

**IMPACT OF IMMUNE-DRIVEN SEQUENCE
VARIATION IN HIV-1 SUBTYPE C GAG-
PROTEASE ON VIRAL FITNESS AND
DISEASE PROGRESSION**

SUBMITTED BY: JACLYN WRIGHT

SUPERVISED BY: PROFESSOR THUMBI NDUNG'U

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in Virology in the Nelson R. Mandela School of Medicine, College of Health Sciences, University of KwaZulu-Natal, Durban.

2011

This project represents original work done by the author and where others have made contributions it has been acknowledged in the text. The experimental work described in this thesis was performed in the HIV Pathogenesis Programme Laboratory, Hasso Plattner Research Laboratory, and Africa Centre Laboratory, in the Doris Duke Medical Research Institute, Nelson R. Mandela School of Medicine, University of KwaZulu-Natal, Durban, South Africa, under the supervision of Professor T. Ndung'u.

J. Wright: _____ Date: _____

Professor T. Ndung'u: _____ Date: _____

As the candidate's supervisor I agree to the submission of this thesis.

Professor T. Ndung'u: _____ Date: _____

DECLARATION

I, Jaclyn Wright, declare that

- (i) The research reported in this thesis, except where otherwise indicated, is my original work.
- (ii) This thesis has not been submitted for any degree or examination at any other university.
- (iii) This thesis does not contain other person's data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
- (iv) This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
 - (a) Their words have been re-written but the general information attributed to them has been referenced;
 - (b) where their exact words have been used, their writing has been placed inside quotation marks, and referenced.
- (v) Where I have reproduced a publication of which I am author, co-author or editor, I have indicated in detail which part of the publication was actually written by myself alone and have fully referenced such publications.
- (vi) This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections,

Signed: _____ Date: _____

PUBLICATIONS AND PRESENTATIONS

Peer-reviewed publications:

1. **Wright, J.K.**, Brumme, Z.L., Carlson, J.M., Heckerman, D., Kadie, C.M., Brumme, C.J., Wang, B., Losina, E., Miura, T., Chonco, F., van der Stok, M., Mncube, Z., Bishop, K., Goulder, P.J.R., Walker, B.D., Brockman, M.A., Ndung'u T. 2010. Gag-Protease-Mediated Replication Capacity in HIV-1 Subtype C Chronic Infection: Associations with HLA Type and Clinical Parameters. *Journal of Virology* 84(20): 10820-10831.
2. **Wright, J.K.**, Novitsky, V., Brockman, M.A., Brumme, Z.L., Brumme, C.J., Carlson, J.M., Heckerman, D., Wang, B., Losina, E., Leshwedi, M., van der Stok, M., Maphumulo, L., Mkhwanazi, N., Chonco, F., Goulder, P.J.R., Essex, M., Walker, B.D., Ndung'u, T. 2011. Influence of Gag-Protease-Mediated Replication Capacity on Disease Progression in Individuals Recently Infected with HIV-1 Subtype C. *Journal of Virology* 85(8): 3996-4006.
3. Radebe, M., Nair, K., Chonco, F., Bishop, K., **Wright, J.K.**, van der Stok, M., Bassett, I.V., Mncube, Z., Altfeld, M., Walker, B.D., Ndung'u, T. 2011. Limited immunogenicity of HIV CD8+ T-cell epitopes in acute clade C virus infection. *Journal of Infectious Diseases* 204(5): 768-76.
4. **Wright, J.K.**, Brumme, Z.L., Julg, B., van der Stok, M., Mncube, Z., Gao, X., Carlson, J.M., Goulder, P.J., Walker, B.D., Brockman, M.A., Ndung'u, T. 2011. Lack of Association between HLA Class II Alleles and In Vitro Replication Capacities of Recombinant Viruses Encoding HIV-1 Subtype C Gag-Protease from Chronically Infected Individuals. *Journal of Virology* 86(2): 1273-1276.
5. **Wright, J.K.**, Naidoo, V.L., Brumme, Z.L., Prince, J.L., Claiborne, D.T., Goulder, P.J.R., Brockman, M.A., Hunter, E., Ndung'u, T. 2012. The impact of HLA-B*81-associated mutations in HIV-1 Gag on viral replication capacity. *Journal of Virology*, in press.

Conference presentations:

1. **Wright, J.K.**, Brockman, M.A., Brumme, Z.L., Goulder, P.J.R., Walker, B.D., Ndung'u, T. Impact of Immune-Driven Sequence Variation in HIV-1 Subtype C Gag-Protease on Viral Fitness and Disease Progression. AIDS Vaccine Conference 2009, 19-22 October, 2009, Paris, France, Abstract 164471.

2. **Wright, J.K.**, Brockman, M.A., Brumme, Z.L., Reddy, S., Maphumulo, L., Bishop, K., Goulder, P.J.R., Walker, B.D., Ndung'u, T. Evidence of Higher HIV-1 Subtype B Gag-Protease-Mediated Viral Replicative Fitness Compared to HIV-1 Subtype C. South African Immunology Society meeting, 9-11 December, 2009, Cape Town, South Africa, Abstract LB-61.

3. **Wright, J.K.**, Brockman, M.A., Brumme, Z.L., Nair, K., Chonco, F., van der Stok, M., Mncube, Z., Bishop, K., Walker, B.D., Ndung'u, T. Gag-Protease-mediated fitness is associated with viral set point in individuals acutely infected with HIV-1 subtype C. Seventeenth Conference on Retroviruses and Opportunistic Infections, 16-19 February, 2010, San Francisco, USA, Abstract F-112.

4. **Wright, J.K.**, Brockman, M.A., Brumme, Z.L., Carlson, J.M., Heckerman, D., Goulder, P.J.R., Walker, B.D., Ndung'u, T. Mutations in Gag-Protease associated with changes in replication capacity in chronic HIV-1 subtype C infection. XVIII International AIDS Conference, 18-23 July, 2010, Vienna, Austria, Abstract TUPE0035.

5. **Wright, J.K.**, Novitsky, V., Brockman, M.A., Brumme, Z.L., Brumme, C.J., Carlson, J.M., Leshwedi, M., Ndung'u, T. Influence of Gag-Protease-Mediated Replication Capacity on Disease Progression in Individuals Recently Infected with HIV-1 Subtype C. 5th South African AIDS conference, 7-10 June, 2011, Durban, South Africa, Abstract 77.

6. **Wright, J.K.**, Brockman, M.A., Brumme, Z.L., Carlson, J.M., Heckerman, D., Naidoo, V.L., Jaggernath, M., Goulder, P.J.R., Walker, B.D., Ndung'u, T. The impact of Gag and specific HLA-B*81-associated Gag mutations on viral replication capacity. 6th IAS Conference on HIV Pathogenesis, Treatment and Prevention, 17-20 July, 2011, Rome, Italy, Abstract MOPE061.

STATEMENT

The following publications, fully referenced here, have been reproduced in part non-continuously throughout the thesis:

1. **Wright, J.K.**, Brumme, Z.L., Carlson, J.M., Heckerman, D., Kadie, C.M., Brumme, C.J., Wang, B., Losina, E., Miura, T., Chonco, F., van der Stok, M., Mncube, Z., Bishop, K., Goulder, P.J.R., Walker, B.D., Brockman, M.A., Ndung'u T. 2010. Gag-Protease-Mediated Replication Capacity in HIV-1 Subtype C Chronic Infection: Associations with HLA Type and Clinical Parameters. *Journal of Virology* 84(20): 10820-10831.

2. **Wright, J.K.**, Novitsky, V., Brockman, M.A., Brumme, Z.L., Brumme, C.J., Carlson, J.M., Heckerman, D., Wang, B., Losina, E., Leshwedi, M., van der Stok, M., Maphumulo, L., Mkhwanazi, N., Chonco, F., Goulder, P.J.R., Essex, M., Walker, B.D., Ndung'u, T. 2011. Influence of Gag-Protease-Mediated Replication Capacity on Disease Progression in Individuals Recently Infected with HIV-1 Subtype C. *Journal of Virology* 85(8): 3996-4006.

The candidate performed the experiments described in these papers, and where others made contributions it has been duly acknowledged in the text. The candidate drafted these publications in full and they were reviewed by co-authors.

J. Wright: _____

Date: _____

Professor T. Ndung'u: _____

Date: _____

ACKNOWLEDGEMENTS

I would like to thank:

- Professor Ndung'u for his supervision, encouragement, and support throughout the project
- Dr. Mark Brockman and Dr. Zabrina Brumme for their hospitality and mentorship
- Jennifer Sela and Pamela Rosato for demonstrating laboratory assays
- Staff and students at the HIV Pathogenesis Programme, Hasso Plattner Research Laboratory and Africa Centre for assistance, in particular Keshni Hiramani, Taryn Green, and Vanessa Naidoo
- Collaborators as acknowledged in the text
- Dr. Johannes Viljoen and Africa Centre laboratory for providing access to tissue culture and sequencing facilities
- Sinikithemba Study, HIV Pathogenesis Programme Acute Infection Study, and Tshedimoso Study teams
- Clinic/hospital staff as well as the study participants
- The National Research Foundation, Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology and Harvard University, and the HIV Pathogenesis Programme for financial assistance

ABSTRACT

Introduction

Understanding of the viral and host factors that determine time for progression to acquired immunodeficiency syndrome (AIDS) in individuals infected with human immunodeficiency virus type 1 (HIV-1) could aid in the design of an effective HIV-1 vaccine. Human leukocyte antigen (HLA) class I profile is strongly and consistently associated with differential rates of HIV-1 disease progression, however the mechanisms explaining this are not well understood. It has been hypothesised that “protective” HLA alleles select escape mutations in functionally important epitopes in the conserved group specific antigen (Gag) protein resulting in HIV-1 attenuation, which may result in slower disease progression. Many of the studies investigating the fitness cost of Gag escape mutations have concentrated on a few pre-selected mutations and have not assessed fitness consequences in the natural sequence background. Furthermore, the majority of studies have focussed on HIV-1 subtype B, while HIV-1 subtype C is the most prevalent subtype worldwide. Therefore, in the present study, a large population-based approach and clinically-derived Gag-protease sequences were used to comprehensively investigate the relationship between immune-driven sequence variation in Gag, viral replication capacity and markers of disease progression in HIV-1 subtype C chronic infection. The influence of Gag function on HIV-1 disease progression was further investigated in early HIV-1 subtype C infection. It was also hypothesised that Gag may contribute significantly to overall HIV-1 fitness and towards fitness differences between HIV-1 subtypes.

Materials and Methods

Recombinant viruses encoding Gag-protease, derived from antiretroviral naïve HIV-1 subtype C chronically (n=406) and recently (n=60) infected patients as well as a small subset of HIV-1 subtype B chronically infected patients (n=25), were generated by electroporation

of an HIV-inducible green fluorescent protein (GFP)-reporter T cell line with plasma-derived *gag-protease* PCR products and linearised *gag-protease*-deleted NL4-3 plasmid. The replication capacities of recombinant viruses, as well as intact HIV-1 isolates from peripheral blood mononuclear cells of patients chronically infected with HIV-1 subtype C (n=16), were assayed in the GFP-reporter T cell line by flow cytometry. Replication capacity was defined as the slope of increase in percentage infected cells from days 3-6 following infection, normalised to the growth of a wild-type NL4-3 control. Replication capacities were related to patient HLA alleles and markers of disease progression (viral load, CD4⁺ T cell count, and rate of CD4⁺ T cell decline in chronically infected patients, and viral set point and rate of CD4⁺ T cell decline in recently infected patients). Replication capacities were compared between isolates and recombinant viruses encoding Gag-protease from the same isolates, as well as between HIV-1 subtype B and C recombinant viruses matched for viral load and CD4⁺ T cell count. Bulk sequencing of patient-derived *gag-protease* amplicons was performed and mutations were identified that were significantly associated with altered viral replication capacity. The fitness effect of some of these mutations was directly tested by site-directed mutagenesis followed by assay of the mutant viruses.

Results

In HIV-1 subtype C chronic infection, protective HLA-B alleles, most notably HLA-B*81 (p<0.0001), were associated with lower replication capacities. HLA-associated mutations at low entropy sites (*i.e.* conserved sites) in or adjacent to Gag epitopes were associated with lower replication capacities (p=0.02), especially the HLA-B*81-associated 186S mutation in the TL9 epitope (p=0.0001). The fitness cost of this mutation was confirmed in site-directed mutagenesis experiments (p<0.001), and the co-varying mutations tested did not significantly compensate for this fitness cost. Replication capacity also correlated positively

with baseline viral load ($p < 0.0001$) and negatively with baseline CD4⁺ T cell count ($p = 0.0004$), but not with subsequent rate of CD4⁺ T cell decline ($p = 0.73$).

In HIV-1 subtype C recent infection, replication capacities of the early viruses did not correlate with subsequent viral set points ($p = 0.37$) but were significantly lower in individuals with below median viral set points ($p = 0.03$), and there was a trend of correlation between lower replication capacities and slower rates of CD4⁺ T cell decline ($p = 0.09$). Overall, the proportion of host HLA-specific Gag polymorphisms in or adjacent to epitopes was negatively associated with replication capacities ($p = 0.04$) but host HLA-B-specific polymorphisms were associated with higher viral set points ($p = 0.01$), suggesting a balance between effective Gag CD8⁺ T cell responses and viral replication capacity in influencing viral set point.

A moderate statistically significant correlation was found between the replication capacities of whole isolates and their corresponding Gag-protease recombinant viruses ($p = 0.04$) and the replication capacities of the subtype C recombinant viruses were significantly lower than that of the subtype B recombinant viruses ($p < 0.0001$). The subtype-specific difference in the consensus amino acids at Gag codons 483 and 484 was found in site-directed mutagenesis experiments to largely contribute to the fitness difference between subtypes, possibly by influencing budding efficiency.

Discussion

The data support that protective HLA alleles, in particular HLA-B*81, attenuate HIV-1 through HLA-restricted CD8⁺ T cell-mediated selection pressure on Gag. Results suggest that viral replication capacity determined by sequence variability in Gag-protease has an impact on HIV-1 disease progression, but also indicate that a balance between HLA-driven fitness costs and maintenance of effective CD8⁺ T cell responses is important in determining

clinical outcome. Gag-protease was observed to significantly contribute to overall HIV-1 replication capacity and variability in this region between HIV-1 subtypes B and C is suggested to partly explain the difference in viral fitness between these subtypes. Specific mutations in Gag-protease associated with viral attenuation were identified and it was also observed that mutations in conserved Gag regions carried the greatest cost to HIV-1 replication capacity. Overall, the data support the concept of, and may assist in the rational design of, an HIV-1 vaccine in which immune responses are directed towards several conserved epitopes, particularly in Gag, with the aim to constrain immune escape (thereby maintaining effective CD8⁺ T cell responses) and attenuate HIV-1 (in the event of partial escape), resulting in slower disease course and reduced HIV-1 transmission at the population level.

TABLE OF CONTENTS

DECLARATION	II
PUBLICATIONS AND PRESENTATIONS	III
STATEMENT	V
ACKNOWLEDGEMENTS	VI
ABSTRACT	VII
TABLE OF CONTENTS	XI
LIST OF FIGURES	XVIII
LIST OF TABLES	XXI
ABBREVIATIONS	XXII
CHAPTER 1 - INTRODUCTION	2
1.1 History and epidemiology of human immunodeficiency virus (HIV) type 1 infection.....	2
1.2 Structure of HIV-1	3
1.2.1 Genome organisation	3
1.2.2 HIV-1 structure.....	4
1.3 HIV-1 replication cycle: role of Gag	5
1.3.1 Virus entry	7
1.3.2 Reverse transcription.....	7
1.3.3 Uncoating.....	9
1.3.4 Nuclear import.....	9
1.3.5 Integration.....	10

1.3.6 Transcription and nuclear export.....	10
1.3.7 Assembly	11
1.3.8 Budding.....	13
1.3.9 Maturation.....	13
1.4 HIV-1 Pathogenesis.....	14
1.4.1 Acute infection	14
1.4.2 Chronic infection.....	18
1.4.3 Advanced disease	19
1.4.4 Host immune responses and viral evasion.....	20
1.4.4.1 Innate immunity	20
1.4.4.1.1 Dendritic cells.....	20
1.4.4.1.2 Macrophages.....	20
1.4.4.1.3 Natural killer (NK) cells	21
1.4.4.1.4 Host restriction factors	21
1.4.4.2 Humoral immunity	22
1.4.4.3 Cell-mediated immunity	23
1.4.4.3.1 CD8+ T cells.....	23
1.4.4.3.2 CD4+ T cells.....	26
1.5 Therapy and preventative strategies.....	27
1.5.1 Antiretroviral therapy and prevention.....	27
1.5.2 Vaccines.....	30
1.6 Factors influencing the rate of disease progression.....	32
1.6.1 Spectrum and measurement of disease progression rate	32
1.6.3 Viral and host factors	33
1.6.3.1 Innate immune responses.....	34

1.6.3.2 Humoral immune responses	35
1.6.3.3 Cell-mediated immune responses.....	36
1.6.3.3.1 CD4+ T cell responses	36
1.6.3.3.2 CD8+ T cell responses	36
1.6.4 HLA class I alleles	38
1.6.4.1 Genetic associations	38
1.6.4.2 HLA: mechanisms of control	39
1.6.4.2.1 Interactions with NK cells	39
1.6.4.2.2 Immune activation.....	39
1.6.4.2.3 CD8+ T cell responses	40
1.6.4.2.3.1 Immunodominance and quality	40
1.6.4.2.3.2 Gag-focussed specificity	41
1.6.4.2.3.2.1 Epitope restriction specificity	41
1.6.4.2.3.2.2 Gag T cell responses associated with viral control	41
1.6.4.2.3.2.3 HLA-associated mutations in Gag associated with viral control	43
1.6.5 Viral fitness.....	44
1.6.5.1 Fitness costs of HLA-associated Gag escape mutations.....	44
1.6.5.2 Viral fitness and disease progression	46
1.6.5.2.1 Whole isolate fitness	46
1.6.5.2.2 Gag-mediated viral fitness.....	47
1.6.5.2.3 Gag-mediated viral fitness/CD8+ T cell response balance.....	48
1.7 Measurement of viral fitness or replication capacity	49
1.7.1 Starting material	49
1.7.1.1 Mutant viruses	49
1.7.1.2 Recombinant viruses	51
1.7.1.3 Whole isolates	52

