

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Hepatitis B Virus infection in the South Pacific

**A thesis presented in partial fulfilment of the
requirements for the degree of**

Master of Science

At

Massey University

Palmerston North, New Zealand.

Gabrielle Louise Harrison

1999

This thesis is dedicated to the scent of Sunshine.

Abstract

An exploratory study into the genetics of the hepatitis B virus and its human host in the South Pacific was undertaken to determine direction for future research. This virus is a serious health concern, especially for the indigenous people of this region. The DNA sequence of 14 complete and 2 partial virus genomes were obtained. The viral sequence mutations were investigated and compared with a collected database and current literature. Phylogenetic analysis of the viral sequences was carried out using version 4.64 of PAUP* and SplitsTree. Using the new sensitive method of sequence based typing, HLA-DPA1 allele's were typed in 51 unrelated Trobriand Islanders.

The viral genomes displayed a great deal of variation with many interesting mutations. The results highlight the affect of quasi-species distribution within a host. Phylogenetic analysis identified two hepatitis B genotypes within the South Pacific, HBV-C and HBV-D. However, the genotype common in northern Europe (HBV-A) was not found. The phylogenetic analysis presented a pattern of evolution that resembled that of its human host. The Trobriand Islanders were found to be an extremely homogeneous population, with 86% homogenous for the HLA-DPA1*02022 allele.

The study proved to be very informative, providing the directions of research we aimed for. The Hepatitis B samples demonstrated an interesting pattern of evolution that parallels that of its host supporting a co-evolutionary relationship between host and pathogen, thus hepatitis B appears to be indigenous in the South Pacific. We are presently establishing research to further investigate this pattern by analysing viral samples from Fiji. We have also established research that will investigate the rate of evolution of this virus. The sequenced based typing method proved to be very informative with the ability to detect new alleles. The allele frequency obtained from the Trobriand Islanders agreed with concurrent research and supports the fast-train model of migration into the Pacific. Further work in Fiji will continue with this theme of research as genetic analysis of Fiji has proved to be more complex.

Acknowledgments

David, the next pint in The Turf is on me

Cheers !

My thanks also to Rosalind Murray-M^cIntosh, for your guidance and perception.

This thesis would not of happened if it had not been for the kind provision of samples to work with. For this I must thank Prof Wolf Schiefenhovel of the Max Planck-Institute for Behavioural Physiology Von-der-Tann, Andechs Germany, for the Trobriand Islander samples and Dr Robert Hickson for being our liaison. Also thanks to Zlatibor Velickovic of the Wellington Medical School for providing HLA-DPA1 primers and training. Dr Bill Abbott of the Pathology Department of Auckland University and Dr Chris Moyes of the Whakatane Child Health and Hepatitis Foundation provided the hepatitis B virus samples. I would also like to thank Chris Moyes for his many words of wisdom throughout this project and beyond.

To the "plebs", "proliferates", and "boffins", Pete Lockhart who taught me how to do it "his way", Richard Winkworth who taught me how to do it the "right way" and Trish M^cLenachan who taught me how to do it "my way", cheers guys, you did an outstanding job in the face of adversary. Well done! Dan Jeffares and Bennet M^cComish, for 'Bennet's smile', I will always be appreciative, you guys can come check out my mailbox anytime ☺. I am most grateful to Anthony Poole for saving a thousand crashes 🍏, Matthew Phillips for his great late night company 🍷, Kerry Slack for being the only person awake when I needed company most; Mark Large, Leon Perrie, and Dave Thomas, for being such good sports (and food for gossip), Carmel Gillman for her serene sanity and Tony Roeven for "world peace". ☺

Mat, Maybee, Mr Pig, Huey, Duey, Leuy (R.I.P.), and George, my menagerie family,
words are no recompense.

Your support, understanding and adventurous soul were, are, and forever will be, my strength.

Thank you.

