

**Evidence of HIV-1 Adaptation to HLA-
Restricted Immune Responses at a Population
Level**

Corey Benjamin Moore

**This thesis is presented for the degree of Doctor of Philosophy of
Murdoch University, 2002**

I declare that this thesis is my own account of my research and contains as its main content work that has not previously been submitted for a degree at any tertiary education institution.

Corey Benjamin Moore
1 January 2003

Abstract

Selection of HIV-1 variants resistant to antiretroviral therapy is well documented. However, the selection *in vivo* of HIV-1 mutant species that can escape host immune system HLA class I restricted cytotoxic T-lymphocyte responses has, to date, only been documented in a few individuals and its clinical importance is not well understood. This thesis analyses the observed diversity of the HIV-1 reverse transcriptase protein in a well characterised, stable, HLA-diverse cohort of HIV-1 infected patients with over two thousand patient-years of observation. The results show that HIV-1 polymorphism is selected within functional constraints and is associated with specific HLA class I alleles. Furthermore, these associations significantly cluster along the sequence and tend to occur within known corresponding HLA-restricted epitopes. Absence of polymorphism is also HLA-specific and more often seen with common HLA alleles. Knowledge of HLA-specific viral polymorphisms can be used to model an individual's viral load from their HLA type and viral sequence. These results suggest that cytotoxic T-lymphocyte escape mutation in HIV-1 is critical to the host at an individual and population level as well as to short and long term viral evolution. This work provides new insights into viral-host interactions and has clinical implications for individualisation of HIV-1 therapy and vaccine design.

Table of contents

Abstract	i
Table of contents	iii
List of figures	vi
List of tables	viii
Abbreviations	x
General acknowledgements	xii
Chapter 1 Introduction	1
Chapter 2 Literature review	9
Introduction and scope	9
The human immunodeficiency virus - type 1	10
Genetics	10
Lifecycle	12
Protein structure and functions	18
HIV-1 pol gene	18
Reverse transcriptase protein	19
Structure	19
Function	21
Protease protein	23
Integrase protein	24
Envelope gene	24
Group-specific antigen gene	25
Matrix protein	26
Capsid protein	26
Nucleocapsid protein	26
p6 protein	26
Accessory genes and proteins	26
Tat	26
Rev	27
Nef	27
Vpr	28
Vpu	28
Vif	28
Tev	28
Host immune defences	29
HLA Class I presentation	29
Molecular structures	32
HLA class Ia genes	33
Pockets	36
Pocket A	36
Pocket B	37
Pocket C	37
Pocket D	37
Pocket E	38
Pocket F	38
TCR structure	38
TCR-HLA-peptide interaction	39
Clinical studies	40
Protection from infection	47

Disease progression	50
Host-viral interactions.....	53
CD8+ cytotoxic T-lymphocytes in HIV-1 disease.....	55
Acute infection.....	56
Clinically asymptomatic stage of disease	59
Advanced stage of disease	60
Immunodominance	61
Measures	63
CD4+ T-cell helper cells in HIV-1 disease.....	64
Cytotoxic T-lymphocyte and T-helper cell avoidance.....	66
Generic forms of avoidance	66
Epitope escape	68
Cytotoxic T-lymphocytes.....	69
T-helper cells	70
Effects of escape	71
Decreased HLA binding	72
TCR recognition.....	73
Antagonism.....	73
Insertions and deletions.....	73
Proteasome and TAP processing	74
Factors affecting escape	74
TCR adaptation	77
Limitations with current methods of measuring escape	78
Effect of antiretroviral therapy.....	79
Antiretroviral therapy.....	79
Drug mechanism	79
Reverse transcriptase inhibitors	80
Nucleoside reverse transcriptase inhibitors	80
Nucleotide reverse transcriptase inhibitors	80
Non-nucleoside reverse transcriptase inhibitors	81
Protease inhibitors.....	81
Effect on HIV-1	82
Viral load	82
Drug resistance mutations and drug failure	82
Effect on the immune control of HIV-1	85
CD8+ cytotoxic T-lymphocytes.....	86
CD4+ T-helper cells.....	87
Drug induced resistance mutations	87
Future treatments	88
Eradication	91
Evolution.....	92
Human evolution to pathogens	92
HIV-1 evolution to the human host	93
Summary	96
Chapter 3 Study group and general methods	97
Laboratory assay methods.....	102
HIV sequencing	102
HLA typing	102
Viral loads.....	102
Database management and data analysis software	103
Chapter 4 Determining the population consensus sequence.....	105