1.7.2 Primary cells versus T cell lines.....	53
1.7.3 Single and multiple cycle fitness assays.....	54
1.7.3.1 Single cycle assays.....	54
1.7.3.2 Multiple cycle assays.....	54
1.7.3.2.1 Pairwise growth competition assays.....	54
1.7.3.2.2 Parallel assays.....	55
1.7.4 Choice of assay for the present study.....	56
1.8 The present study: rationale, aims and objectives.....	56
CHAPTER2 – MATERIALS AND METHODS.....	61
2.1 Ethics.....	61
2.2 Measurement of replication capacities of viruses encoding Gag-protease isolated from chronically and recently HIV-1 infected subjects.....	61
2.2.1 Study subjects.....	61
2.2.1.1 HIV-1 subtype C chronically infected subjects.....	61
2.2.1.2 HIV-1 subtype B chronically infected subjects.....	62
2.2.1.3 HIV-1 subtype C recently infected subjects.....	65
2.2.2 Generation of Gag-protease NL4-3 recombinant virus stocks.....	68
2.2.2.1 Amplification and sequencing of patient-derived <i>gag-protease</i>	68
2.2.2.2 Preparation of the pNL4-3Δ <i>gag-protease</i> backbone.....	70
2.2.2.3 Thawing, culturing, and storing of CEM-GXR25 cells.....	71
2.2.2.4 Co-transfection of GXR cells with <i>gag-protease</i> and pNL4-3Δ <i>gag-protease</i>	73
2.2.2.5 Monitoring of virus production by flow cytometry.....	75
2.2.2.6 Harvesting virus stocks.....	76
2.2.2.7 Validation of virus stocks.....	76
2.2.3 Titration and replication capacity measurement of virus stocks.....	78

2.2.3.1 Titration	78
2.2.3.2 Replication capacity assay	79
2.2.4 Variability in measuring Gag-protease-mediated replication capacity	80
2.2.5 Data Analysis.....	80
2.3 Measurement of replication capacities of intact viral isolates	86
2.3.1 Isolation of HIV-1 and measurement of replication capacity	86
2.3.2 Data analysis.....	88
2.4 Measurement of replication capacities of mutant viruses	88
2.4.1 Construction of mutant plasmids.....	88
2.4.1.1 The site-directed mutagenesis process.....	88
2.4.1.2 Preparation of plasmid vectors	89
2.4.1.3 Introduction of mutations into plasmid vectors.....	91
2.4.2 Generation of mutant viruses and measurement of replication capacity	93
2.4.3 Data analysis.....	93
CHAPTER 3 - RESULTS	95
3.1 Quality control.....	95
3.1.1 Validation of generated Gag-protease NL4-3 recombinant viruses.....	95
3.1.2 Validation of replication capacity assay	97
3.1.3 Overall method variability: recombinant virus generation and replication capacity assay	97
3.2 HIV-1 subtype C chronic infection.....	100
3.2.1 Distribution of replication capacities	100
3.2.2 Replication capacities versus HLA class I alleles	100
3.2.3 Replication capacities versus viral loads and CD4+ T cell counts	104
3.2.4 Replication capacities versus rates of CD4+ T cell decline	104

3.2.5 Replication capacities versus Gag-protease sequences.....	106
3.2.5.1 Gag-protease polymorphisms	106
3.2.5.2 HLA-associated polymorphisms	106
3.2.5.3 Amino acid variants	110
3.2.5.4 HLA-B*81-associated 186S	112
3.2.5.5 Residues co-varying with 186S	115
3.3 HIV-1 subtype C recent infection	116
3.3.1 Distribution of replication capacities	116
3.3.2 Replication capacities versus markers of disease progression	116
3.3.3 Replication capacities versus protective HLA alleles	119
3.3.4 Replication capacities/viral set points versus Gag-protease sequences	120
3.3.4.1 Amino acid variants	120
3.3.4.2 Number of HLA-associated polymorphisms in Gag.....	124
3.3.4.2.1 Replication capacity	124
3.3.4.2.2 Viral set point	126
3.3.5 Longitudinal analyses of replication capacities, CD8+ T cell responses and clinical parameters	127
3.4 Contribution of Gag-protease to overall replication capacity	128
3.5 HIV-1 subtype B versus subtype C	131
3.5.1 Gag-protease NL4-3 recombinant viruses	131
3.5.2 Subtype-specific Gag mutations.....	133
3.6 HLA-B*81-associated mutations	135
CHAPTER 4 – DISCUSSION.....	141
4.1 Impact of HLA on HIV-1 replication capacity.....	141
4.1.1 Protective HLA alleles: HLA-B*81	141

4.1.2 HLA-B*81 associated mutations: 186S	143
4.2 Impact of HIV-1 replication capacity on disease progression	147
4.2.1 Chronic infection: cross-sectional and longitudinal analyses	147
4.2.2 Recent infection: longitudinal analyses	148
4.3 Identification of mutations with impact on HIV-1 replication capacity	154
4.3 Impact of Gag-protease on overall HIV-1 replication capacity.....	157
4.4 Impact of Gag-protease on HIV-1 replication capacity in different subtypes	158
4.5 Study limitations	160
4.6 Implications for HIV-1 vaccine design	162
4.7 Summary and conclusions	165
APPENDIX.....	168
REFERENCES	176

LIST OF FIGURES

Figure 1.1	Genome organisation, structure, and replication cycle of HIV-1.....	6
Figure 1.2	Pathogenesis of HIV-1 infection.....	28
Figure 1.3	Spectrum and markers of HIV-1 disease progression rate and factors influencing the rate of progression.....	50
Figure 2.1	Generation of Gag-protease NL4-3 recombinant virus stocks.....	77
Figure 2.2	Titration and measurement of replication capacities of Gag-protease NL4-3 recombinant viruses.....	81
Figure 2.3	Isolation of HIV-1 from infected PBMCs and generation of concentrated virus stocks.....	87
Figure 2.4	Site-directed mutagenesis procedure.....	90
Figure 3.1	Neighbour-joining tree (Paup 4.0) of HIV-1 <i>gag</i> sequences.....	96
Figure 3.2	Replication capacity assay quality control analyses.....	98
Figure 3.3	Overall method variability: recombinant virus generation and replication capacity assay.....	99
Figure 3.4	Replication capacities of Gag-protease NL4-3 recombinant viruses encoding Gag-protease from HIV-1 subtype C chronically infected patients.....	101
Figure 3.5	Associations between HLA alleles and replication capacities of recombinant viruses encoding Gag-protease from HIV-1 subtype C chronically infected patients.....	102
Figure 3.6	Relationships between replication capacities of recombinant viruses encoding patient Gag-protease and markers of disease progression in HIV-1 subtype C chronically infected patients.....	105
Figure 3.7	Associations between Gag-protease polymorphisms and replication capacities of recombinant viruses encoding Gag-protease from HIV-1 subtype C	

	chronically infected patients	107
Figure 3.8	Associations between HLA-associated polymorphisms and replication capacities of recombinant viruses encoding Gag-protease from HIV-1 subtype C chronically infected patients	109
Figure 3.9	Association between Gag-protease amino acid variants and altered replication capacities of recombinant viruses encoding Gag-protease from HIV-1 subtype C chronically infected patients.....	111
Figure 3.10	Single HLA-B*81-associated Gag amino acid associations with altered replication capacities of recombinant viruses encoding Gag-protease from HIV-1 subtype C chronically infected patients.....	114
Figure 3.11	Replication capacities of Gag-protease NL4-3 recombinant viruses encoding Gag-protease from HIV-1 subtype C recently infected patients	117
Figure 3.12	Relationships between replication capacities of recombinant viruses encoding patient Gag-protease and markers of disease progression in HIV-1 subtype C recently infected patients.....	118
Figure 3.13	Relationship between protective HLA alleles and replication capacities of recombinant viruses encoding Gag-protease from HIV-1 subtype C recently infected patients.....	121
Figure 3.14	Associations between HLA-associated Gag polymorphisms, replication capacities of recombinant viruses encoding patient Gag-protease, and viral set points in HIV-1 subtype C recently infected patients.....	125
Figure 3.15	Contribution of Gag-protease to overall HIV-1 fitness.....	130
Figure 3.16	Comparison of replication capacities of recombinant viruses encoding subtype B and C Gag-protease from patients chronically infected with HIV-1.....	132
Figure 3.17	Replication capacities of HIV-1 encoding subtype-specific mutations in Gag... ..	134

Figure 3.18	Structure of the HIV-1 Gag p24 capsid.....	136
Figure 3.19	Replication capacities of HIV-1 encoding HLA-B*81-associated mutations in Gag.....	137
Figure 3.20	Replication capacities of 186S mutants generated by the original and modified electroporation protocols.....	139
Figure 4.1	Impact of immune escape with fitness costs on HIV-1 pathogenesis and implications for HIV-1 vaccine design.....	166

LIST OF TABLES

Table 1.1	Role of Gag proteins in the HIV-1 replication cycle.....	15
Table 2.1	Clinical and demographic characteristics of HIV-1 subtype C chronically infected patients.....	63
Table 2.2	Clinical characteristics of matched HIV-1 subtype B and C chronically infected patients.....	64
Table 2.3	Clinical and demographic characteristics of HIV-1 subtype C recently infected patients.....	67
Table 2.4	Forward primers used to introduce specific Gag mutations into SK-254 HIV-1 Gag-protease and pNL4-3	92
Table 3.1	Amino acids in HIV-1 subtype C Gag-protease from chronically infected subjects associated with altered replication capacities by multivariate analysis ($p < 0.001$, $q < 0.25$).....	113
Table 3.2	Amino acids in HIV-1 subtype C Gag-protease from recently infected patients associated with alterations in replication capacity ($p < 0.05$).....	122
Table 3.3	Longitudinal measurements of Gag-protease-mediated viral replication capacities, Gag CD8+ T cell responses, and clinical measures from patients recently infected with HIV-1 subtype C (including patient HLA profiles)..	129
Table A1	HLA-associated amino acids in HIV-1 subtype C Gag-protease	168
Table A2	Best-defined (A-list) optimal Gag epitopes from the Los Alamos HIV molecular immunology database.....	171
Table A3	Amino acids in Gag-protease from HIV-1 subtype C chronically infected subjects associated with altered replication capacities.....	173

ABBREVIATIONS

μg, microgram	CCR5, C-C chemokine receptor 5
μl, microlitre	CD, cluster of differentiation
μM, micromolar	CO ₂ , carbon dioxide
°C, degrees Celsius	CRF, circulating recombinant form
A, deoxyadenosine nucleotide	CsA, cyclosporin A
aa, amino acid	CXCR4, CXC chemokine receptor 4
Ab, antibody	CypA, cyclophilin A
ADCC, antibody-dependent cell-mediated cytotoxicity	D', linkage disequilibrium
Ag, antigen	DC-SIGN, dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin
AIDS, acquired immunodeficiency syndrome	DEPC, diethyl pyrocarbonate
Alix, apoptosis-linked gene 2-interacting protein X	DMSO, dimethylsulfoxide
ANOVA, analysis of variance	DNA, deoxyribonucleic acid
AP, clathrin adaptor protein	dNTP, deoxyribonucleotide triphosphate
APC, antigen presenting cell	E, enzyme-linked immunosorbent assay (ELISA)
APOBEC3G, apolipoprotein B messenger RNA editing catalytic subunit-like protein 3G	EC, elite controller
ART, antiretroviral therapy	EDTA, ethylenediaminetetraacetic acid
bp, base pair	ELISA, enzyme-linked immunosorbent assay
C, carboxy	ELISPOT, enzyme-linked immunospot assay
Cap, capsid	Env, envelope glycoprotein

ESCRT, endosomal sorting complexes required for transport	KIR, killer cell immunoglobulin-like receptor
FBS, foetal bovine serum	LB, Luria-Bertani
Fc, fragment crystallisable	L-domain, late domain
FI, entry/fusion inhibitor	LEDGF, lens epithelium-derived growth factor
Gag, group specific antigen	LTNP, long-term non-progressor
GFP, green fluorescent protein	LTR, long terminal repeat
gp(X), glycoprotein of X kDa	lys, lysine
GXR, CEM-GXR25	M, matrix
H ₂ SO ₄ , sulphuric acid	MHC, Major Histocompatibility Complex
hd, heterodimer	MIP, macrophage inflammatory protein
HEPES, N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid	ml, millilitre
HIV, human immunodeficiency virus	mm, millimetre
HIV-1, human immunodeficiency virus type 1	mM, millimolar
HLA, human leukocyte antigen	MOI, multiplicity of infection
HRP, horse-radish peroxidase	N, amino
HTA, heteroduplex tracking assay	Nef, negative regulation factor
IFN, interferon	ng, nanogram
IgA, immunoglobulin A	NK, natural killer
IL, interleukin	NLS, nuclear localisation signal
IN, integrase	NNRTI, non-nucleoside reverse transcriptase inhibitor
IQR, interquartile range	No., number
kb, kilobase	NRTI, nucleoside reverse transcriptase inhibitor
kDa, kilodalton	

Nu, nucleocapsid	SNP, single-nucleotide polymorphism
p(X), protein of X kDa	ss, single-stranded
PBMC, peripheral blood mononuclear cell	T, deoxythymidine nucleotide
PBS, primer binding site	Tat, transactivator of transcription factor
PCR, polymerase chain reaction	TCR, T cell receptor
PFA, paraformaldehyde	TIP47, tail-interacting protein of 47kDa
PHA, phytohaemagglutinin	TLR, toll-like receptor
PI, protease inhibitor	TMB, tetramethylbenzidine
pNL4-3, NL4-3 plasmid	Topo E, topoisomerase enzyme
pNL4-3 Δ <i>gag-protease</i> , <i>gag-protease-</i> deleted NL4-3 plasmid	TRIM5 α , tripartite motif-containing 5 α
Pol, polymerase	tRNA, transfer ribonucleic acid
Pr(X), precursor of X kDa	Tsg101, host tumour susceptibility gene 101
PR, protease	U, unit
Ref1, restriction factor 1	V, volt
Rev, regulator of virion protein	Vif, viral infectivity factor
RN, RNase H	Vpr, viral protein R
RNA, ribonucleic acid	Vpu, viral protein U
rpm, revolutions per minute	WB, Western blot
RT, reverse transcriptase	x g, times gravity
RT-PCR, reverse transcriptase polymerase chain reaction	Zn, zinc
SD, standard deviation	α , alpha
SFC, spot forming cell	β , beta
SIV, simian immunodeficiency virus	γ , gamma
SK, Sinikithemba	μ F, microfarad

CHAPTER 1

INTRODUCTION

CHAPTER 1 - INTRODUCTION

1.1 History and epidemiology of human immunodeficiency virus (HIV) type 1 infection

HIV is a retrovirus belonging to the *Retroviridae* family and *Lentivirus* genus [1]. It infects human immune cells that express the cluster of differentiation (CD)4 protein, causing progressive immune dysfunction and deficiency that results in a syndrome of opportunistic infections and cancers, known as acquired immunodeficiency syndrome (AIDS) [1]. HIV infection is a modern disease that was first recognised in 1981 [2, 3]. Homosexual men, blood transfusion recipients, intravenous drug users, as well as children and sexual partners of these individuals, were first described as affected by this new immunodeficiency syndrome, suggesting that the etiologic agent was transmitted in body fluids [2]. By 1984, HIV was established as the causative agent [1, 2, 4-7], and was later renamed HIV type 1 (HIV-1) to distinguish it from a related, but considerably less prevalent, AIDS-causing retrovirus, HIV-2, isolated in 1986 [2].

HIV-1 is closely related to simian immunodeficiency viruses (SIVs) that infect chimpanzees in west-central Africa [3]. The origin of the three major strains of HIV-1, namely groups M, N and O, can be traced to three independent transmissions of SIV from chimpanzees (and possibly gorillas in the case of group O) to humans in the region of the Cameroon where these apes were butchered for bush meat [3, 8]. Recently (in 2009), another HIV-1 strain, thought to be transmitted from gorillas to humans, was identified with the proposed designation of group P [9]. Using some early HIV-1 group M sequences (1959 and 1960), which already indicated significant diversification of the common ancestor, and statistical models estimating the rate of evolution, the date of origin of the HIV-1 group M strain was estimated to be 1908 [3, 8]. Since then, especially with the growth of cities in Africa and increasing ease of global travel from about the mid-twentieth century onward [3], the

epidemic has grown exponentially, resulting in 33 million people living with HIV at the end of 2009, with 1.8 million deaths in 2009 alone [10], and approximately 2.6 million new infections diagnosed per year [11].

The HIV-1 group M strain accounts for greater than 95% of HIV infections [3, 8]. It is considerably diverse, consisting of 9 different subtypes (A-D, F-H, J-K) and at least 48 circulating recombinant forms (CRFs) [8]. Subtypes and CRFs are unevenly distributed globally (probably due to founder effects), with the greatest diversity found in west-central Africa. Subtype B predominates in Europe, the Americas and Australia; subtypes B, C and CRF-01 (a recombinant of subtypes A and E) are found in Asia; while subtype C predominates in sub-Saharan Africa [1, 2, 8]. Since approximately 68% of the global burden of HIV-1 disease is carried by sub-Saharan Africa [10], HIV-1 subtype C, which is the focus of the present study, is the most prevalent subtype worldwide [8].

A discussion on the structure and function of HIV-1 and its interaction with the human host to cause disease follows.

