachida19@yahoo.fr
SUB INVESTIGATOR	Not Applicable
STUDY COORDINATOR	Not Applicable
SUPERVISOR	**Name & Surname:** Prof MB Taylor **E-Mail:** maureen.taylor@up.ac.za
STUDY DEGREE	MSc
SPONSOR COMPANY	Not Applicable
CONTACT DEATAILS OF SPONSOR	Not Applicable
SPONSORS POSTAL ADDRESS	Not Applicable
MEETING DATE	21/11/2012

The Protocol was approved on 21/11/2012 by a properly constituted meeting of the Ethics Committee subject to the following conditions:

1. The approval is valid for 2 years period [till the end of December 2014], and
2. The approval is conditional on the receipt of 6 monthly written Progress Reports, and
3. The approval is conditional on the research being conducted as stipulated by the details of the documents submitted to and approved by the Committee. In the event that a need arises to change who the investigators are, the methods or any other aspect, such changes must be submitted as an Amendment for approval by the Committee.

Members of the Research Ethics Committee:

Prof M J Bester	(female)BSc (Chemistry and Biochemistry); BSc (Hons)(Biochemistry); MSc(Biochemistry); PhD (Medical Biochemistry)
Prof R Delport	(female)BA et Scien, B Curationis (Hons) (Intensive care Nursing), M Sc (Physiology), PhD (Medicine), M Ed Computer Assisted Education
Dr NK Likibi	MBB HM – Representing Gauteng Department of Health) MPH
Dr MP Mathebula	(female)Deputy CEO: Steve Biko Academic Hospital; MBCHB, PDM, HM
Prof A Nienaber	(female) BA(Hons)(Wits); LLB; LLM; LLD(UP); PhD; Dipl.Datametrics(UNISA) – Legal advisor
Mrs MC Nzeku	(female) BSc(NUL); MSc(Biochem)(UCL, UK) – Community representative
Prof L M Ntlhe	MbChB (Natal) FCS (SA)
Snr Sr J Phatoli	(female) BCur(Eet.A); BTec(Oncology Nursing Science) – Nursing representative
Dr R Reynders	MBChB (Prêt), FCPaed (CMSA) MRCPCH (Lon) Cert Med. Onc (CMSA)
Dr T Rossouw	(female) MBChB (cum laude); M.Phil (Applied Ethics) (cum laude), MPH (Biostatistics and Epidemiology (cum laude), D.Phil

DR R SOMMERS; MBChB; MMed(Int); MPharmMed.
Deputy Chairperson of the Faculty of Health Sciences Research Ethics Committee, University of Pretoria

♦Tel:012-3541330 ♦Fax:012-3541367 / 0866515924 ♦E-Mail: manda@med.up.ac.za
♦Web: //www.healthethics-up.co.za ♦H W Snyman Bld (South) Level 2-34 ♦ Private Bag x 323, Arcadia, Pta, S.A., 0007