Contents

Abstract	ii
Acknowledgements	iii
Dedication	iv
Table of Content	v
List of Figures	viii
List of Tables	ix
List of Abbreviations	xi

Chapter one: Introduction

1.1	Introduction to the thesis	1
1.2	Literature review	3
1.2.1	The hepatitis B virus	3
1.2.2	Hepadnaviridae phylogenetic analysis	12
1.2.3	Hepatitis B virus and the people of the Pacific	22
1.2.4	The people of the Pacific	24
1.3	Aim	34

Chapter two: Methods and Software packages

2.1	DNA extraction	35
2.2	PCR amplification	36
2.2.1	HBV samples	37
2.2.2	Trobriand samples	39
2.2.3	Confirmation of PCR product	39
2.3	PCR template purification	39
2.3.1	Qiagen QIAquick™ PCR purification protocol	40
2.3.2	Qiagen QIAquick™ Gel extraction purification kit protocol	40
2.3.3	Quantification of DNA	41
2.4	Cloning	41
2.4.1	Ligation	41
2.4.2	Transformation	41
2.4.3	Recombinant DNA extraction	42

2.5	Sequencing	42
2.5.1	HBV samples	42
2.5.2	Trobriand samples	43
2.6	Sequence analysis	44
2.6.1	HBV samples	45
2.6.2	Trobriand samples	45
2.7.	Phylogenetic analysis	49
2.7.1	The data	50
2.7.2	Tree reconstruction	50
2.7.3	Models: transformations	51
2.7.4	Tree selection criterion	52
2.7.5	Search strategies	53
2.7.6	Tree testing and evaluation	54
2.8	Software packages and database	56

Chapter three: Results

3.1	Pacific HBV samples; sequences, features, and phylogenetic analysis	57
3.1.1	HBV sequencing results	57
3.1.2	Pacific HBV sequence features	69
3.1.3	Pacific HBV phylogenetic analysis results	74
3.2	Hepadnaviridae complete genome Genbank search	78
3.2.1	Database and Pacific samples phylogenetic analysis	78
3.2.2	Covarion structure analysis	84
3.3	HLA-DPA1 data	85

Chapter four: Discussion

4.1	Pacific HBV samples, sequence features	89
4.1.1	Transcriptional control element variations	89
4.1.2	Viral gene variations	92
4.1.3	Summary	94

4.2	Hepadnaviridae phylogenetic analysis	94
4.2.1	Pacific HBV samples	94
4.2.2	Covarian variation analysis	95
4.2.3	Summary	96
4.3	HLA-DPA1 allele typing data	96
4.4	Conclusion	97
4.5	Skills	98
	Bibliography	99
	Appendices	
	Appendix A	115
	Appendix B	124
	Appendix C	128
	Appendix D	132

List of figures

Chapter one: Introduction

1.1	Replication cycle of the hepatitis B virus	5
1.2	Human hepatitis B virus genomic characteristics	9
1.3	Trees representing phylogenetic relationship of Hepadnaviruses	16
1.4	Trees demonstrating different divergence patterns of Human hepatitis B virus	18
1.5	Stable tree topology of Human hepatitis B virus	20
1.6	Ambiguous tree topology of Hepadnaviridae	21
1.7	Map of the Pacific region	25
1.8	The Lapita language dispersal	29

Chapter two: Methods and Software packages

2.1	Example of MT Navigator PPC alignment of Trobriand HLA-DPA1 alleles	46
2.2	Examples of Mt Navigator PPC alignment of Trobriand HLA-DPA1 alleles (with electroperograms showing)	47
2.3	AMI Match Tools final analysis out put file for Trobriand HLA-DPA1 alleles	49

Chapter Three: Results

3.2	SplitsTree graph of the of the sixteen Pacific hepatitis B virus samples	75
3.3	SplitsTree graph showing phylogenetic network of the database and Pacific Hepadnaviridae samples	80
3.4	SplitsTree graph of the database and Pacific Hepadnaviridae samples	81