1.2 Structure of HIV-1

1.2.1 Genome organisation

The 9719 base pair (bp) HIV-1 genome (reference strain HXB2) is made up of nine overlapping genes flanked by two identical 634 bp long terminal repeat (LTR) regions: group specific antigen (*gag*, 1,503 bp in HXB2 reference strain), polymerase (*pol*, 3,012 bp), envelope glycoprotein (*env*, 2,571 bp), transactivator of transcription factor (*tat*, 306 bp),

regulator of virion protein (*rev*, 351 bp), viral infectivity factor (*vif*, 579 bp), viral protein R (*vpr*, 292 bp), viral protein U (*vpu*, 249 bp), and negative regulation factor (*nef*, 621 bp) [12]. Tat and Rev, also known as the regulatory proteins, control viral gene expression and nuclear export, respectively, while the accessory proteins Vif, Vpr, Vpu, and Nef antagonise host defences and enhance pathogenicity [13]. The three major proteins shared by all *Retroviridae*, Gag, Pol, and Env, are synthesised as polyproteins, namely precursor of 55 kDa (Pr55^{gag}; Gag polyprotein), Pr160^{gagpol} (Gag-Pol polyprotein minus Gag p6), and glycoprotein of 160 kDa (gp160; Env polyprotein) [13]. Viral protease cleaves the Gag polyprotein into protein of 17 kDa (p17) matrix (132 amino acids [aa]), p24 capsid (231 aa), p7 nucleocapsid (55 aa), and p6 (52 aa), as well as the spacer peptides, p2 (14 aa) and p1 (16 aa) [12, 13]. Protease also cleaves the enzymatic Pol proteins from the Gag-Pol precursor, namely protease (99 aa), reverse transcriptase (560 aa), and integrase (288 aa) [12, 13]. The Env polyprotein is cleaved by a host cellular enzyme into surface envelope gp120 (511 aa) and transmembrane envelope gp41 (345 aa) [12, 13].

1.2.2 HIV-1 structure

The outermost layer of the HIV particle is the envelope, which consists of the viral envelope glycoproteins inserted into lipid membrane derived from the host cell plasma membrane [2]. The outer shell immediately underneath and bound to the lipid bilayer of the envelope is composed of up to 5,000 Gag p17 matrix molecules [14, 15] that are assembled as hexamers of trimers [16]. The matrix molecule is folded into five alpha (α)-helices and a three-stranded beta (β)-sheet [14]. Immediately beneath the matrix is the conical structure formed by approximately 1,500 Gag p24 capsid molecules arranged in hexamers [15]. The capsid units are composed of an amino (N)-terminal domain (seven α -helices, two β -hairpins, and a cyclophilin A [CypA] binding loop) [17, 18] connected by a short linker to a carboxy (C)-

terminal domain that is mostly helical [19]. Enclosed within the capsid structure are approximately 1,500 Gag p7 nucleocapsid molecules that are in tight association with the two copies of plus single-stranded viral genomic ribonucleic acid (RNA) through their two zinc (Zn) finger domains, which are themselves connected by a highly basic linker domain [20, 21]. Incorporated into the nucleocapsid are also about 50 copies of reverse transcriptase, protease, and integrase in dimeric forms as well as approximately 100-200 molecules of Vpr [21].

The genome organisation and the major components of the virus structure are shown in Figure 1.1.

1.3 HIV-1 replication cycle: role of Gag

The replication cycle of HIV-1 may be categorised into several steps (depicted in Figure 1.1): (i) virus entry, (ii) reverse transcription of genomic RNA to double stranded deoxyribonucleic acid (DNA), (iii) uncoating of the capsid particle to release the viral pre-integration complex, (iv) nuclear import of the pre-integration complex, (v) integration of viral DNA into the host cell genome, (vi) transcription and nuclear export of RNA, (vii) viral protein synthesis and virus assembly, (viii) budding and cell exit, and (ix) maturation. Each of these events in the replication cycle will be briefly outlined below with a focus on the important role of Gag, the major viral structural protein that is of particular interest in the present study (summarised in Table 1.1, presented at the end of Section 1.3, page 15). Note that Gag residues referred to in the text are numbered sequentially from the start codon (residue 1) according to the HXB2 reference strain.

Figure 1.1 Genome organisation, structure, and replication cycle of HIV-1

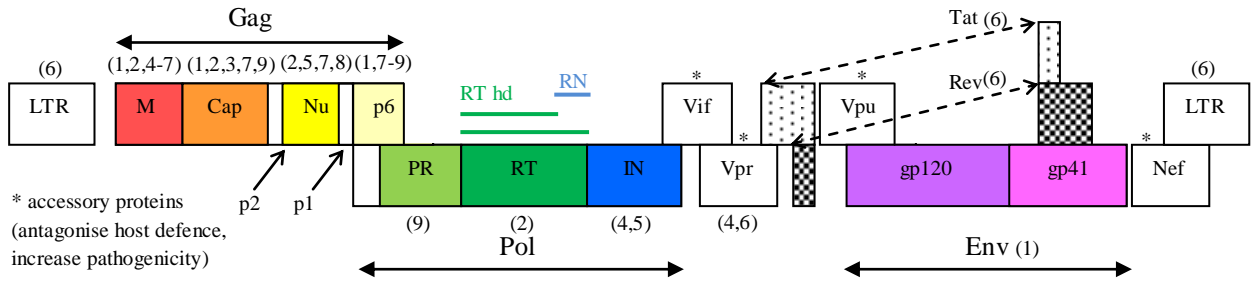
(i) The organisation of the 9 overlapping genes, encoding HIV-1 proteins, flanked by the LTRs is shown, with the regions coding for the major proteins (Gag, Pol, and Env) highlighted in colour (adapted from <http://www.hiv.lanl.gov/content/sequence/> [267]).

(ii) The main elements of the mature virion are depicted with colours corresponding to coding regions of the HIV-1 genome as shown in (i) (modified from Freed (2001) [13]).

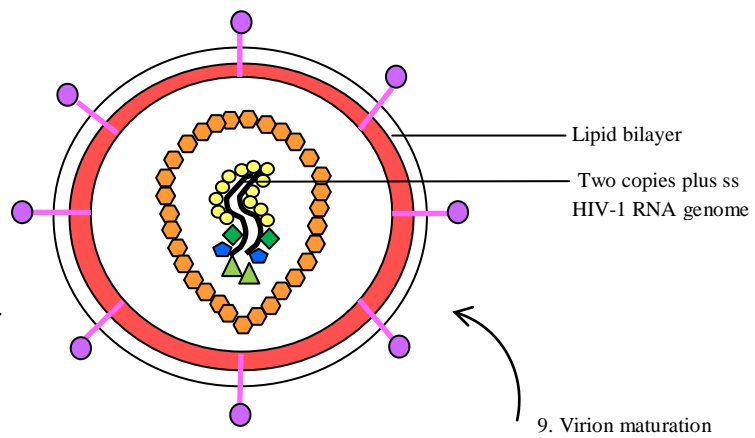
(iii) The various steps of the HIV-1 replication cycle are illustrated from cell entry of the virus to virion maturation (modified from Freed (1998) [18]). The direct or indirect role of the HIV-1 proteins in mediating various steps of the replication cycle is indicated by numbering in (i).

HIV-1 – human immunodeficiency virus type 1; LTR – long terminal repeat region; Gag – group specific antigen; M – matrix; Cap – capsid; p2 – protein of 2 kDa; Nu – nucleocapsid; p1 – protein of 1 kDa; p6 – protein of 6 kDa; Pol – polymerase; PR – protease; RT – reverse transcriptase; hd – heterodimer; RN – RNase H; IN – integrase; Vif – viral infectivity factor; Vpr – viral protein R; Tat – transactivator of transcription factor; Rev – regulator of virion protein; Vpu – viral protein U; Env – envelope glycoprotein; gp120 – glycoprotein of 120 kDa; gp41 – glycoprotein of 41 kDa; Nef – negative regulation factor; ss – single-stranded; RNA – ribonucleic acid.

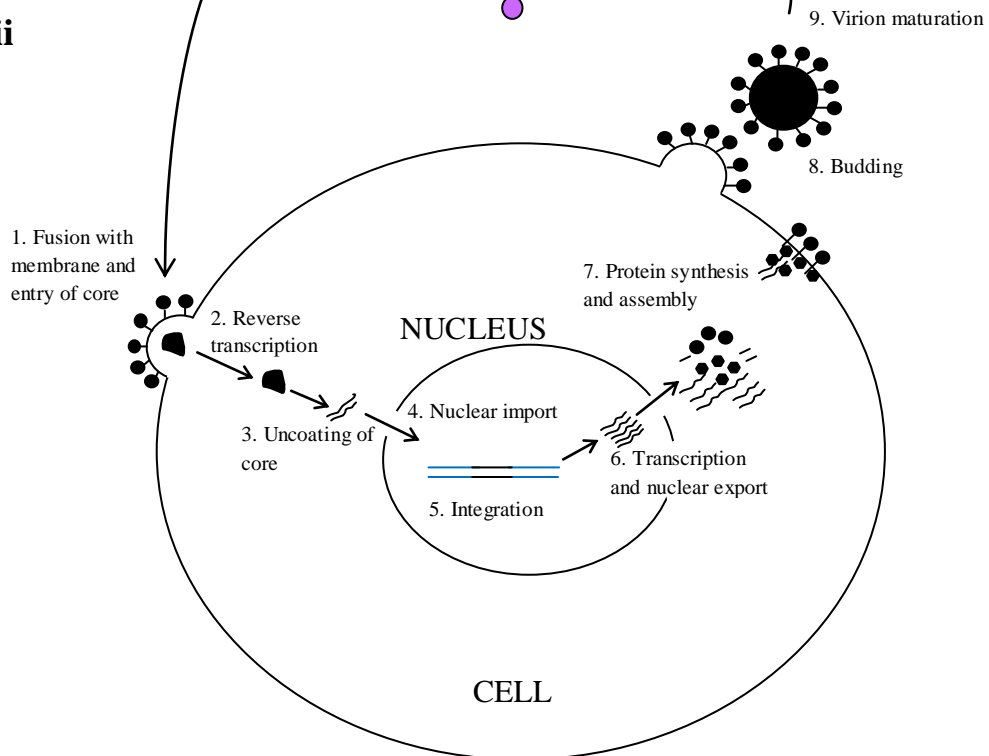
i



ii



iii



1.3.1 Virus entry

Gp120 binds to the CD4 receptor protein, which is present on the surface of T helper cells, T regulatory cells, monocytes, macrophages, dendritic cells, and Langerhans cells [1]. This causes a conformational change in gp120, allowing a secondary co-receptor binding site to engage either the C-C chemokine receptor 5 (CCR5) protein (found on T cells, monocytes, macrophages, dendritic cells, and Langerhans cells) or the CXC chemokine receptor 4 (CXCR4) protein (found on T cells) [13, 22]. HIV-1 may use other co-receptor proteins to a much lesser extent [1]. Co-receptor binding triggers fusion pore formation and a conformational change in gp41 that results in the formation of a helical bundle structure, which completes fusion of the viral and cellular membranes [22, 23]. Viral fusion is dependent on virion maturation, and Gag interaction with Env gp41 is thought to couple fusion to maturation [24, 25].

1.3.2 Reverse transcription

Following virus entry, the virus shell alternately binds to actin filaments and a microtubule network that mediates its transit to the nuclear membrane [26]. Reverse transcription is thought to predominantly take place at the nuclear pore and occurs within an intact capsid shell prior to uncoating [26].

Reverse transcription is initiated by the binding of host cell transfer (t)RNA^{lys} to the 5' end primer binding site (PBS) of the plus single-stranded HIV-1 RNA [2], which is chaperoned by the Gag nucleocapsid [18]. The packaging of tRNA^{lys} into virions is mediated through the binding of Gag (at Gag residues 308-362) to lysyl-tRNA synthetase which then interacts with tRNA^{lys} [27]. Following the synthesis of a segment of minus strand DNA from the PBS

to the 5' end by viral reverse transcriptase, viral RNase H (cleaved from reverse transcriptase) degrades the RNA portion of the DNA/RNA hybrid [2]. The synthesised DNA segment then binds to the 3' end of the RNA and acts as a primer for minus strand DNA synthesis (known as the first strand transfer). RNase H simultaneously degrades RNA, leaving two RNA segments which act as primers for complementary plus strand DNA synthesis. Following displacement of tRNA^{lys} from the PBS by RNase H, the complementary PBS on both strands can hybridise and both strands of DNA are synthesised to completion (known as the second strand transfer). These strand transfer reactions are chaperoned by the Gag nucleocapsid and mutations at Gag residues 400 and 421 in the Zn fingers (residues 392-405 and 413-426) of the nucleocapsid cause defects in reverse transcription [28].

These nucleocapsid mutations [28], as well as mutations at Gag residues 384, 387, 388, and 391 in the nucleocapsid [29], also cause decreased stability of the newly synthesised DNA indicating a role for Gag nucleocapsid in protection of newly synthesised DNA. The Gag matrix protein may also be involved in stabilising newly synthesised DNA or in facilitating reverse transcription as a mutation at residue 21 reduced viral DNA synthesis [30]. The mechanism underlying the effects of this mutation may be due to increased membrane binding of the matrix protein. In addition, Vif protects the newly synthesised DNA from hypermutation by the host restriction factor apolipoprotein B messenger RNA editing catalytic subunit-like protein 3G (APOBEC3G) though targeting it for degradation [2].

The final step of reverse transcription is the formation of a DNA flap upon the termination of plus strand synthesis [26]. This is thought to trigger uncoating of the capsid [26], which still remains an ill-defined process [31].

1.3.3 Uncoating

The regulation of uncoating is important for the success of reverse transcription as evidenced by the impaired reverse transcription in virions with unstable capsid cores [32]. The Gag capsid and its interaction with host factors, such as CypA, tripartite motif-containing 5 α protein (TRIM5 α), and possibly other restriction factors, regulates uncoating [31]. The Gag capsid binds CypA through the CypA binding loop region (Gag residues 217-225 [33]) which results in conformational changes to the C-terminal domain of the capsid and may therefore play a role in destabilisation of the capsid during uncoating [17, 27]. CypA is also thought to protect the capsid from host restriction factors that promote premature uncoating [31, 34]. Rhesus and human TRIM5 α were shown to restrict HIV-1 [35] through accelerated uncoating [36] and human CypA was suggested to block restriction factor 1 (Rf1; a human TRIM5 α variant) activity in human cells [37]. However, human TRIM5 α has only a modest effect on HIV-1 and this was later found to be independent of capsid-CypA interaction [38]. Therefore, CypA through interaction with the Gag capsid may protect HIV-1 from an unknown host restriction factor [31, 34, 39].

1.3.4 Nuclear import

Uncoating of the capsid is required prior to nuclear entry of newly synthesised DNA [26]. The pre-integration complex is composed of Gag matrix, Gag nucleocapsid, reverse transcriptase, integrase, Vpr, double-stranded DNA, and several host proteins, including lens epithelium-derived growth factor (LEDGF) [14, 26, 40]. Vpr, integrase, the DNA flap, and the matrix protein have nuclear localisation signals (NLSs) [14], which may be recognised by importins in the nuclear pore complex, resulting in transport through the complex [40]. However, whether NLS of viral components are required for nuclear import remains controversial [40]. The NLS in matrix are at residues 25-32 and 110-114 [14, 41].

However, mutation of residues 26 and 27 did not change nuclear transport but impaired infectivity and integration, suggesting a role for matrix in integration [42].

1.3.5 Integration

The HIV-1 integrase protein catalyses the integration of viral DNA with host cell DNA. The 3' ends of viral DNA are cleaved and the cellular target DNA is cleaved in a staggered manner, followed by the joining of these [2]. The nucleocapsid is involved in 3' end processing since mutation of Gag residues 400 and 421 results in defective processing and impaired integration [28]. Cellular repair enzymes then fill in the gaps between integrated viral DNA and host DNA [2]. Intact pre-integration complexes including the host cellular factors that are part of the complex are required for efficient integration [40].

1.3.6 Transcription and nuclear export

HIV-1 Tat recruits a complex of host cellular factors to the integrated viral LTR, resulting in the phosphorylation of RNA polymerase II and stimulation of transcription [40]. Vpr may also activate transcription and regulate viral gene expression [43]. Singly and multiply spliced RNAs encoding viral proteins are transcribed, and unspliced RNA which constitutes the HIV-1 genome is also produced [40]. Rev interacts with the cellular machinery to export these viral RNAs into the cytoplasm where they are translated into viral proteins and polyprotein precursors [2]. The Gag matrix protein possesses a nuclear export signal at the N-terminus, which when disrupted by mutation at residues 18 and 22 results in the accumulation of viral Gag RNA and genomic RNA in the nucleus [44].

1.3.7 Assembly

The Env glycoproteins are synthesised on ribosomes attached to the endoplasmic reticulum membrane, while Gag and Gag-Pol precursors are synthesised on free ribosomes in the cytoplasm [21]. The trafficking of Env glycoproteins to the plasma membrane occurs via the secretory pathway during which time the cellular enzyme, furin, cleaves Env into the gp41 and gp120 subunits [21]. The gp120/gp41 complex segregates to lipid rafts where virion assembly is targeted [45], and are actively recruited into virions via interaction of Gag matrix with the gp41 cytoplasmic tail [46]. The N-terminal 100 amino acids of the matrix protein are required for the incorporation of envelope glycoproteins into virion particles [46], including specific interacting residues at Gag codons 8, 9, 13, 16, 17, 18, 31, 35, and 63 [47-50]. More recently, this interaction was found to be mediated by host tail-interacting protein of 47 kDa (TIP47), and the binding of this factor by matrix residues 6-17 was shown to be essential for envelope glycoprotein incorporation into virions [50].