Introduction.....	105
Methods.....	105
Results.....	106
Discussion.....	109
Chapter 5 Functional constraints of HIV-1 reverse transcriptase.....	111
Introduction.....	111
Methods.....	112
Results.....	113
Discussion.....	116
Chapter 6 Determining putative escape mutation sites in HIV-1 reverse transcriptase.....	119
Introduction.....	119
Methods.....	124
Covariates considered before modelling.....	125
Step 1. Power calculations.....	126
Step 2. Removing covariates with small numbers of cases.....	126
Step 3. Removing covariates with high univariate <i>P</i> -values.....	126
Step 4. Logistic regression forward selection.....	127
Step 5. Logistic regression backwards elimination.....	127
Step 6. Calculating exact <i>P</i> -values.....	127
Results.....	128
Discussion.....	132
Chapter 7 Correlation between putative escape mutation sites and known HLA-restricted CTL epitopes.....	137
Introduction.....	137
Methods.....	138
Results.....	139
Discussion.....	142
Chapter 8 Clustering of putative escape mutations sites.....	143
Introduction.....	143
Methods.....	144
Results.....	144
Discussion.....	145
Chapter 9 Putative escape mutation sites flanking known HLA-restricted CTL epitopes.....	149
Introduction.....	149
Methods.....	151
Results.....	151
Discussion.....	154
Chapter 10 Determining putative HLA-restricted CTL epitopes from the putative escape mutation sites and HLA binding motifs.....	157
Introduction.....	157
Methods.....	159
Results.....	160
Discussion.....	168
Chapter 11 Adjustment for multiple comparisons.....	171
Introduction.....	171
Methods.....	172
Results.....	174
Discussion.....	177
Chapter 12 Overall test of significance.....	179
Introduction.....	179

Methods.....	179
Results.....	180
Discussion.....	181
Chapter 13 Population evolution of HIV-1.....	183
Introduction.....	183
Methods.....	184
Results.....	186
Discussion.....	187
Chapter 14 HLA molecular subtyping: improving the strength of associations and biological importance.....	191
Introduction.....	191
Methods.....	196
Results.....	196
Discussion.....	201
Chapter 15 Longitudinal analysis of putative escape mutation sites.....	205
Introduction.....	205
Methods.....	206
Results.....	207
Discussion.....	208
Chapter 16 Putative escape mutations and viral load: clinical implications.....	211
Introduction.....	211
Methods.....	213
Results.....	215
Discussion.....	220
Chapter 17 General discussion.....	227
Future directions.....	232
References.....	239
Appendices.....	289

List of figures

Figure 2.1. The HIV-1 genome.....	10
Figure 2.2. The HIV-1 virion.....	11
Figure 2.3. Lifecycle of HIV-1.....	13
Figure 2.4. Virions of HIV-1 budding from a CD4+ T-cell.....	13
Figure 2.5. HIV-1 cell infection and killing.....	14
Figure 2.6. Half-lives of HIV-1 in cellular and extracellular compartments.....	15
Figure 2.7. Schematic of the kinetics of HIV-1 in known compartments <i>in vivo</i>	15
Figure 2.8. HIV-1 tropism and coreceptor usage for cell entry.....	16
Figure 2.9. HIV-1 phenotypes throughout infection.....	17
Figure 2.10. Cellular dynamics of HIV-1 infection in CD4+ T cells.....	18
Figure 2.11. HIV-1 RT and its precursor proteins.....	19
Figure 2.12. Structure of the HIV-1 RT subunits.....	20
Figure 2.13. Schematic of HIV-1 RT creating an RNA-DNA hybrid from an RNA template.....	21
Figure 2.14. Ribbon model of the tertiary structure of HIV-1 protease homodimer.....	23
Figure 2.15. Mechanism for coreceptor function in HIV-1.....	25
Figure 2.16. The HLA class I and II pathways.....	30
Figure 2.17. Antigen-presenting cell, TCR and co-signal interactions.....	31