Appendix C

C.1	Neighbour-joining tree of Pacific samples and 30 database sequences	128
C.2	Neighbour-joining tree of the Pacific samples and 22 database HBV sequences	129
C.3	Maximum likelihood tree of the Pacific samples and 22 database HBV sequences	130
C.4	Neighbour-joining tree of Pacific samples with database genotypes HBV-C and HBV-D	131

Appendix D

D.1	MT Navigator layout with test sequences aligned to DPA1 exon 2 consensus	142
-----	--	-----

List of tables

Chapter one : Introduction

1.1	The four genomic transcripts of the hepatitis B virus	5
1.2	Transcriptional control elements of the hepatitis B virus as described by current literature	6
1.3	Prevalence of hepatitis B virus surface antigen in adults within Asia and the Pacific	24

Chapter two: Methods and Software packages

2.1	Materials solutions and cell strain and genotype used	36
2.2	Hepatitis B virus primers	38
2.3	ABI dye primer sequencing protocol	44
2.4	Covarian Inequality test character descriptions	55
2.5	Software packages and data bases	56

Chapter three: Results

3.1	Cloned fragments of the Pacific hepatitis B samples	58
3.2	Background information on the hepatitis B serum donors	59
3.3	The genomic sequences of the Pacific hepatitis B virus samples	60
3.4	The Pacific hepatitis B virus samples', sequence features	73
3.5	Bootstrap values for the Pacific hepatitis B virus samples' clusters (sites 1-910)	76
3.6	Nearest neighbour bootstrap one step rearrangement values for the Pacific hepatitis B virus samples	77
3.7	Hepadnaviridae sequences obtained from Genbank used in the phylogenetic analysis	79
3.8	Bootstrap values from the combined database and Pacific hepadnaviridae samples' trees	82
3.9	Hepadnaviridae sequence numbers	83
3.10	Covarian Inequality test values from the hepatitis B virus genotypes' analysis	84
3.11	Trobriand Islander HLA-DPA1 allele typing results	85
3.12	Trobriand Islander HLA-DPA1 allele sequences	86

Appendix A

A.1	HBeAg and HBcAg amino acid sequences	115
A.2	Polymerase amino acid sequences	116
A.3	HBsAg amino acid sequences	121
A.4	X protein amino acid sequences	123

Appendix B

B.1	Relevant background information the collected hepadnaviridae genomes from Genbank	124
-----	--	-----

List of Abbreviations

%	percent
ALT	alanine aminotransferase
bp	base pairs
ddH ₂ O	deionised and distilled water
DMSO	di-methyl sulphoxide
DNA	deoxyribonucleotide acid
dNTP	deoxyribonucleotide triphosphate
E.R.	endoplasmic reticulum
EDTA	ethylenediamine tetraacetic acid
g	grams
HBcAg	hepatitis B core antigen
HB _e Ag	hepatitis B <i>e</i> antigen
HB _s Ag	hepatitis B surface antigen
HBV	hepatitis B virus
HBV-A	hepatitis B virus genotype A
HBV-B	hepatitis B virus genotype B
HBV-C	hepatitis B virus genotype C
HBV-D	hepatitis B virus genotype D
HBV-E	hepatitis B virus genotype E
HBV-F	hepatitis B virus genotype F
HLA	Human leukocyte antigen
Kb	Kilobase
M	Molar
mg	milligram
MgCl	Magnesium Chloride
MHC	Major Histocompatibility Complex
μl	microliter
ml	milliliter
mM	microMolar
mtDNA	mitochondrial DNA
NaCl	Sodium Chloride
NNB	nearest-neighbour bootstrap
p mol	picomoles

PCR	polymerase chain reaction
RNA	ribonucleotide acid
SDS	Sodium Dodecyl Sulphate
ypb	years before present
°C	degrees Celsius