The nucleocapsid domain of the Gag polyprotein binds to full-length viral genomic RNA and this complex trafficks to the lipid rafts of the plasma membrane via an endosomal pathway [51, 52]. Gag-RNA complexes may also be trafficked to sites of cell-cell contact by the endosomal pathway, thereby facilitating cell-to-cell transmission [52]. Gag trafficking is largely dependent on the matrix protein [14, 18, 53]. The first N-terminal α -helical matrix domain (residues 11-19 [54]) and Gag residues 132 and 135 interact with host proteins of the endosomal pathway clathrin adaptor protein (AP)-3 and AP-2, respectively [55, 56]. Domains in the matrix protein may also interact with other cellular factors of the endosomal pathway [21]. Residues 85-89 of the matrix protein are involved in directing assembly to the plasma membrane, rather than to intracellular membranes [57]. The matrix protein is also responsible for binding to and therefore targeting assembly to the plasma

membrane. The highly basic domain at the N-terminus of matrix (residues 15-31 [58]) is positively charged and specifically interacts electrostatically with a negatively charged acidic phospholipid found in the plasma membrane [59]. This binding, as well as Gag nucleocapsid binding to viral genomic RNA and Gag multimerisation, increases the exposure of the previously sequestered myristate moiety bound to the N-terminus of matrix (the first 7 amino acids are required for myristate recognition [57]) to promote anchoring of myristate into the inner leaflet of the plasma membrane [59]. A salt bridge formed between residues 12 and 89 is important for this myristate exposure and pH, which may vary in subcellular locations, may regulate this [60]. A hydrophobic domain (residues 7-9 [61]) synergising with the basic domain and myristate also plays a role in membrane binding. Importantly, membrane binding can also be regulated by phosphorylation: upon viral entry phosphorylation of serine residues 9, 67, 72, and 77 may promote membrane dissociation, facilitating early steps of the life cycle [62, 63].

The binding of the Gag nucleocapsid to RNA occurs through the hydrophobic plateau [53], which is located between residues 390 and 423 and encompasses Zn finger domains and basic residues that flank the first Zn finger [21]. This binding is important for the initiation of Gag multimerisation, through the localising and concentrating of Gag monomers as well as providing an RNA scaffold for assembly [2, 53]. Some multimerisation may initiate before the complex reaches the plasma membrane, but extensive multimerisation occurs at the plasma membrane with the accumulation of Gag-RNA complexes in the lipid raft domains [59]. Multimerisation is mediated largely by interactions between the capsid domains to form hexamers (that are joined mainly through the capsid C-terminal domains) [15]. Approximately 5,000 Gag subunits interact to form a spherical capsid that encloses 2 copies of the viral RNA genome as well as other viral and cellular proteins [15].

The role of the Gag p6 protein during virus assembly is the incorporation of Vpr via residues 463-466, 482-484, and 489-493 [64]. It also plays an important role in the budding of virus particles and their release from the cell [65].

1.3.8 Budding

Budding of the new virion particle is thought to be mediated primarily by the interaction of Gag p6 late (L)-domain PTAP (residues 455-458) with the host tumour susceptibility gene 101 (Tsg101) protein [66]. Another L-domain which binds the host apoptosis-linked gene 2-interacting protein X (Alix) was identified in Gag p6, namely **LYPLASLRSL** (residues 483-492; essential residues in bold) [64, 67]. Tsg101 and Alix recruit the endosomal sorting complexes required for transport (ESCRT) to direct viral budding [65]. Mutations in the Alix binding region of p6 impact negatively on virus replication, and more severely so when Tsg101 (part of the ESCRT-I complex) binding is simultaneously inactivated [68]. Recently it was also shown that the basic residues of the Gag nucleocapsid (380, 384, 387, 388, 391, 397, 403, 406, 409-411, 415, 418, 424, 429) are required for budding, probably through allowing interaction with host ESCRT machinery downstream of Tsg101 [65].

1.3.9 Maturation

Upon budding, the viral protease enzyme cleaves the Gag and Gag-Pol precursors, resulting in virion maturation. The Gag matrix remains associated with the envelope and 1,000-1,500 subunits of the Gag capsid re-assemble into a cone-shaped structure [15]. In the mature capsid, interfaces between N-terminal domains and between N-terminal domains and C-terminal domains form the hexamers; while interfaces between C-terminal domains connect the hexamers [15]. In general, mutations in the capsid C-terminal domain (residues 283-

363) impair virus assembly, but mutations in the capsid N-terminal domain (residues 133-277) impair maturation [69]. Virus maturation is essential for virus infectivity [70] (Section 1.3.1). Recently, mutation of a conserved Gag residue in p6, 488, was demonstrated to reduce Gag cleavage between the capsid and spacer peptide resulting in irregular core structure and consequently reduced viral infectivity [64]. Refer to Table 1.1 for a summary of the role of Gag in the HIV-1 replication cycle.

Virion maturation completes the viral replication cycle, which requires approximately 2 days in total [1, 71]. In untreated chronic infection, the replication cycle results in the production of approximately 10^{10} to 10^{11} virions per day, each with a cell-free half-life of 30 to 60 minutes [1]. Productively infected CD4+ T cells have a half-life of 1 day and are therefore rapidly eliminated [1, 72]. However, the pathogenesis of HIV-1 is complex and the disease is likely the result of both direct cytopathic and indirect effects [73].

1.4 HIV-1 Pathogenesis

1.4.1 Acute infection

HIV-1 is most commonly transmitted across a mucosal surface, although other important modes of transmission include parenteral (*e.g.* intravenous drug use, blood transfusion) and mother-to-child transmission (intrapartum, perinatally or via breast milk) [1, 74]. HIV-1 may gain access to the genital or rectal submucosa through minor breaches of the epithelium induced by sexual intercourse [75]. Alternative possibilities include transcytosis through the epithelium, contact with intraepithelial dendritic cells, or movement through intercellular spaces in the epithelium [72]. Studies conducted on rhesus macaques infected with SIV and

Table 1.1 Role of Gag proteins in the HIV-1 replication cycle

Gag protein	Residues	Role	Replication step
Matrix (p17): residues 1-132	9, 67, 72, 77	Serine residues phosphorylated on viral entry, which may promote membrane dissociation.	Early post-entry steps
	21	Mutation increased membrane binding, reducing viral DNA synthesis.	Early post-entry steps
	25-32, 110-114	NLSs which may be involved in nuclear import.	Nuclear import
	26, 27	Impaired integration when mutated.	Integration
	18, 22	Impaired nuclear export when mutated.	Nuclear export
	1-100: 6-17, 18, 31, 35, 63	Envelope glycoprotein incorporation. Residues 6-17 bind TIP47, which mediates the incorporation. This interaction couples Gag maturation to fusion.	Assembly/Entry
	11-19, 132	Trafficking of Gag proteins.	Assembly
	85-89	Targeting virus assembly to plasma membrane.	Assembly
	15-31	Basic domain mediating membrane binding.	Assembly
	7-9	Hydrophobic domain synergising with basic domain and myristate to promote membrane binding.	Assembly
1-7	Binding of myristate required for membrane binding.	Assembly	
12, 89	Regulation of myristate exposure to control membrane binding.	Assembly	
Capsid (p24): residues 133-363	308-362	Binds to lysyl-tRNA synthetase that mediates packaging of tRNA ^{lys} which is required for initiation of reverse transcription.	Reverse transcription
	217-225	Binds cyclophilin A which protects HIV-1 from an unknown restriction factor and facilitates proper uncoating of the capsid.	Uncoating
	135	Trafficking of Gag proteins.	Assembly
	283-363	Mutations in this domain mainly impair virus assembly.	Assembly
	133-277	Mutations in this domain mainly impair virus maturation.	Maturation/Entry
Nucleocapsid (p7): residues 378-432	400, 421	Impaired reverse transcription when mutated. Zn fingers chaperone strand transfers in reverse transcription.	Reverse transcription
	384, 387, 388, 391, 400, 421	Involved in the protection of newly synthesised DNA.	Steps following reverse transcription
	400, 421	Mutations impaired 3' end processing of viral DNA.	Integration
	390-423	Hydrophobic plateau binding RNA genome to nucleocapsid. This regulates myristate exposure promoting membrane binding and provides a scaffold for virus assembly.	Assembly
	380, 384, 387, 388, 391, 397, 403, 406, 409-411, 415, 418, 424, 429	Basic residues required for budding possibly through interaction with ESCRT machinery.	Budding
p6: residues 449-500	463-466, 482-484, 489-493	Mediates incorporation of Vpr into virions.	Assembly
	455-458 (PTAP)	Interacts with Tsg101 protein (part of ESCRT-I) to recruit ESCRT machinery for budding.	Budding
	483-492: 483-485, 489, 492	Interacts with Alix protein to recruit ESCRT machinery for budding.	Budding
	488	Mutation disrupts Gag cleavage by protease.	Maturation /Entry

HIV-1 – human immunodeficiency virus type 1; NLS - nuclear localisation signal; TIP47 – tail-interacting protein of 47 kDa; tRNA – transfer ribonucleic acid; Zn - zinc; DNA – deoxyribonucleic acid; ESCRT - endosomal sorting complexes required for transport; p6 – protein of 6 kDa; Vpr – viral protein R; Tsg101 protein – tumour susceptibility 101 protein; Alix – apoptosis-linked gene 2-interacting protein X.

sequencing of HIV-1 transmitted viruses indicate that the first cells infected are sub-optimally activated CD4⁺ CCR5⁺ memory T cells [72, 75, 76] and that the transmitted virus is dependent on CCR5 (not CXCR4) for cell entry [77]. Other cells found in the mucosa that could be infected in the acute phase include Langerhans cells (mucosal dendritic cells) and macrophages [73, 78, 79]. The initial infection is supported by signalling from the mucosal epithelial cells (triggered by pathogen-mediated activation of toll-like receptors [TLRs]) that recruits dendritic cells, which in turn secrete cytokines that attract activated CD4⁺ T cells susceptible to infection [75, 76]. Thus, the innate immune response promotes establishment of infection (Section 1.4.4.1).

After approximately 10 days (the eclipse phase [72, 80, 81]), the expansion of infection is sufficient to result in dissemination of cell-free virus or infected cells to the draining lymph node and consequently the bloodstream and secondary lymphoid organs [72, 75]. The dendritic cells in the mucosa may also internalise HIV-1 through dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin (DC-SIGN) receptors and retain the infectious particles, migrate to the draining lymph node, and make contact with CD4⁺ T cells in the lymph node, resulting in efficient transfer of HIV-1 that assists in dissemination of the virus [1, 73, 79]. Furthermore, trapping of HIV-1 virions by dendritic cells in the lymph node exposes susceptible CD4⁺ T cells to infection as they migrate into the lymph node follicle to provide help to B cells [1]. Since many susceptible cells are found in close proximity within the lymphoid tissues, a massive burst of HIV-1 replication occurs in these tissues. This is particularly true in the gut-associated lymphoid tissue which harbours a high concentration of activated CD4⁺ CCR5⁺ memory T cells [72]. Interestingly, it was recently noted that successfully transmitted viruses have a high affinity for $\alpha 4\beta 7^+$ CD4⁺ T cells, which are present in the vaginal/rectal mucosa and home to the gut (where the exponential

replication is best facilitated), suggesting this characteristic may be crucial to the successful establishment of infection [11]. A massive depletion of CD4⁺ T cells occurs during this early exponential phase of viral replication. During this replication burst, at around 21-28 days post-infection, peak plasma viremia (>1 million HIV-1 RNA copies/ml) is reached, and more than 50% of individuals may experience acute viral symptoms, such as fever, rash, and/or lymphadenopathy [72, 82].

Prior to this replication burst, HIV-1 RNA is first detectable in the blood (Fiebig stage I, 10-17 days post-infection) followed by p24 antigen (Fiebig stage II, 17-22 days post-infection) [80, 81]. The first HIV-specific antibodies become detectable (seroconversion) during the phase of peak viremia (Fiebig stage III, 22-25) [80, 81], but they are non-neutralising (*i.e.* they are unable to block virus entry into cells) (Section 1.4.4.2). Viremia then declines during the following Fiebig stages (stage IV [25-31 days post-infection], stage V [31-101 days post-infection], and stage VI [open-ended, early chronic infection phase]), which are characterised by the detection of different antibodies by Western blot, and a steady state (viral set point) is established during Fiebig stages V or VI [80, 81]. The CD8⁺ T cell responses emerge as plasma viremia approaches peak and peaking of these responses is coincident with decline in viremia, suggesting that the suppression of viral replication is largely mediated by CD8⁺ T cells [72] (Section 1.4.4.3.1).

During the acute phase of infection a reservoir of latently infected cells is established [72, 74]. These cells are primarily resting memory CD4⁺ T cells [83], but macrophages are also important reservoirs [73, 78]. These latently infected cells are invisible to the immune response, safe from antiretroviral drugs, long-lived (some persist for decades), and produce infectious virus only when re-activated; thereby preventing elimination of the infection [78,

83] (Section 1.5.1). Productively infected macrophages are also an important reservoir as they can survive for several weeks and efficiently transmit HIV-1 to CD4⁺ T cells via cell-to-cell synapses [84].

1.4.2 Chronic infection

The viral set point is maintained by a balance between ongoing virus replication and host immune responses as the host enters a chronic phase of infection [1, 72]. During this time there is a gradual decline in circulating CD4⁺ T cells and loss of immune function. The decline in immune function may be due to direct manipulation of immune cells by HIV-1 proteins as well as the depletion of CD4⁺ T cells, since these cells provide help to the humoral and cell-mediated arms of the immune system (Section 1.4.4.3.2). The depletion of CD4⁺ T cells is not only caused by direct infection but also as a result of chronic immune activation and inflammation.

Chronic immune activation plays a central role in HIV-1 pathogenesis, and this is highlighted by the lack of immune activation and symptomatic disease in SIV natural hosts despite uncontrolled viremia [85-87]. This state of chronic immune activation may be due to several factors: immune responses to ongoing HIV-1 replication, direct activation of T cells and macrophages by HIV-1 Nef and gp120, microbial translocation across the gut mucosa (due to depletion of CD4⁺ T cells in the gut) resulting in production of pro-inflammatory cytokines by innate immune cells, reactivation of viral infections such as cytomegalovirus and Epstein Barr virus, and molecular mimicry of human leukocyte antigen (HLA) molecule domains resulting in allogeneic non-specific T cell activation [85-87]. The pro-inflammatory cytokines released during chronic immune activation lead to activation of

latently infected cells resulting in virus replication, and activation of uninfected CD4⁺ T cells, which leads to rapid apoptosis in most cases and is perhaps the greatest contributor of CD4⁺ T cell loss [85]. As the immune system becomes progressively exhausted, there is a decline in its ability to regenerate depleted CD4⁺ T cells. The persistence of HIV-1 replication stimulates the proliferation and differentiation of HIV-1-specific CD8⁺ T cells and they become progressively exhausted, losing cytolytic and cytokine-secreting ability, which ultimately results in their deletion [82]. Thus the ability to control HIV-1 replication is further impaired, leading to further CD4⁺ T cell depletion; however, high levels of virus replication continue to be supported at this stage by macrophages [78]. The high level of pro-inflammatory cytokines may also promote a range of disorders, such as osteoporosis, atherosclerosis, and neuronal injury [85].

1.4.3 Advanced disease

When CD4⁺ T cell counts decrease below 200 cells/mm³, the host becomes highly susceptible to opportunistic infections (*e.g.* oral candidiasis, pneumococcal infections, and tuberculosis) and certain cancers resulting mainly from viral infections (*e.g.* Kaposi's sarcoma caused by human herpes virus-8 and lymphomas caused by Epstein-Barr virus) [74]. Once CD4⁺ T cell counts are less than 200 cells/mm³ and/or one or more AIDS-defining illnesses are present, the patient is considered to have AIDS [1]. The average time from infection to AIDS is 8-10 years, but this may vary considerably due to host and viral factors (Section 1.6). At this stage of advanced disease, the immune system is overwhelmed and death ensues as a result of infections.

1.4.4 Host immune responses and viral evasion

Although host immune responses are mounted to HIV-1 from the earliest stage, they are unsuccessful in eliminating infection. The responses of different immune cells during acute and chronic infection and the ability of HIV-1 to subvert, evade, or use these immune responses to sustain infection are next discussed.

1.4.4.1 Innate immunity

1.4.4.1.1 Dendritic cells

In acute infection, dendritic cells are activated through cytokines secreted by mucosal epithelial cells or directly through TLR-mediated recognition of HIV-1 RNA [88] (Section 1.4.1). Activation results in the secretion of pro-inflammatory cytokines, importantly interferon- α (IFN- α), which activates a range of antiviral pathways [89] and upregulates restriction factors (Section 1.4.4.1.4). IFN- α as well as antigen presentation by dendritic cells activates CD4⁺ T cells, which co-ordinate adaptive immune responses [79, 88, 89] (Section 1.4.4.3.2). HIV-1 counteracts these antiviral effects by Vpr-mediated downregulation of IFN- α production and through impairing the function of dendritic cells [89, 90]. However, IFN- α also contributes to disease progression by recruiting susceptible cells to the site of infection and by promoting chronic immune activation [75, 89].

1.4.4.1.2 Macrophages

Macrophages are also involved in activating innate and adaptive immune responses through cytokine secretion and antigen presentation [78]. However, HIV-1 can interfere with antigen presentation by macrophages, shift macrophages to a pro-inflammatory phenotype to promote chronic immune activation, impair macrophage phagocytic function, and induce

(via Nef) macrophages to secrete factors that permit the infection of resting T cells, thereby increasing the latent reservoir [73, 78].

1.4.4.1.3 Natural killer (NK) cells

In early infection, NK cells may be recruited and activated by cytokines secreted by dendritic cells and macrophages [88, 89, 91]. NK cells may produce CCR5 ligands (such as macrophage inflammatory protein-1 α [MIP-1 α]) that compete with HIV-1 for the CCR5 co-receptor and they may also kill infected cells via release of perforin (which punches holes in the plasma membrane) and granzymes (which initiate cell death by apoptosis) [74, 88, 89]. The binding of inhibitory or activating killer cell immunoglobulin-like receptors (KIRs) on NK cells to HLA molecules on infected cells will determine whether killing is inhibited or promoted [89, 91]. HIV-1 evades this recognition by Nef-mediated selective downregulation of the expression of HLA-A and HLA-B molecules, but not HLA-C or HLA-E molecules (which interact with NK cells to avoid killing) [91, 92]. NK cells may also recognise infected cells by binding via their fragment crystallisable (Fc) receptors to antibody bound to the surface of infected cells (antibody-dependent cell-mediated cytotoxicity [ADCC]), however production of antibody only occurs approximately 3 weeks following infection [1] (Section 1.4.4.2). NK function becomes increasingly impaired in chronic infection, with the accumulation of a dysfunctional CD56⁻ CD16⁺ NK cell population [88].