Figure 2.18. The organisation and location of the HLA complex.	33
Figure 2.19. Ribbon model of the tertiary structure of an HLA class Ia molecule with a bound peptide.	34
Figure 2.20. The structure of the 20 common amino acids.	35
Figure 2.21. The HLA molecule peptide-binding groove and peptide.	36
Figure 2.22. Interaction between peptide and HLA molecule.	37
Figure 2.23. Interaction between TCR and HLA-peptide complex.	39
Figure 2.24. CCR5 genotype, CCR5 expression and disease progression.	52
Figure 2.25. Schematic model of viral dynamics over the course of HIV-1 infection.	53
Figure 2.26. Immune responses and viral load throughout HIV-1 infection.	55
Figure 2.27. CTL evasion mechanisms used by HIV-1.	72
Figure 2.28. Schematic representation of the theoretical chance of a single induced mutation.	76
Figure 2.29. Distribution of HIV subtypes throughout the world.	94
Figure 4.1. The consensus sequence of the cohort compared to the HXB2CG clade B reference sequence in HIV-RT for amino acids 20 to 227.	108
Figure 5.1. Polymorphism frequency of catalytic, functional, stability and external residues in HIV-1 RT.	113
Figure 5.2. The polymorphism frequencies in pre-antiretroviral treatment HIV-1 RT sequences at amino acid positions 20 to 227 of HIV-1 RT.	115
Figure 6.1. The exact <i>P</i> -value versus the original inexact <i>P</i> -value for association of HLA-A and HLA-B with polymorphism.	128
Figure 6.2. The difference in inexact and exact <i>P</i> -values for association of HLA-A and HLA-B with polymorphism.	129
Figure 6.3. The polymorphism frequencies, known HLA restricted CTL epitopes and HLA associations with polymorphism for all positions between 20 and 227 of HIV-1 RT.	131
Figure 9.1. Predicted cleavage site versus polymorphism frequency in HIV-1 RT prior to antiretroviral treatment.	152
Figure 11.1. HLA correction factor versus HLA allele frequency.	174
Figure 11.2. The correction factor versus the number of models for which an HLA allele was not eliminated.	175
Figure 11.3. Minimum amount of polymorphism and HLA frequency required to detect an association.	178
Figure 13.1. The number of negative associations versus HLA phenotypic frequency.	185
Figure 14.1. HLA-B5 molecular subtypes.	194
Figure 14.2. HLA-B35 molecular subtypes.	195
Figure 14.3. Distribution of I135x in HIV-1 RT sequence in all HLA-B5 individuals.	199
Figure 14.4. Distribution of D177x in HIV-1 RT sequence in all HLA-B35 individuals.	200
Figure A.1. The standard opening form of EpiPop.	314
Figure A.2. The standard form with advanced options.	314
Figure A.3. The Advanced form.	315
Figure A.4. The Population Residues Summary form.	317
Figure A.5. The Single Residue Analysis form.	321
Figure A.6. The Multiple Site Analysis form.	323
Figure A.7. The Results Display form.	324
Figure A.8. Schematic diagram of flow of information from the various departments to the system.	331

Figure A.9. The Import and Update form.	333
Figure A.10 The Filename form.	333
Figure A.11. The Defaults form.	337
Figure A.12. The Edit (patient) Profile form.	340
Figure A.13. The Edit (patient) Visits form.	341
Figure A.14. The Edit (patient) Drugs form.	342
Figure A.15. The Edit Drug Names form.	343
Figure A.16. The Change Drug Code form.	343
Figure A.17. The Patient Trials and Complications edit form.	344
Figure A.18. The Edit Drug Interactions form.	345
Figure A.19. The MS Access query wizard.	346
Figure A.20. The Main Form for the system.	347
Figure A.21. The patient Search form (with the Main Form to show that selecting the Select button updates the Main Form).	348
Figure A.22. An example preview of all three patient reports.	349
Figure A.23. The Patient Profile Report.	350
Figure A.24. Patient Visit Report.	351
Figure A.25. Opportunistic Infections Report.	351