Chapter One: Introduction

The Haganemmons of Asistus three have the most impatient chromosomes of any life form in the Galaxy. Whereas most races are content to evolve slowly and carefully over thousands of generations, discarding a prehensile toe here, nervously hazarding another nostril there, the Haganemmons would do for Charles Darwin what a squadron of Aroturan stunt apples would have done for Sir Isaac Newton. Their genetic structure, based on the quadruple sterated octo-helix, is so chronically unstable that far from passing their basic shape onto their children, they will quite frequently evolve several times over lunch. But they do this with such reckless abandon, that, if they are sitting at a table, they are unable to reach a coffee spoon, they are liable, without a moments consideration, to mutate into something with far longer arms, but which is probably quite incapable of drinking the coffee. This, not unnaturally, produces a terrible sense of personal insecurity and a jealous resentment all stable life forms or "filthy rotten stinking samelings" as they call them. They justify this by claiming that as they have personally experienced what it is like to be virtually everybody else they can think of, they are in a very good position to appreciate all their worst points. This "appreciation" is usually military in nature, and is carried out with unmitigated savagery from the gun rooms of their horribly beweaponed "chameleoid" death flotilla. Experience has shown that the most effective way of dealing with any Haganemmons you may meet is to run away. Terribly fast!

(The Hitch-Hiker's Guide to the Galaxy; Adams 1978)

1.1 Introduction to the thesis

This thesis is an exploratory study into the genetics of a serious health concern in the South Pacific, the hepatitis B virus (HBV). It is a pilot study to determine directions for future research and the genetics of both the virus and host are investigated. The long term focus of the project is to investigate the genomic sequence of the virus and its phylogeny, as well as the history of the people of the South Pacific and their major histocompatibility complex (MHC) genes.

DNA sequences have become powerful tools in epidemiological studies, for example in Human Immuno-Deficiency disease and Hepatitis C virus (Holmes *et al.* 1995). They enable a greater understanding of the ecology of the diseases. This thesis uses this tool to study HBV in the South Pacific, some aspects of the genetics of both the virus and their human hosts within the Pacific will be investigated. Complete viral genomes from infected Polynesians will be sequenced and analysed, and a start will be made on human leukocyte antigen (HLA) typing, using a new sequence based method.

HBV is the smallest-double stranded DNA virus known to infect humans; it contains multiple overlapping reading frames and replicates through an RNA intermediate by reverse transcriptase. The lack of proof-reading ability of this reverse transcriptase accounts for the high mutational rate of this virus however, the overlapping reading frames place constraints on viable replicants, which restricts the viral evolution. Combined, these factors give HBV a unique mode of evolution which, along with the viral/host evolutionary relationship, remains unresolved (Miller and Robinson 1986; Gojobori *et al.* 1990; Bollyky and Holmes 1998). Co-evolution, cross-species, and recent transfer hypotheses have been suggested but current more sophisticated studies are inconclusive (Mandart *et al.* 1984; Orito *et al.* 1989; Norder *et al.* 1996; Bollyky and Holmes 1998). Clarification of these relationships is of interest, for not only is the molecular biology of the hepatitis B virus unique, but also as it is the major cause of hepatocyte carcinoma world wide, it is a serious health concern.

Infection with HBV can result in a range of clinical states ranging from asymptomatic to fatal. The differences appear to be related to genetics of both the host immune system and the virus (Chisari and Ferrari 1995; Milne *et al.* 1995; Gust 1996). It is estimated that 75% of the world's estimated carriers are within the Western Pacific and South East Asia. Here, where the virus is hyper-endemic (that is, there is a very high proportion of carriers in the population see 1.2.3), patterns of infection vary considerably between villages, cities, countries and ethnic groups (Gust 1996). This observation has relevance when it is noted that archaeological, linguistic and biological studies have shown human migrations through the Near and Far Oceania to be of two separate lineages (Hill and Serjeantson 1989; Kirch 1997; Spriggs 1997). To date little research has been undertaken to clarify the level of genetic admixture between these populations — a relevant issue when considering apparent mode of disease transmission and clearance.