1.4.4.1.4 Host restriction factors

To date, three major retroviral restriction factors have been described, including APOBEC3G (Section 1.3.2), TRIM5 α (Section 1.3.3), and tetherin (acts to tether virions to

the cell surface, preventing their release) [36]. These restriction factors are upregulated by IFN- α (and other type 1 IFNs), which is secreted during acute infection [90] (Section 1.4.4.1.1). However, mutations in the Gag capsid protein may result in evasion from TRIM5 α activity [36]. Furthermore, specific viral factors antagonise the action of restriction factors, namely Vif targets APOBEC3G for degradation (Section 1.3.2) and Vpu sequesters tetherin from the site of viral budding [36].

1.4.4.2 Humoral immunity

Immune complexes consisting of viral proteins and antibody can first be detected at approximately 18 days post-infection, followed by free antibody to Env gp41 (23 days post-infection) and Env gp120 (38 days post-infection) [72]. These initial antibody responses are non-neutralising and fail to block HIV-1 entry into cells. However, non-neutralising antibodies can function to recruit (via their constant region) NK cells (via their Fc receptors) to lyse infected cells (ADCC) [91] (Section 1.4.4.1.3). With the progression of disease and persistent viral replication, the ability of antibodies to induce ADCC is impaired [93].

Neutralising antibodies that block viral entry into cells develop at approximately 12 weeks post-infection. However, this response is often narrowly directed to the variable Env regions of the infecting viral strain [72], and HIV evades this response by the selection of viral variants that are not easily neutralised. Mutations in the viral sequence arise at a high rate due to the low fidelity of HIV-1 reverse transcriptase (1 error per 10,000 nucleotides) and the high replication rate [36], leading to viral variants. The mutations that are present in the viral variants selected by the immune response are known as “escape mutations”. As novel

neutralising antibodies develop, the mutant virus escapes again, thus the response continues to lag behind viral evolution [93, 94].

Broad-specificity neutralising antibodies (that target conserved Env regions shared by diverse viral isolates) are rare, may only develop after 20-30 months of infection [72], and even then are generally not associated with control of viremia [93]. The rarity of broadly neutralising antibodies can be explained by the variable carbohydrates (which can shift by point mutations) masking the Env glycoprotein and external loops masking the more conserved Env domains, which are only accessible for a short time during viral entry [36, 79]. The delay in the development of broadly neutralising antibodies could also be due to impaired antibody maturation (required for high affinity antibody development) resulting from extensive damage to germinal centres in lymphoid tissue during acute infection as well as impaired CD4⁺ T cell help [93]. Further dysregulation in humoral immunity in HIV-1 infection includes abnormal B cell activation and the development of autoantibodies [93].

1.4.4.3 Cell-mediated immunity

1.4.4.3.1 CD8⁺ T cells

HIV-1-specific CD8⁺ T cell responses have been detected as early as Fiebig stage I and II, however these responses appear to peak as viral load begins to decline [72] (Section 1.4.1). The coincident peak of CD8⁺ T cell responses with viral decline suggests that they are an important contributor to the inhibition of virus replication in early infection [95]. This is further supported by experiments that demonstrated a marked increase in SIV viremia following depletion of CD8⁺ T cells in rhesus macaques [96, 97], while depletion of

antibody-generating B cells in an animal model had little impact on viral load decline in primary SIV infection [74, 98].

Cytotoxic CD8⁺ T cells recognise HIV-1 infected cells and subsequently eliminate them via the secretion of perforin and granzymes (reviewed in [91]). This recognition is dependent on the presentation of viral epitopes (typically 8-10 amino acid sequences produced by processing of viral proteins within the cytoplasm) by HLA class I molecules (encoded by the Major Histocompatibility Complex [MHC] locus) on the surface of infected cells to the CD8⁺ T cell. The CD8 molecule on the T cell interacts with a constant region on the HLA molecule, and assists binding of the T cell receptor (TCR) to the epitope-HLA complex. This interaction is sufficient to trigger killing if the CD8⁺ T cell is already a differentiated effector cell, however naïve CD8⁺ T cells also require co-stimulatory signals (and this may be aided by CD4⁺ T cells) to become activated (Section 1.4.4.3.2). It should be noted that activated CD8⁺ T cells also produce antiviral cytokines (*e.g.* IFN- α) and chemokines (*e.g.* MIP-1 α and MIP-1 β , which are CCR5 ligands), and recent evidence suggests that these non-cytolytic mechanisms of viral inhibition may in fact be more important than direct cell killing [99, 100].

The HLA class I molecule is a heterodimer of β 2-microglobulin, which is a constant region shared by all class I proteins, and an α -chain, which is the variable region that recognises and non-covalently binds epitope [91]. Every individual expresses 6 HLA class I alleles (two HLA-A, two HLA-B, and two HLA-C alleles) and 6 HLA class II alleles (two HLA-DR, two HLA-DP, and two HLA-DQ alleles) (Section 1.4.4.3.2), of which there are currently 4,721 and 1,353 variant alleles, respectively, in total [101]. Each HLA molecule has a defined range of epitopes to which it can bind that is largely dependent on the epitope

residues that make direct contact with the HLA molecule (anchor residues) [91]. Thus, the HLA class I profile of an individual determines which viral epitopes mediate cytotoxic CD8⁺ T cell recognition and killing of infected cells, as well as other CD8⁺ T cell-mediated antiviral effects.

HIV-1 can counteract CD8⁺ T cell recognition through Nef-mediated selective downregulation of HLA class I molecules [36, 92]. Another important mechanism of evasion is the development of escape mutations in or adjacent to epitopes presented by the host HLA molecules. These mutations may impair intracellular processing, disrupt the binding of the epitope to the HLA molecule, or disrupt recognition of the presented epitope by the TCR [102-105]. CD8⁺ T cell escape mutations are rapidly selected as viremia declines, indicating that CD8⁺ T cell-mediated immune pressure makes an important contribution to this decline, while Env mutations selected by neutralising antibodies only appear approximately 12 weeks following infection [72, 106].

Mathematical modelling has estimated that 15-35% of HIV-infected cells are killed by a single CD8⁺ T cell response (of a few responses present) in early infection [72, 106] while only 4-6% of infected cells are eliminated by CD8⁺ T cells during chronic infection [107]. Therefore, the CD8⁺ T cells present in later infection are not as efficient as those in early infection, despite targeting a broader array of viral sequences (approximately 10 epitopes) [82, 108]. There may be several reasons for this impairment. The substantial loss of CD4⁺ T cells in acute infection results in decreased HIV-1-specific CD4⁺ T cell responses within the first few months and, therefore, a lack of CD4⁺ T cell help for subsequent CD8⁺ T cell responses [82]. CD4⁺ T cells are required for the maintenance of CD8⁺ T cell responses after acute infection [109], for their progression into long-term memory cells [72], and to

prevent rapid exhaustion [82]. Chronic high levels of HIV-1 antigen may also result in exhaustion, loss of function, and elimination of CD8⁺ T cells (Section 1.4.2). The resulting immune activation from HIV-1 infection may also impair CD8⁺ T cell responses by upregulating expansion of Fox P3⁺ CD25⁺ regulatory T cells, which inhibit the immune system by undetermined mechanisms that may include interleukin (IL)-10 and transforming growth factor- β secretion [82]. HIV-1 Nef might also aid in the evasion of CD8⁺ T cell responses through impairment of CD4⁺ T helper cell responses [36].

1.4.4.3.2 CD4⁺ T cells

HIV-1-specific CD4⁺ T cell responses are present in acute infection but decline rapidly [72, 82]. They provide help to innate immune cells, CD8⁺ T cells and B cells (reviewed in [91]). Antigen presenting cells (APC; macrophages, dendritic cells, and B cells) engulf HIV-1 antigen into a phagolysosome where it is processed into peptides of 13-17 amino acids. These peptides are then presented by HLA class II molecules on the APC surface to CD4⁺ T cells. The CD4 molecule on the T cell interacts with the HLA class II constant region and enhances binding of the TCR to the epitope-HLA complex. A co-stimulatory signal, principally the binding of B7 on the APC to CD28 on the CD4⁺ T cell, is required for activation of naïve CD4⁺ T cells. The activated CD4⁺ T helper 1 cells produce IL-2 which is required for the proliferation and differentiation of CD8⁺ T cells that have encountered antigen and other cytokines, such as IFN- γ , which promote the activity of macrophages. The CD4⁺ T helper 1 cells may also initiate co-stimulatory signals that are required for full activation of naïve CD8⁺ T cells encountering antigen. The activated CD4⁺ T helper 2 cells produce cytokines (*e.g.* IL-4 and IL-5) which stimulate proliferation of B cells that have encountered antigen and the production of soluble antibodies.

HIV-1 impairs all arms of the immune system through the depletion of CD4+ T helper cells (Section 1.4.2) with preferential infection and elimination of HIV-1-specific CD4+ T cells [110]. The depletion of uninfected CD4+ T cells is partly mediated by Nef-mediated upregulation of Fas ligand on infected cells that triggers apoptosis when binding to Fas on uninfected cells [1]. HIV-1 Nef further impairs CD4+ T cell function by inhibiting HLA class II epitope presentation and downregulating CD28 [36]. The preferential infection of the memory subset of CD4+ T cells also impairs immunity to previously encountered pathogens [111].

The general impairment of immune responses contributes to the gradual decline in immune function, the end result of which is death. Section 1.4 is summarised in Figure 1.2.

1.5 Therapy and preventative strategies

1.5.1 Antiretroviral therapy and prevention

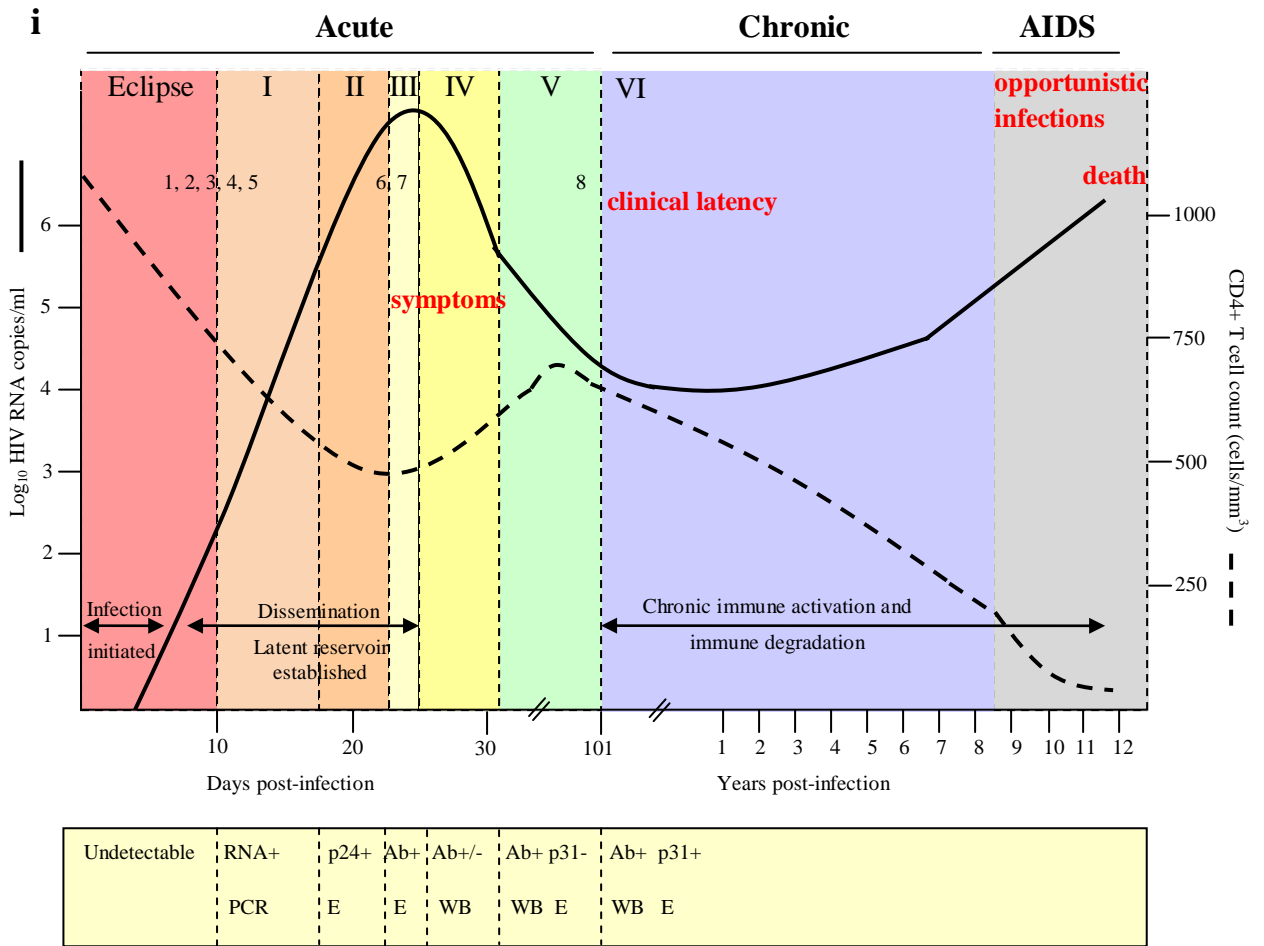
Since combination antiretroviral therapy (ART) was introduced in 1996, HIV-related morbidity and mortality has fallen and life-expectancy has increased significantly for those receiving treatment [83, 112]. Currently, there are 25 single drugs in 6 different classes, targeting various stages of the HIV-1 replication cycle (Section 1.3), approved for use: nucleoside reverse transcriptase inhibitors (NRTIs), non-nucleoside reverse transcriptase inhibitors (NNRTIs), protease inhibitors (PIs), entry/fusion inhibitors (FIs), CCR5 antagonists, and integrase inhibitors [112, 113]. First line therapy consists of a combination of 3 drugs - 2 NRTIs and a third drug from another class [112]. Multiple options are available for second line therapy should drug toxicity, drug-drug interactions, and/or resistance necessitate this [112].

Figure 1.2 Pathogenesis of HIV-1 infection

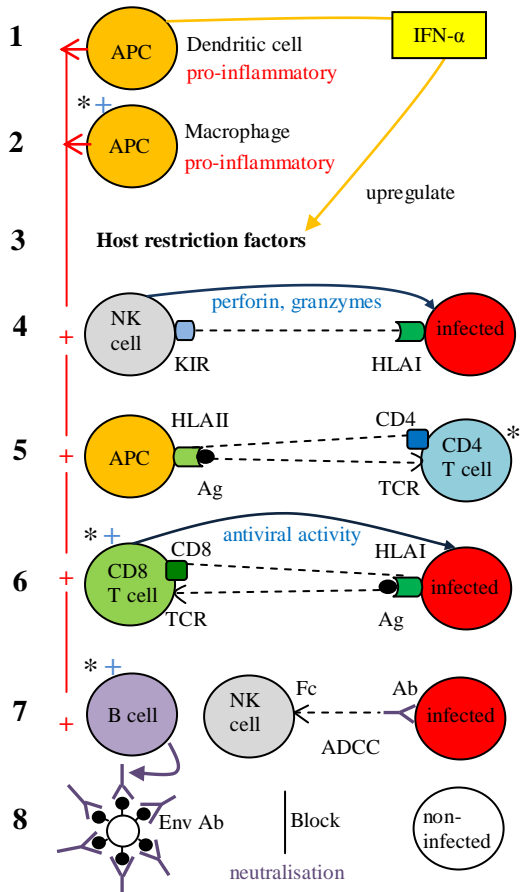
(i) The graph shows viral load and CD4⁺ T cell count patterns, pathogenic events, approximate timing of initiation of immune responses, clinical manifestations and approximate time frames of the different stages of HIV-1 disease (acute phase, chronic phase, AIDS). Roman numerals indicate the Fiebig stages I-V of acute infection and the open-ended stage VI of early chronic infection. These Fiebig stages are characterised by the detection of HIV-1 RNA by PCR, p24 antigen by ELISA, and antibodies by Western blot or ELISA (shown below the graph). The graph is adapted from McMichael *et al.* (2010) [72] and Kuritzkes and Walker (2007) [74].

(ii) The different host immune responses are depicted and some of the mechanisms by which HIV-1 evades, damages, and uses the host immune responses to its advantage are listed. Dotted lines in immune response diagrams indicate binding, plus signs (+) indicate enhancement of activity and the asterisks (*) indicate the cells influenced by CD4⁺ T helper cells.