List of tables

Table 2.1. Viral and cellular dynamics of HIV-1.	14
Table 2.2. The incidence of established HIV-1 infection after different types of exposure.	40
Table 2.3. Studies that have shown association with susceptibility to, or protection from, HIV-1 infection and disease progression.	42
Table 2.4. Proteins encoded in viruses that have been shown to interfere with the HLA/MHC class I and II pathways.	67
Table 2.5. Some well-defined examples of HIV-1 CTL escape mutations in individuals.	71
Table 2.6. Conditions that favour mutation.	75
Table 3.1. Demographics for the cohort.	98
Table 3.2. HLA allele frequencies in the cohort and the estimated frequency of each molecular subtype of each broad HLA group.	99
Table 6.1. Multivariate analysis of positions 135 and 162 with and without adjustment for polymorphisms at other positions.	124
Table 7.1. The actual and expected numbers of positive associations (based on all amino acid positions) within their corresponding epitopes between residues 20 and 227.	140
Table 7.2. The actual and expected numbers of positive associations (based on those amino acid positions with power to detect an association) within their corresponding epitopes between residues 20 and 227.	141
Table 8.1. Clustering of HLA associations with polymorphism in HIV-1 RT.	145
Table 9.1. Predicted proteasome cleavage sites.	152
Table 9.2. The number of putative cleavage sites corresponding to putative CTL escape mutation sites.	154
Table 10.1. Known CTL epitopes in HIV-1 RT and how well they fit their corresponding HLA motifs.	161

Table 10.2. Putative epitopes best fitting around the positive putative CTL escape mutation sites, based on their corresponding HLA motifs.....	163
Table 10.3. Putative epitopes best fitting around the negative putative CTL escape mutation sites, based on their corresponding HLA motifs.....	165
Table 10.4. Putative CTL epitopes best fitting the negative CTL escape mutation sites after putative escape, based on their corresponding HLA motifs.	166
Table 11.1. Multiple comparisons HLA correction factors for each HLA allele.	175
Table 12.1. Association between HLA and polymorphism in HIV-1 RT.	180
Table 13.1. The frequency and the number of negative associations found for each HLA allele.	187
Table 14.1. Change in strength of associations of polymorphic sites associated with HLA-B5 after considering the molecular subtypes.....	197
Table 14.2. Change in strength of associations of polymorphic sites associated with HLA-B35 after considering the molecular subtypes.....	198
Table 15.1. Summary of mutations that occurred over time at sites with HLA associated polymorphism in those individuals with and without the corresponding HLA allele.	208
Table 16.1. Association between viral load and a polymorphism at position 166 in individuals with HLA-A11.	215
Table 16.2. Association between viral load and a polymorphism at position 162 in individuals with HLA-B7.	216
Table 16.3. Association between viral load and a polymorphism at position 43 in individuals with HLA-A3.	217
Table 16.4. Comprehensive viral load model.	219
Table A.1. HLA alleles remaining after each step of the modelling process.	291
Table A.2. Relative increase over time of polymorphism in the positive associations.	303
Table A.3. Relative decrease over time of polymorphism in the negative associations.	307
Table A.4. Inexact, exact and corrected <i>P</i> -values, and odds ratio for each association with an exact <i>P</i> -value less than 0.05.	309
Table A.5. Associations considered and eliminated from the comprehensive viral load model.....	311