HIV-1 – human immunodeficiency virus type 1; AIDS – acquired immunodeficiency syndrome; RNA – ribonucleic acid; CD – cluster of differentiation; PCR – polymerase chain reaction; p24 – protein of 24 kDa; Ab - antibody; p31 - protein of 31 kDa; E – enzyme-linked immunosorbent assay (ELISA); WB – Western blot; APC – antigen presenting cell; IFN- α – interferon alpha; NK cell – natural killer cell; KIR – killer cell immunoglobulin-like receptor; HLA – human leukocyte antigen; Ag - antigen; TCR – T cell receptor; Fc – fragment crystallisable; ADCC – antibody-dependent cell-mediated cytotoxic activity; Vpr – viral protein R; Nef – negative regulation factor; Vif – viral infectivity factor; Vpu – viral protein U; Env – envelope glycoprotein.



ii Immune responses



Immune evasion/degradation

- Downregulation of IFN- α (Vpr), impaired dendritic function, inflammation assists the establishment of infection
- Pro-inflammatory environment facilitates establishment of infection, impaired phagocytosis, secretions promote resting T cell infection (Nef)
- Evade TRIM5 α by capsid mutations, evade APOBEC3G by degradation (Vif), evade tetherin by sequestration (Vpu)
- Selective downregulation of HLA molecules (Nef)
- Preferential infection of HIV-specific CD4+ memory T cells, triggering of apoptosis in uninfected cells (Nef), impaired antigen presentation by APC (Nef)
- Downregulation of HLA (Nef), escape mutations, exhaustion and deletion of CD8+ T cells, impaired CD4+ T cell help
- Germinal centre damage, abnormal B cell activation, production of autoantibodies, impaired CD4+ T cell help
- Escape mutations, variable loops and carbohydrates masking conserved epitopes, germinal centre damage, impaired CD4+ T cell help

There are several challenges with ART. Even with highly potent therapy that suppresses viral load to undetectable levels there is low-level persistent residual HIV-1 replication from latently infected cells that become periodically activated, long-lived macrophages, and cells in anatomical sites such as the brain where penetration of drugs is limited [73, 78, 83, 112]. It is estimated that it would take 70 years of highly effective ART to purge this reservoir [74]. There are intensive efforts investigating the possibility of curing HIV-1 by activating latent reservoirs in combination with ART [83], and there is already an example of an HIV-1 cure in a man who received a bone marrow transplant from a CCR5 negative donor [114]. However, in the absence of a cure, complex treatment regimens that are difficult to maintain for long periods must be adhered to for life [115], with the high likelihood of drug resistance since greater than 95% adherence is required to avoid this. Continuous use of ART together with the persistent immune activation despite therapy increase the risk for cardiovascular disease, metabolic disorders, neurocognitive abnormalities, liver and renal disease, bone disorders, malignancy, and frailty [83]. Therefore, people on treatment do not reach a full life expectancy. In Denmark, an HIV-1 infected individual on therapy is half as likely to reach age 70 [116]. Furthermore, long-term ART for every infected individual for their lifetime is not sustainable due to high costs, and current coverage in low- and middle-income countries is only 40% [83].

Some novel ideas for therapeutics are being investigated, such as therapeutic vaccines and gene therapy. Two recent therapeutic vaccine trials of an adenovirus serotype 5 HIV-1 Gag vaccine and autologous dendritic cells pulsed with autologous heat-inactivated HIV-1 resulted in modest reductions in viral load and are therefore promising [117]. Delivery to mice of hematopoietic stem cells in which CCR5 alleles were disrupted (by a Zn finger nuclease) at a mean frequency of 17%, resulted in preferential expansion of CCR5-negative

cells and HIV-1 control [118]. Thus, Zn finger nuclease treatment of patient hematopoietic stem cells is also a possible treatment approach.

Since delivery of ART is failing to keep pace with the number of infected individuals, several preventative approaches are being investigated to reduce the number of new HIV-1 transmissions. These include male circumcision, risk reduction counselling, condom use, and pre-exposure or post-exposure ART prophylaxis [119]. Recently 39% efficacy of protection from HIV-1 acquisition was shown in women using a 1% tenovifir gel before and after sexual intercourse [120]. However, the ideal way to combat the spread of HIV-1 is with an effective preventative vaccine. [94, 119, 121-123]. Historically vaccines have proven to be the most successful and cost-effective way to reduce the incidence of infectious diseases, such as polio, measles, mumps, rubella, hepatitis B, and influenza [94, 123]. The development of an effective HIV-1 vaccine has been elusive thus far.

1.5.2 Vaccines

Successful vaccines to infectious agents have resulted from empirical testing of immunogens leading to the development of a product that induces a natural, effective immune response, which serves to protect recipients from disease caused by future infections [94, 123]. The challenges to developing an HIV-1 vaccine are numerous [1, 94, 112, 123]. A protective HIV-1 vaccine would have to induce effective immune responses to clear the infection within the narrow window of a few days before a latent reservoir is established (Section 1.4.1). Further, these immune responses would have to be effective against the enormous diversity of HIV-1 strains, which far exceeds the diversity of influenza for which a new vaccine is needed every year [94]. Broadly neutralising antibodies could provide complete

protection from infection, however there has not yet been success in inducing these antibodies since the conserved epitopes to which they are directed are masked (Section 1.4.4.2) and poorly immunogenic [119]. Complete protection from HIV-1 infection may not be achieved, but even a vaccine that could elicit immune responses to suppress viral load to 1,000-2,000 HIV RNA copies/ml, could substantially slow disease progression, markedly reduce transmission, and therefore have enormous public health benefit [94, 122]. Such a vaccine would pose the challenge of continued suppression of viral load despite the high mutation rate of HIV-1 which leads to escape of immune responses [123] (Section 1.4.4.3.1).

To date, greater than 30 candidate HIV-1 vaccines have been tested in human clinical trials, the most recent ones being the Step/Phambili and RV144 trials [119, 124-126]. All vaccine trials, with the exception of the RV144 trial, have failed to impact HIV-1 acquisition, and all failed to influence disease progression in vaccinees who acquired HIV-1 [119]. The Step (conducted in America and Australia) and Phambili (conducted in South Africa) trials evaluated the efficacy of a T cell only vaccine, consisting of an adenovirus serotype 5 vector containing the HIV-1 *gag*, *pol*, and *nef* genes [127]. The reasons for the failure of this T cell only vaccine are not fully understood, but may be related to non-optimal specificity and breadth of CD8⁺ T cell responses elicited [121, 127]. The RV144 trial, which tested a recombinant HIV-1 (*env-gag-protease*)-canarypox vector prime and a recombinant gp120 plus alum boost, was more successful, showing a 31% efficacy in protection from HIV-1 acquisition [126]. CD4⁺ T cell Env responses, ADCC, antibody binding to gp120, and low titre neutralising antibodies were detected [126, 127]. It has been suggested that the modest protective effect was likely mediated by non-neutralising antibodies, suggesting the possible importance of innate immunity in preventing infection [93]. The general consensus is that an effective vaccine will have to elicit B cell, CD4⁺ T cell and CD8⁺ T cell responses, and

include adjuvants that stimulate innate immune responses [119]. However, the results of the vaccine trials highlight the lack of understanding of what constitutes an effective immune response against HIV-1 and how to elicit it.

There is a wide range in the time of progression to AIDS amongst HIV-1 infected individuals (Section 1.6.1) with a small minority spontaneously suppressing viral loads to low or undetectable levels indefinitely. Understanding the mechanisms of differential disease progression rates and natural control of HIV-1 replication to low levels could result in the identification of effective immune responses and the design of an effective vaccine. Therefore, a review of the factors determining degree of control of viral replication and HIV-1 disease progression follows.

1.6 Factors influencing the rate of disease progression

1.6.1 Spectrum and measurement of disease progression rate

There is a wide range in the ability of infected individuals to control HIV-1 replication and therefore the rate of disease progression varies greatly between individuals. The average time for progression to AIDS is 8-10 years [128]. However, approximately 5% of individuals progress to AIDS within 2-3 years [129, 130] while another 5-8%, termed long-term non-progressors (LTNPs), are able to maintain low viremia (mostly defined as less than 2,000 HIV RNA copies/ml) and stable CD4+ T cell counts for at least 5 years without ART [131, 132]. A subset (less than 0.2% of the total HIV-1-infected population) of LTNPs, termed elite controllers, suppress viral load to undetectable levels (less than 50 HIV RNA copies/ml) indefinitely without treatment and generally do not have clinical symptoms [131, 132].

Many studies assessing the influence of various factors on HIV-1 disease progression have used time to AIDS as a direct endpoint to measure disease progression rate. Others have used indirect markers of disease, such as the viral set point following seroconversion or within the first 6-12 months, which are strong predictors of progression to AIDS [74, 133]. Viral load and immune activation (as measured by the CD38 marker on CD8+ T cells) are the strongest predictors of time to AIDS (predicting 47% and 40% variability in time to AIDS, respectively) followed by CD4+ T cell count (predicting 29%) [130, 134-136]. Thus, early viral set point, viral load, and CD4+ T cell count are often used as surrogate markers of disease progression [137]. CD4+ T cell decline has an inverse relationship to plasma viral load [130] but recently was found to explain only about 3% of variability in time to AIDS due to the high variability in the CD4+ T cell decline measurement [135]. CD4+ T cell decline is also sometimes used as a marker of disease progression, although usually in conjunction with other markers.

1.6.3 Viral and host factors

Several host genetic and immune factors as well as viral genetic factors may influence the rate of HIV-1 disease progression [79, 129, 132, 138-140]. For example, individuals who are homozygous for a 32-bp deletion in the CCR5 co-receptor gene show almost complete protection against HIV-1 acquisition (since the transmitted virus is highly dependent on this co-receptor for cell entry), and those who are heterozygous for the deletion or have polymorphisms that decrease CCR5 expression have delayed progression to AIDS [129, 141]. Polymorphisms in the CypA (Section 1.3.3) and Tsg101 (Section 1.3.8) host genes involved in HIV-1 replication have also been associated with differences in susceptibility to HIV-1 infection or in the rate of progression to AIDS [129, 142]. However, the majority of host genetic factors linked to altered disease progression mediate their effect through

influencing host immune responses to HIV-1. Similarly most viral genetic factors associated with slower or faster rates of disease progression are themselves consequences of host immune responses and/or affect pathogenesis through altering the effectiveness of host immune responses. Therefore the host and viral factors influencing HIV-1 disease progression will be discussed in terms of the innate, humoral and cell-mediated immune responses.

1.6.3.1 Innate immune responses

LTNPs maintain higher levels of dendritic cells [131], which play an important role in the innate immune response (Section 1.4.4.1.1). Dendritic cells produce IFN- α (Section 1.4.4.1.1) and two single-nucleotide polymorphisms (SNPs) in the IFN- α receptor have been associated with susceptibility to HIV-1 infection [143]. Polymorphisms in restriction genes upregulated by IFN- α and other type 1 IFNs, namely TRIM5 α and APOBEC3G, have been linked to altered disease progression [144, 145]. However, it appears that polymorphisms in identified restriction factors are not the cause of viral control in the majority of elite controllers [131].

Increased NK cell activity has been detected in exposed but seronegative individuals [146] and factors influencing NK cell activity have also been linked to altered disease progression. A SNP in the HLA-C promoter gene has been identified as a major genetic determinant associated with lower viral set point, and may mediate its effect through influencing interactions with NK or CD8⁺ T cells leading to more efficient cell lysis [147]. However, this SNP does not explain elite control in the majority of elite controllers [131, 148]. The expression of the activating KIR3DS1 (on NK cells) in combination with HLA-Bw4

molecule ligands (including HLA-B*57 and HLA-B*27) is associated with slower progression to AIDS [149], and is overexpressed in LTNPs but not elite controllers [131]. Long-term non-progression in blood transfusion recipients of a Nef-deleted virus [150, 151] may be partly accounted for by lack of Nef-mediated downregulation of HLA, which antagonises both NK cell and CD8⁺ T cell immune responses (Sections 1.4.4.1.3 and 1.4.4.3.1), although viral attenuation could also be a contributing factor (Section 1.6.5). However, the majority of elite controllers do not harbour viruses with gross genetic defects such as Nef-deletions [152].

1.6.3.2 Humoral immune responses

There is not much data to suggest that antibody responses determine control of HIV-1 replication, although there is more evidence that they may be important in preventing infection [93, 131, 132]. In a macaque-SIV/HIV model, passively administered antibodies can protect against infection and disruption of ADCC activity dramatically increases susceptibility to SIV/HIV chimera infection [153]. Further, the modest protection against HIV-1 acquisition by the RV144 vaccine may have been elicited by non-neutralising antibodies [93] (Section 1.5.2). Mucosal immunoglobulin A (IgA) antibodies capable of neutralisation were also identified in exposed but seronegative individuals [146].

It has been shown that ADCC activity is more potent in elite controllers than in progressors [93]. However, since equally potent ADCC activity was shown in some acutely infected individuals and individuals on ART, potency may be a consequence rather than cause of level of viremia [93]. Broadly neutralising antibodies are less common in aviremic LTNPs [154] and elite controllers have lower titres of broadly neutralising antibodies and similar

levels of autologous neutralising antibodies when compared with progressors [155, 156]. Collectively these data suggest that neither ADCC nor neutralising antibodies play a major role in maintaining viremic control in HIV-1 controllers.

1.6.3.3 Cell-mediated immune responses

1.6.3.3.1 CD4+ T cell responses

HIV-specific CD4+ T cell responses of elite controllers and LTNPs have a higher proliferative potential than those of progressors, and also result in the secretion of multiple cytokines, including IL-2, upon stimulation, while CD4+ T cells from progressors mostly secrete IFN- γ [90, 131, 157]. Further, there are preserved central memory and activated effector memory CD4+ T cell subsets in HIV-1 controllers [158]. The preservation of a vigorous CD4+ T cell response in HIV-1 controllers may be important for CD8+ T cell-mediated control of virus replication, but whether or not it is crucial is unknown [82]. However, a recent study has shown that IL-21-secreting CD4+ T cells (preserved in elite controllers) may contribute to viral control through enhancing CD8+ T cell function [159]. It is also unclear whether preserved CD4+ T cell responses in controllers are a cause or consequence of low viremia and there is conflicting data in this regard [82, 90, 131]. It is clear at least that the proliferative capacity of HIV-specific CD4+ T cells can be restored by ART to levels observed in LTNPs, suggesting that this characteristic is influenced by the level of viremia [90, 160].

1.6.3.3.2 CD8+ T cell responses

As with HIV-specific CD4+ T cell responses, there are qualitative differences in HIV-specific CD8+ T cell responses between elite or viremic controllers and progressors. HIV-

specific CD8⁺ T cells from elite controllers and/or LTNPs are polyfunctional (secreting multiple cytokines) [161], have a high proliferative capacity when stimulated [162], are more efficient at lytic granule loading, and have a high per-cell killing capacity [163]. Some studies have found restoration of CD8⁺ T cell polyfunctionality by ART, suggesting that polyfunctionality is a consequence rather than cause of low viremia [164, 165]. However, these studies measured polyfunctionality in response to antigen stimulation and did not compare to LTNPs. This may account for the different result obtained in a recent study in which CD8⁺ T cell polyfunctionality in response to autologous HIV-1-infected CD4⁺ T cells was compared in patients on ART versus LTNPs with matched suppressed viral loads, where polyfunctionality was greater in LTNPs [160]. However, these authors argue that polyfunctionality is not likely to be an important determinant of immune control as polyfunctional cells form a small subset of the total HIV-specific CD8⁺ T cell response [160]. Proliferative and cytotoxic capacities of CD8⁺ T cells were superior in LTNPs when compared with patients on ART and these characteristics may contribute to immune control of HIV-1 [131, 160]. Interestingly, HIV-1 specific CD8⁺ as well as CD4⁺ T cell responses have been detected in HIV-1 exposed but seronegative individuals [146, 166].

The important role that the CD8⁺ T cells play in reducing viral load from a peak in early infection to a lower viral set point (Section 1.4.4.3.1) and the strong association of this set point with progression to AIDS (Section 1.6.1), as well as the CD8⁺ T cell depletion experiments in monkeys (Section 1.4.4.3.1), provide evidence that the CD8⁺ T cell responses are important in determining viral control and disease progression. However, perhaps the strongest evidence supporting this is the consistent, strong association of different HLA class I alleles, which present viral epitopes to activate CD8⁺ T cell mediated killing of infected cells (Section 1.4.4.3.1), with the rate of progression to AIDS [132, 167-

172]. A recent genome-wide association study on nearly 1,000 HIV-1 controllers revealed 313 SNPs associated with control, all of which were within the region coding for HLA concentrated in and around class I genes, suggesting it is the major factor involved in determining viral control [169].

1.6.4 HLA class I alleles

1.6.4.1 Genetic associations

Although there are some associations between HLA class II alleles and rate of HIV-1 disease progression (*e.g.* DRB1*13 is associated with slower disease progression) [167, 172, 173], HLA class I alleles are clearly the strongest and most consistent markers of disease progression [132, 167]. HLA-B*57 (especially B*5701) and HLA-B*27 have been most robustly associated with slower progression to AIDS in European and North-American individuals, while HLA-B*35-Px is consistently associated with rapid progression to AIDS in these populations [167]. In African populations, alleles consistently and strongly identified with low viral loads and/or high CD4+ T cell counts and/or are overrepresented in LTNPs include HLA-B*57 (especially B*5703), HLA-B*5801, and HLA-B*8101 [170, 174, 175]. In contrast, in these African populations, HLA-B*5802 and HLA-B*18 are significantly associated with higher viral loads [170, 174, 176]. In addition, rare HLA class I types have been associated with slower disease progression and HLA class I homozygotes have been associated with faster disease progression [132, 167]. There is also some evidence that HLA types may be associated with differing resistance to infection [167], for example HLA-B*57 expressed in combination with KIR3DL1*h/*y is more prevalent in exposed seronegative individuals [177]. The underlying basis for the strong association between HLA class I alleles and disease progression (and/or resistance to infection) is not fully understood. Some possible mechanisms are outlined below.

1.6.4.2 HLA: mechanisms of control

1.6.4.2.1 Interactions with NK cells

HLA-B class I alleles may have either a Bw4 or Bw6 specificity. Most HLA alleles associated with protective effects have Bw4 specificity while those associated with rapid disease progression more often have Bw6 specificity [167]. Since co-expression of Bw4 alleles, like HLA-B*57, and KIR3D molecules (on NK cells) are associated with slower disease progression (Section 1.6.3.1) or resistance to infection (Section 1.6.4.1), this suggests that an underlying mechanism of HLA protection may be interaction with NK cells.