Abbreviations

A	Adenine
Ab	Antibody
AH	Ancestral haplotype
AIDS	Autoimmune deficiency syndrome
APC	Antigen presenting cell
C	Cytidine
CCR	CC chemokine receptor
CDR	Complementary-determining region
cps	Copies
CTL	Cytotoxic T-lymphocyte
CXCR	CXC chemokine receptor
D	Diversity (gene segment)
DBMS	Database management system
df	Degrees of freedom
DNA	Deoxyribonucleic acid
DT	Decay time
EBV	Epstein-Barr virus
eCTL	Effector CTL
env	Envelope
ER	Endoplasmic reticulum
G	Guanine
gag	Group-specific antigen
HBV	Hepatitis B virus
HCMV	Human cytomegalovirus
HIV	Human immunodeficiency virus
HLA	Histocompatibility leukocyte antigen
HSV	Herpes simplex virus
IC	Inhibitory concentration
IDU	Immunodeficiency unit
IFN	Interferon
IL	Interleukin
IQR	Inter-quartile range
J	Joining (gene segment)
kDa	Kilo Dalton
LTNP	Long term nonprogressor
LTR	Long terminal repeat
MCMV	Murine cytomegalovirus
MHC	Major histocompatibility complex
MIP	Macrophage inflammatory protein
mL	Millilitre
M-tropic	Macrophage tropic
n	Number
NC	Non-conservative
nef	"Negative factor"
NK	"Natural killer"
NNRTI	Non-nucleoside reverse transcriptase inhibitor
NRTI	Nucleoside reverse transcriptase inhibitor
NS	Non-synonymous

NSI	Non-syncytium inducing
NtRTI	Nucleotide reverse transcriptase inhibitor
OR	Odds ratios
P	<i>P</i> -value
PBMC	Peripheral blood mononuclear cell
PCR	Polymerase chain reaction
PI	Protease inhibitor
RANTES	"Regulated on activation, normal T expressed and secreted"
RNA	Ribonucleic acid
RRE	Rev-responsive element
RT	Reverse transcriptase
SDF	"Stromal-derived factor"
SHIV	Simian-human immunodeficiency virus
SI	Syncytium-inducing
SIV	Simian immunodeficiency virus
STI	Structured treatment interruptions
T	Thymine
TAP	"Transporter associated with antigen presentation"
TCR	T-cell receptor
tRNA	Transfer RNA
T-tropic	T-cell tropic
V	Variable (gene segment)
WABMRDB	Western Australian bone marrow registry database
WAHIVCS	Western Australian HIV cohort study

For amino acid abbreviations see Figure 2.20 on page 35.

General acknowledgements

First of all I would like to thank Professor Simon Mallal, without whom I may not have started this PhD. His foresight to accept me - a computer science graduate - as firstly an honours student and then encourage me to start a PhD has taken me down a path I never expected. His enthusiasm and perseverance to research in the specific field of this thesis and his ability to verbalise the relevant information from the literature with the forethought of how it relates clinically was instrumental to the publication of the core of this thesis in *Science*.

I would also like to thank Professor Ian James for accepting me as a student in immunology despite the fact that I did not have a statistical or immunological background. His remarkable ability to make complicated statistics make sense and to translate biological problems into statistics made my introduction to statistics easier and provided the comprehensive and unique statistics required for this thesis.

Thanks also go to my fellow PhD student Dr Mina John. Her ability to effortlessly put anything in writing with the relevant references is admired. Without her cooperation, my results may not have been published in *Science*. The long hours we spent together working on similar problems and sharing the ups and downs that come with doing a PhD helped make my life a little easier.

Further thanks go to my parents, who like me, started this thesis without a background in immunology, but managed to make sense of it. Their assistance also made my life during my PhD more comfortable.

My appreciation also goes to Dr David Nolan and Professor Frank Christiansen, who both kindly offered to proofread this thesis and provided critical review; and to

Professor Sarah Rowland-Jones who made helpful comments on the published manuscript.

Finally, I am indebted to the willingness of the people in the Department of Clinical Immunology and Biochemical Genetics and the participants of the WA HIV Cohort Study for providing the masses of data to analyse. The collective resources available through the infrastructure provided by Murdoch University and Royal Perth Hospital have been instrumental to this thesis.