1.6.4.2.2 Immune activation

One suggested mechanism of protection of certain HLA alleles is association with lack of immune activation. For example, the protective alleles HLA-B*57 and HLA-B*27 have been associated with lower levels of markers of immune activation and inflammation which may contribute to a clinical benefit in these individuals [178]. It has also been proposed that the sequence homology between HIV gp120 and regions of both HLA class I and II molecules may promote allogeneic T cell reactions and immune activation, and that this may differ depending on the HLA repertoire, thereby partly explaining HLA allele association with disease progression [87]. Consistent with this, antibodies to a conserved region of gp120 displaying structural homology with HLA were found in the sera of slow progressors [179]. Further supporting this idea, there is much evidence that immune activation plays a central role in determining development of AIDS: CD38 (a marker of activation) expression on CD8⁺ T cells predicts progression to AIDS to a similar degree as HIV-1 viral load in early infection and is the strongest predictor in later infection [134], CD4⁺ and CD8⁺ T cell activation predict progression to AIDS [180], polymorphisms in the CXCR6 receptor (a mediator of inflammation) are strongly associated with long-term non-progression to AIDS

[181], polymorphisms in genes encoding pro-inflammatory (*e.g.* tumour necrosis factor- α) and anti-inflammatory cytokines (*e.g.* IL-10) have been associated with altered rate of progression to AIDS [129], elite controllers have lower levels of HIV-specific CD8⁺ and CD4⁺ T cell activation [182], elite controllers have immune activation restricted to the T cell effector compartment and not a generalised pattern of immune activation [158], and natural hosts of SIV (African green monkeys and sooty mangabeys) have high viremia but low levels of immune activation and non-progression to AIDS [85, 86].

1.6.4.2.3 CD8⁺ T cell responses

However, Mamu-B*08 (similar to HLA-B*27) positive macaques were less likely to achieve elite control when infected with a strain of SIV encoding escape mutations in Mamu-B*08-restricted epitopes, suggesting that the protective mechanism of certain HLA alleles is likely CD8⁺ T cell-based [131]. This may be related to the quality or specificity of CD8⁺ T cell responses restricted by certain HLA alleles.

1.6.4.2.3.1 Immunodominance and quality

HLA alleles associated with slower progression to AIDS mediate strong, immunodominant CD8⁺ T cell responses (that make a major contribution to the total CD8⁺ T cell response) in early infection, which may explain their protective effect [183]. Further, individuals who display a larger contribution of the most immunodominant CD8⁺ T cell response for a given HLA allele to their total CD8⁺ T cell response in early infection generally have lower viral set points (even after exclusion of protective HLA alleles) [184]. Thus, strong CD8⁺ T cell responses in early infection may mediate viral control. The protective HLA-B*57 and HLA-B*27 alleles have a greater frequency of polyfunctional T cell responses, which may be due

to low avidity T cell receptors on these populations [185]. Another recent finding is that the nature of the HLA-B*57 peptide binding results in greater cross-reactivity and therefore greater potential to recognise epitopes containing point mutations at certain residues, which may allow more effective CD8⁺ T cell responses to emerging mutant viruses [186].

1.6.4.2.3.2 Gag-focussed specificity

1.6.4.2.3.2.1 Epitope restriction specificity

A feature of protective HLA alleles is the restriction of conserved epitopes in Gag [174, 187, 188]; in contrast to unfavourable HLA alleles, such as HLA-B*5802, which preferentially present more variable Env-derived epitopes to CD8⁺ T cells [176]. In addition, protective HLA alleles have been associated with immunodominant targeting of Gag [189, 190]. It is also of interest to note that even after excluding protective HLA alleles, CD8⁺ T cell responses to conserved epitopes across the HIV-1 proteome have been associated with lower viral loads [191, 192].

1.6.4.2.3.2.2 Gag T cell responses associated with viral control

While the overall breadth and magnitude of the CD8⁺ T cell response does not correlate with rate of disease progression [108], several studies have shown an association between preferential Gag-specific CD8⁺ T cell responses and favourable clinical outcomes. Kiepiela *et al.* (2007) showed an inverse correlation between the number of CD8⁺ T cell responses to the conserved Gag protein and viral load as well as a correlation between the number of responses to the highly variable Env and higher viral loads in chronic subtype C HIV-1 infection [193]. Similarly, the magnitude and breadth of CD8⁺ T cell responses to Gag or Gag capsid (p24) correlated inversely with viral loads and directly with CD4⁺ T cell counts

[194] and broad Gag-biased epitope repertoires were associated with lower viral loads [195]. Zuñiga and colleagues (2006) showed that the contribution of Gag p24 CD8+ T cell responses to the total CD8+ T cell response is positively associated with low viral loads and high CD4+ T cell counts [196]. Recently, Pereyra *et al.* (2008) showed that HIV-1 controllers preferentially target Gag over other viral proteins [156]. Superior *in vitro* suppression of viral replication in autologous CD4+ T cells using inhibition assays is mediated by Gag-specific rather than Nef- or Env-specific CD8+ T cell responses [197, 198]. However, it should also be noted that while some acute infection studies have shown that early CD8+ T cell responses to Gag result in lower viral set points [199], others have shown no correlation with viral set point [200] or no correlation between Gag-specific CD8+ T cell responses in early chronic infection and progression to AIDS [201].

The link between Gag-specific CD4+ T cell responses and viral control further supports the important role of Gag-specific CD8+ T cell responses in disease progression. HIV-1 controllers have a high avidity memory Gag-specific CD4+ T cell population that proliferate extensively in response to minimal amounts of antigen [202]. Gag-specific CD4+ T cell responses are maintained in LTNPs (although these responses after seroconversion had no prognostic value for progression to AIDS in that study) [203], and Gag p24-specific CD4+ T cell responses in early infection predicted lower viral set points [204]. CD4+ T helper 1 cells (which promote CD8+ T cell mediated immunity) that are predominantly directed to Gag are associated with elite control while the opposite is true for those predominantly directed towards Env, suggesting that Gag-specific CD8+ T cell responses are important in mediating viral control [205].

1.6.4.2.3.2.3 HLA-associated mutations in Gag associated with viral control

Protective HLA alleles are not only associated with presentation of Gag epitopes and preferential Gag-specific CD8⁺ T cell responses, but also with strong selection pressure resulting in the development of escape mutations in these key Gag epitopes [174, 187]. HLA-driven mutations that were associated with lower viral loads in chronic subtype C HIV-1 infection were found predominantly in Gag and were restricted by protective HLA alleles [206]. Further, in acutely HIV-1-infected individuals, increasing numbers of transmitted HLA-B-associated polymorphisms [207], or more specifically the presence of transmitted polymorphisms associated with the protective HLA-B*57 and HLA-B*5801 alleles [208], were associated with lower viral set points. Protective HLA-B alleles (linked to lower viral loads) were also associated with a higher number of mutations in Gag that revert following transmission to hosts without the relevant HLA alleles to select them when compared with HLA-B alleles linked to higher viral loads [209]. Reversion of HLA-associated mutations when transmitted to HLA-mismatched recipients (*i.e.* in the absence of their selection pressure) implies that these mutations carry a cost to viral fitness or replication capacity [209, 210].

Collectively, these studies indirectly suggest an important mechanism of HLA association with disease progression rate: CD8⁺ T cell responses restricted by favourable HLA alleles select for escape mutations in conserved epitopes, particularly those in Gag, resulting in a fitness cost to HIV-1 and this may at least in part explain slower disease progression in individuals with these alleles. In the next section, direct evidence for fitness cost of HLA-driven Gag mutations, as well as direct evidence for impact of viral fitness on HIV-1 disease progression, is discussed.

1.6.5 Viral fitness

1.6.5.1 Fitness costs of HLA-associated Gag escape mutations

A fitness cost for some escape mutations in Gag p24 capsid, a particularly conserved region of the Gag protein, that are associated with protective HLA alleles has been directly demonstrated through insertion of the escape mutation (by site-directed mutagenesis) into a laboratory-adapted strain of HIV-1 followed by *in vitro* measurement of viral replication. The HLA-B*5703/B*5801-associated T242N mutation in the TW10 epitope reduces recognition by CD8+ T cells [210], reverts on transmission to HLA-mismatched recipients [210], and decreases replication capacity *in vitro* [33, 187, 211]. Similarly, the A146P mutation alone or in combination with I147L in the HLA-B*5703/B*5801-restricted ISW9 epitope, and the A163G mutation (which arises following escape in TW10 and ISW9 epitopes) in the HLA-B*5703-restricted KF11 epitope, significantly reduced recognition by CD8+ T cells as well as viral replication *in vitro* [102, 104, 187, 212]. The protective allele HLA-B*27 is associated with selection of the escape mutation R264K, which dramatically reduces viral replication capacity *in vitro* [213].

Interestingly, HLA-B*57/B*5801-associated mutations are often selected early in infection while the HLA-B*27-associated R264K mutation is typically selected very late in infection and correlates with the onset of AIDS [168, 187, 213-215]. This suggests that the mechanism of slower disease progression mediated by different HLA alleles may be different. It has been argued that HLA-B*57 and HLA-B*5801 may mediate their protective effect through strong early immunodominant CD8+ T cell responses that result in rapid selection of Gag escape mutations and early viral attenuation, while HLA-B*27 may mediate its protective effect through an effective CD8+ T cell response that is maintained through a high genetic barrier to escape (due to the associated high viral fitness cost) [168, 187, 213].

A recent study has indicated that escape mutations in more conserved regions, such as Gag p24, often result in viral fitness costs while escape mutations in the more variable Env region often carry no fitness cost or actually increase viral fitness *in vitro* [104]. Mutations arising in Gag may be more likely to result in fitness costs since it is a key structural component of the virion and plays important roles during several stages in the HIV-1 replication cycle (Sections 1.2 and 1.3), therefore requiring conservation. Escape mutations occurring in conserved regions of Gag that are not restricted by the protective HLA alleles have also been shown to result in viral fitness costs, including: E207D alone or in combination with V215L (in Gag p24) in the A*25-restricted EW10 epitope [104], E17K (a conserved p17 residue) in a B*4002-restricted epitope, and K26R (a conserved p17 residue) in a B*08 restricted epitope [216].

For most of the escape mutations listed above, secondarily arising co-varying mutations that partially compensate for fitness costs have been described. Brockman *et al.* (2007) show that mutations at Gag codons 219, 223, 228, and 248 cumulatively increase the replication capacity of T242N mutants, although fitness is not restored to wild-type levels [33]. Interestingly, the underlying mechanism of the T242N fitness defect is disrupted interaction with CypA and these compensatory mutations conferred independence from CypA [33]. Troyer *et al.* (2009) showed that while I147L increased the replication capacity of a A146P mutant, the double mutant was still significantly less fit than the wild-type [104]. Similarly, S165N partially compensates for the fitness cost of a subtype C A163G mutant [212]. The S173A mutation is strongly linked to and arises in conjunction with the R264K HLA-B*27-associated escape mutation, and restores viral replication capacity to near wild-type levels [213].

1.6.5.2 Viral fitness and disease progression

1.6.5.2.1 Whole isolate fitness

Some of the first studies to suggest an impact of viral fitness on HIV-1 disease progression described unusually slow progression (or non-progression) to symptomatic disease in recipients of an attenuated strain of HIV-1 containing a deletion in the *nef* gene [150, 151]. However, the long-term outcome in these individuals might also be explained by the absence of the many antagonistic effects of Nef (Section 1.4.4). Some reports have shown distinct differences in the fitness of HIV-1 isolates from LTNPs and progressors (although isolates from both groups are replication competent) [217-219]. Further indicating that viral replicative ability influences rate of disease progression, a study showed that differences in viral set point after rebound following treatment interruption were driven by viral fitness [220].

There is also evidence that difference in fitness between HIV-1 isolates of different subtypes may influence disease progression. Slower rates of CD4⁺ T cell decline in HIV-1 subtype C infection when compared with subtype A and D infection have been reported, but slower disease progression in HIV-1 subtype C infected individuals has not been directly shown [221]. In a recent report, HIV-1 subtype C isolates were shown to have equal transmissibility but were dramatically less fit when compared to other M group subtypes [221]. These authors suggest that the global prevalence of HIV-1 subtype C may be explained by the lower pathogenic viral fitness of this subtype, which may result in slower disease progression and consequently greater opportunity for transmission.

1.6.5.2.2 Gag-mediated viral fitness

Some studies suggest that Env is the strongest determinant of HIV-1 fitness, explaining as much as 86% of the variability in viral fitness [222, 223], however Campbell *et al.* (2003) showed that protease-reverse transcriptase may explain 53% of the variability in HIV-1 fitness [224]. Recently, recombinant viruses encoding Gag p24 correlated with HIV-1 isolate replication capacity, providing direct evidence that Gag p24 is also a significant determinant of HIV-1 fitness [219].

There is evidence that reduced Gag function as a result of HLA-driven mutations, and thereby lowered viral fitness, might impact on the rate of disease progression. Recombinant viruses encoding Gag-protease from HIV-1 subtype B infected elite controllers were significantly attenuated [225]. In a small number of LTNP infants without HLA-B*57/B*5801 alleles, whole viral isolates as well as recombinant viruses encoding Gag p24 (from the infant's virus) had classical Gag p24 mutations associated with these protective HLA alleles (and a lack of compensatory mutations) as well as significantly reduced fitness [219]. In contrast, in progressor infants who received a virus with these Gag p24 mutations, compensatory mutations were present and the viral isolates were fitter [219]. A more recent study demonstrated reduced viral replication capacity, attributed to B*57/B*5801-associated mutations in Gag as well as drug-resistance mutations in protease, in HIV-1 subtype B acutely infected individuals of various HLA profiles who subsequently became HIV-1 controllers [226]. While these studies suggest long-term benefit in HLA-mismatched recipients of an attenuated virus encoding HLA-B*57/B*5801-associated escape mutations in Gag, other studies investigating HLA-mismatched recipients have observed reversion of these mutations within 2 years, resulting in restored fitness and loss of viral control [187, 227]. Reversion and loss of viral control suggest that an effective CD8⁺ T cell response

mediated by possession of the protective HLA allele is thus also required for maintenance of less fit viral variants and long-term control.

1.6.5.2.3 Gag-mediated viral fitness/CD8+ T cell response balance

There is much evidence to suggest that viral fitness alone cannot explain disease progression but rather that the balance between CD8+ T cell responses to the virus and mutations that arise as a result of this contributes to clinical outcome. Kawada *et al.* (2006) demonstrated loss of virus control in monkeys with the accumulation of Gag escape mutations in SIV, although SIV was progressively attenuated *in vitro* [228]. Similarly, Crawford *et al.* (2009) showed a decrease in *in vitro* fitness with the sequential accumulation of HLA-B*5703-associated mutations in Gag (A146P, I147L, T242N, A163G) and a loss in viral control simultaneous with evasion from effective CD8+ T cell responses to all of the HLA-B*5703-restricted Gag p24 epitopes [187]. Further, although elite controllers have viruses with significantly reduced Gag-protease mediated fitness [225], they may also select rare fitness-reducing mutations in Gag and have the ability to recognise and make strong *de novo* CD8+ T cell responses to these variants [229].

Contrasting the importance of the balance between CD8+ T cell responses and viral fitness in determining disease progression is a recent study by Schneidewind *et al.* (2009) investigating a small number of infants [227]. This study showed favourable clinical outcome and viral control in HLA-B*57 haplo-identical infants receiving a virus with HLA-B*57-selected mutations in Gag p24 even when CD8+ T cell responses to these epitopes were undetectable. It should however be noted that reversion of these mutations did not occur in these infants, suggesting their maintenance by low level CD8+ T cell responses that

were not detectable in the blood but may have been present in the lymph nodes or other compartments. Also, CD8+ T cell responses to B*57-associated variants have previously been detected in HLA-B*57 positive infants [230]. Further, in contrast to the findings of Schneidewind *et al.* (2009) [227], other groups have reported that individuals with HLA-B*57 or HLA-B*27 who received a virus with escape mutations in Gag from individuals with these same HLA alleles did not display a favourable clinical course [187, 231].

An overview of the factors influencing HIV-1 disease progression is illustrated in Figure 1.3.

1.7 Measurement of viral fitness or replication capacity

Depending on the question addressed, the fitness consequences of a particular viral mutation, viral genomic region, or whole viral isolate can be measured. The advantages and limitations of each of these measurements, as well as the advantages and disadvantages of different cell types and assays used to obtain viral fitness measurements, are discussed below.

1.7.1 Starting material

1.7.1.1 Mutant viruses

To measure the fitness effect of a particular mutation, this mutation may be introduced by site-directed mutagenesis into a laboratory-adapted strain of HIV-1 and the consequence determined by comparison of the mutant versus wild-type viruses [33, 187, 211]. Since a standard laboratory-adapted strain is used to introduce mutations, a reporter gene can be engineered into the strain to allow for ease of virus detection [232]. This may be for example, firefly luciferase (which is a luminescent marker) [232, 233] or jellyfish green

Figure 1.3 Spectrum and markers of HIV-1 disease progression rate and factors influencing the rate of progression

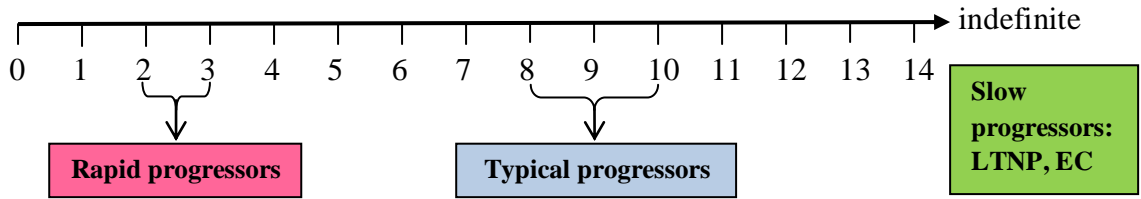
(i) Timeline indicating the wide variability in time of progression to AIDS in individuals infected with HIV-1.

(ii) The graph shows early viral set point as a predictor of HIV-1 disease progression and other commonly used and/or significant markers of disease progression rate are listed. The graph was adapted from Kuritzkes and Walker (2007) [74].

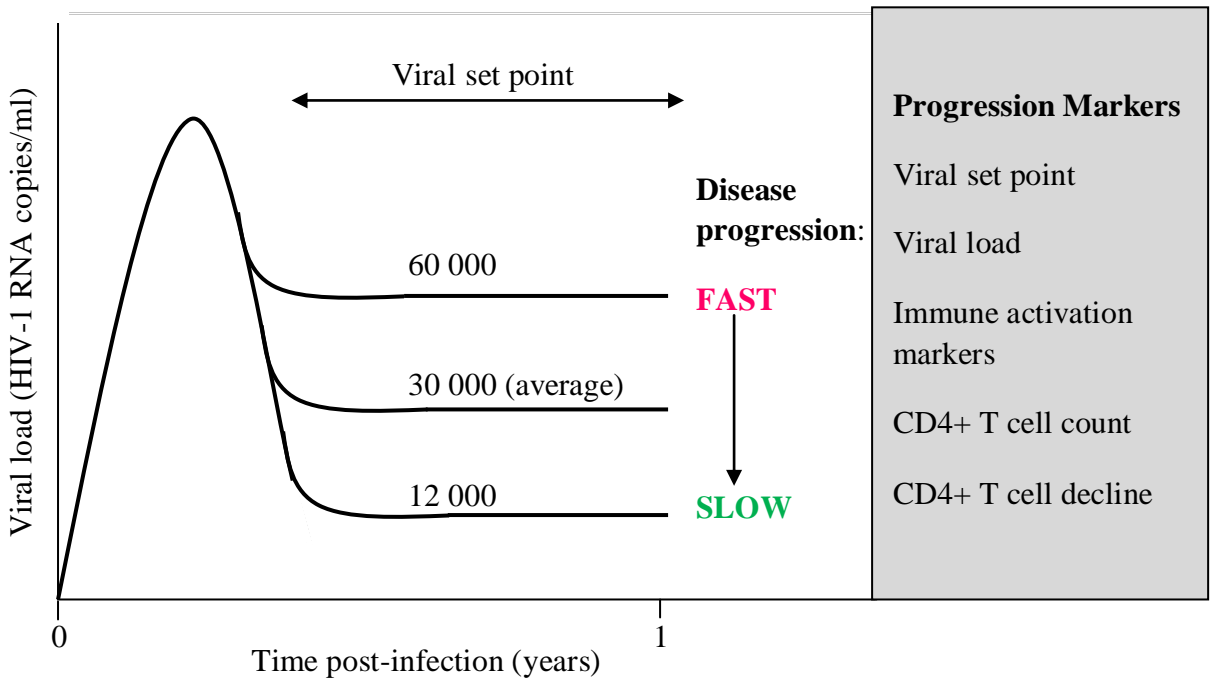
(iii) Overlapping and interacting host genetic, host immune and viral factors influencing the rate of HIV-1 disease progression are outlined. HLA class I alleles are most strongly associated with disease progression rate and interactions between HLA alleles and other immune factors which may partly explain the association are shown. Highlighted with red arrows is the interaction pathway suggesting that certain HLA alleles restrict Gag-specific CD8⁺ T cell responses which select for escape mutations in Gag that impact viral fitness and therefore disease progression. This pathway is highlighted, and HLA and viral fitness factors are underlined, since the focus of the present study is to investigate the influence of HLA-driven mutations in Gag-protease on HIV-1 fitness and therefore disease progression.

HIV-1 – human immunodeficiency virus type 1; AIDS – acquired immunodeficiency syndrome; LTNP – long-term non-progressors; EC – elite controllers; RNA – ribonucleic acid; CD – cluster of differentiation; CCR5 – C-C chemokine receptor 5; Tsg101 – tumour susceptibility gene 101 protein; IFN- α - interferon alpha; APOBEC3G - apolipoprotein B messenger RNA editing catalytic subunit-like protein 3G; TRIM5 α - tripartite motif-containing 5 alpha; KIR – killer cell immunoglobulin-like receptor; HLA – human leukocyte antigen; ADCC – antibody-dependent cell-mediated cytotoxic activity; Gag – group specific antigen; Nef – negative regulation factor.

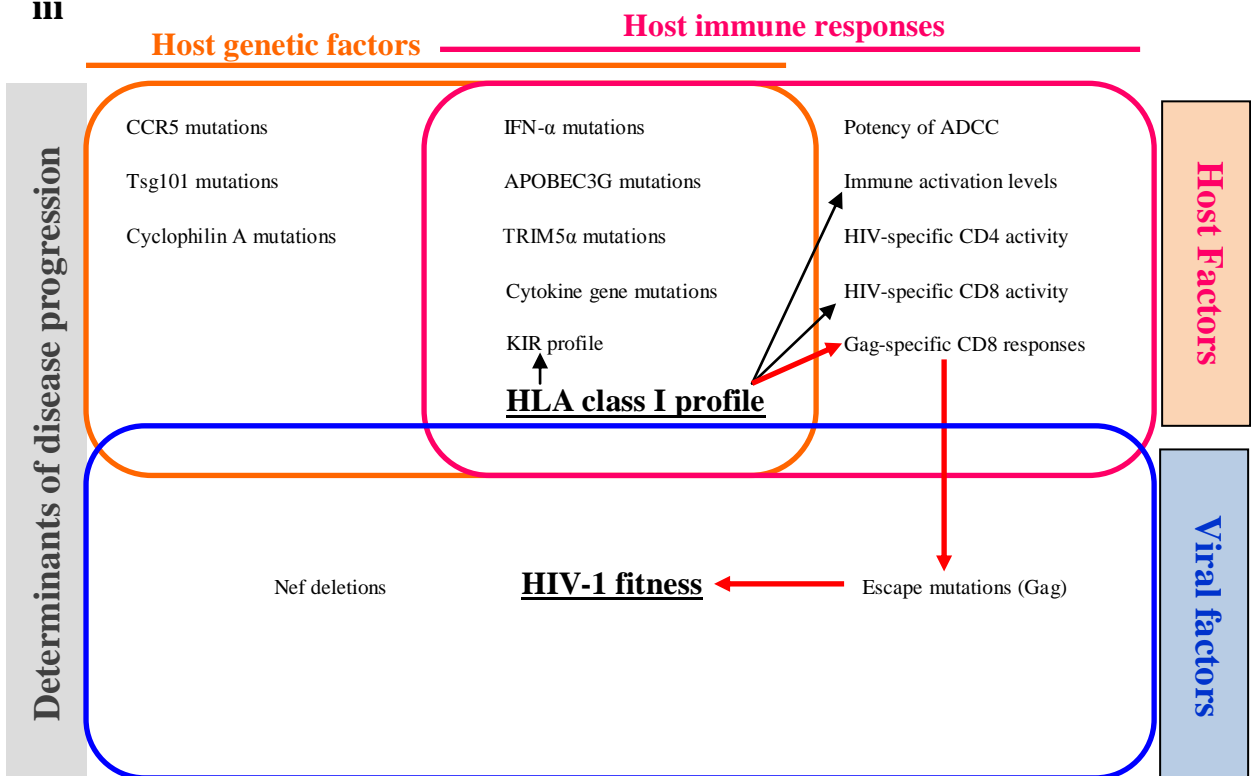
i Time to AIDS following HIV-1 infection (years)



ii



iii



fluorescent protein (GFP; which can be detected by flow cytometry) [219, 234, 235]. Other methods of detection include direct detection of viral gene products/activities or reporter genes engineered into cells [232] (Section 1.7.1.3). Fitness differences can be directly attributed to the mutation and results are therefore easily interpreted. However, a limitation of this approach is the consideration of the fitness consequences of a particular mutation outside of the context of the natural viral sequence background, where co-existing mutations are likely to occur [236]. Introduction of numerous secondary mutations through site-directed mutagenesis is possible, however this process is labour-intensive and still may not recapitulate natural conditions.

1.7.1.2 Recombinant viruses

In contrast to point mutations, an entire genomic region may be inserted into a standard viral backbone (with or without a reporter gene for detection), allowing for a direct link between measured fitness and the viral protein or region of interest [236].

The genomic region analysed may be a single clone or a pool of sequences amplified from clinical samples or other sources [236]. The advantage of using a single clone is that the precise sequence is known, while a population sequence reflects mutations present in the virus pool but not necessarily occurring in combination in one virus strain. However, analysis of a virus pool may be more representative of the diversity *in vivo*.

The construction of recombinant viruses may be performed using restriction enzymes, a yeast recombination system, or homologous recombination of the viral genomic region (amplified using primers complementary to the vector) and vector in mammalian cells [237,

238]. Due to the variability of HIV-1, there is often a lack of convenient, unique restriction sites [238, 239]. A yeast recombination system does not require unique restriction sites, but (with the exception of a very recently developed system [238]) sub-cloning is required to achieve full infectivity [238-240]. Homologous recombination is less labour-intensive, although it is time-consuming as eukaryotic recombination is of poor efficiency [238]. Another alternative is gene complementation involving transfection of cells with a gene expression vector and a full length vector with the relevant gene deleted to produce pseudovirions [233, 241]. However, this introduces foreign genetic elements [238] and since pseudovirions may only complete one replication cycle, this may only be applied in single cycle fitness assays (Section 1.7.3.1).

Although viral fitness effects measured by this approach can be attributed to a genomic region, it may not be clear which mutations are primarily responsible, necessitating further investigations [236]. Another limitation is that since the genomic region is not in the natural context, interactions with other genes are not taken into account [232, 236]. For example, Gag interactions with gp41 in Env (Section 1.3.7) may be disrupted by examining either protein in isolation. A measurement of whole viral isolates is therefore likely to yield the most reliable fitness results [232, 236].

1.7.1.3 Whole isolates

HIV-1 may be isolated from peripheral blood mononuclear cells (PBMCs) or plasma, with similar or higher rates of success from PBMCs and greater efficiency from PBMCs [242, 243]. There are also methods to extract clones from viral isolates [244] to allow fitness of specific clones to be measured. Despite accuracy of measurement, the genetic determinant

for the fitness effect is not known and isolation of virus requires extra costs, is time-consuming and difficult for some strains [236]. If the fitness of whole isolates is measured in primary cells (*e.g.* PBMCs), measurement of fitness is limited to the direct detection of viral products (*e.g.* p24 antigen by enzyme-linked immunosorbent assay [ELISA]) or activity (*e.g.* reverse transcriptase activity) [232], neither of which assess particle infectivity [71]. However, if isolates are grown in a cell line, the cell line may be manipulated to express a reporter gene, such as GFP, on infection, thereby allowing an easy, convenient, and relatively inexpensive measure of viral spread in the culture [71, 235].

1.7.2 Primary cells versus T cell lines

Viral fitness measured in primary cells (cells cultured directly from human subjects) should yield results representative of the *in vivo* situation. PBMCs are the primary cells most commonly used in HIV-1 fitness assays to investigate pathogenic fitness; however, other cell types may be used for specific research questions. For example, Langerhans cells may be used to directly assess HIV-1 transmission fitness [222]. It is important to note that fitness measurements obtained in PBMCs may differ from those obtained using established T cell lines, and may also vary between different T cell lines [232]. For example, reduced viral fitness due to the M184V NRTI resistance mutation was evident in PBMCs but not in T cell lines, due to lower nucleotide concentrations present in PBMCs [245]. Although the fitness defect of T242N was detected in a CEM-GXR cell line [246], this fitness cost was significantly more pronounced in a Jurkat cell line and in PBMCs, possibly because this mutation disrupts capsid-CypA interaction (Section 1.6.5.1) and CEM-GXR cells display higher CypA levels [33]. However, no significant influence of cell type was detected in studies investigating the fitness of resistance mutants to NNRTI [232] and another study showed similar results for fitness effects of escape and compensatory mutations in PBMCs

and cell lines [213]. Although PBMCs are the cells infected *in vivo*, they are highly variable between donors, require stimulation before use and cannot be maintained for long periods *in vitro* [71, 236]. Cell lines overcome these difficulties.

1.7.3 Single and multiple cycle fitness assays

1.7.3.1 Single cycle assays

Single cycle assays are commonly used to measure infectivity and generally involve infection of cell lines with pseudovirions or recombinant viruses encoding reporter genes, followed by reporter gene detection in cells (luminescence/fluorescence) between 24-72 hours post-infection [211, 219, 241]. Single cycle assays have the advantage of yielding results quickly. However, they have the disadvantage of not measuring the full replication cycle and are often less sensitive [71]. Multiple cycle assays can amplify fitness differences of variants over many cycles, thus increasing sensitivity, and are able to measure spread of infection in a culture [71].

1.7.3.2 Multiple cycle assays

1.7.3.2.1 Pairwise growth competition assays

In pairwise growth competitions, two viral variants compete for the same resources under the same culture conditions, making this an accurate and sensitive measure of viral fitness [236]. The fitter variant outcompetes the other and the relative proportion of the variants at different time points can be measured by a heteroduplex tracking assay (HTA), population sequencing, or real-time polymerase chain reaction (PCR) [232, 236]. HTA requires that the variants are sufficiently different to design unique probes allowing resolution of the two

probe-PCR product heteroduplexes on a gel [218, 247]. Quantitation of relative variant proportions by population sequencing has a limited linear range, however clonal sequence analysis is labour-intensive [232]. Real-time PCR is higher throughput, however this assay has to be optimised for each primer/probe used to appropriately differentiate the variants [232]. Therefore, the methods of detection are generally complex, technically demanding, labour-intensive, costly, require prior sequence knowledge, and results are difficult to quantify [71, 232]. These difficulties may be largely overcome if recombinant viruses instead of whole viral isolates are competed, by engineering different fluorescent genes (or reporter genes detected by fluorescent antibodies) into the virus backbones used, thereby allowing relative quantification by flow cytometry [234, 248]. However, another theoretical potential problem may be recombination of the variants compared since they are in a single culture, although this risk can be minimised by low inoculation dose and a limited culture period [232].

1.7.3.2.2 Parallel assays

Parallel mono-culture replication assays are simpler and less labour-intensive than growth competitions [71, 232]. Cell populations in parallel cultures may grow at slightly different rates and could therefore influence calculated viral fitness, which may be important when fitness differences are very small [236]. They are therefore less sensitive than growth competitions, although more sensitive than single cycle assays. Viral gene products or activity may be directly detected or reporter genes in the viral backbone may be detected. Alternatively cell reporter genes can be used to measure spread of infection.

1.7.4 Choice of assay for the present study

For the present study, a flow cytometry-based, multiple cycle, parallel mono-culture infection assay using a GFP-reporter CEM-GXR T cell line [71] was chosen to investigate fitness consequences of sequence variability in the Gag-protease region, specific Gag mutations, as well as the fitness of whole HIV-1 isolates. This assay was chosen because of its relative sensitivity (due to measuring multiple cycles), simple and inexpensive detection techniques (that may be applied to recombinant viruses or whole isolates), and consistency between experiments (due to its use of a homogenous population of cells) (Sections 1.7.2 and 1.7.3.2.2). The measure of viral spread overcomes the potential problem of detecting defective virus particles by p24 ELISA for example (Section 1.7.1.3).

Homologous recombination in the CEM-GXR cell line was chosen to construct the Gag-protease recombinant viruses [225] since the protocol is simple and not labour-intensive (Section 1.7.1.2). The patient-derived virus pool was measured as opposed to clones, since this is likely to be more representative of *in vivo* conditions and is less labour-intensive than cloning for the large number of individuals tested (Section 1.7.1.2).

1.8 The present study: rationale, aims and objectives

A better understanding of the determinants that affect the rate of HIV-1 disease progression and why some individuals control HIV-1 to low viral loads is important for designing an effective HIV-1 vaccine, for which the need is urgent (Section 1.5.2). Variation in HLA class I alleles is the strongest correlate of progression to AIDS, however, the mechanisms underlying this association are incompletely understood (Section 1.6.4), and may differ between HLA alleles (Section 1.6.5.1). One hypothesised mechanism is that certain HLA

alleles select for mutations in conserved, structurally/functionally important Gag epitopes through CD8⁺ T cell mediated pressure, resulting in reduced viral fitness, lower viral loads and slower disease progression in these individuals (Section 1.6.4.2.3.2.3). Further, individuals without protective HLA alleles who control viral load tend to make CD8⁺ T cell responses to conserved regions of HIV-1, such as Gag, or may be recipients of a virus with reduced Gag-protease function, suggesting that viral attenuation might possibly also explain viral control in individuals without protective HLA alleles.

These hypotheses require further investigation, since many previous studies investigating fitness cost of Gag escape mutations and their clinical relevance have concentrated only on a few pre-selected escape mutations associated with protective HLA alleles, have not assessed fitness consequences in the natural sequence background (in the presence of other escape and compensatory mutations), and/or have focused on a limited number of patients. The majority of studies have also focused on HIV-1 subtype B, rather than HIV-1 subtype C which contributes most to the global epidemic. Furthermore, the clinical relevance or impact of Gag function on HIV-1 disease progression requires further clarification (Section 1.6.5). It is also not known which Gag mutations carry the greatest fitness cost, whether certain mutations (or combinations thereof) selected by common/non-protective HLA alleles might impact on viral fitness and HIV-1 disease progression, whether Gag function contributes substantially to overall HIV-1 fitness (given its important role in replication [Section 1.3]), and whether Gag function might explain differences in fitness (and pathogenicity) between HIV-1 subtypes, all of which have important implications for design of a vaccine aiming to drive HIV-1 to a less fit state. Such a vaccine will also require knowledge of compensatory pathways developed by HIV-1, which may be revealed by fitness studies using clinically-derived sequences.

Thus, the main aim of the present study was to examine the hypothesis that HLA-driven variation in HIV-1 Gag significantly impacts on viral fitness and disease progression. This was assessed through a comprehensive investigation using clinically-derived Gag-protease sequences collected from a large population of chronically HIV-1 subtype C infected individuals. Protease was included in order to maintain the natural sequence interaction between Gag and protease, *viz.* site-specific cleavage of the Gag polyprotein by protease enzymatic activity (Section 1.3.9), the importance of which is indicated by the co-evolution of these proteins [249, 250]. The specific aims of the present project were to:

- Assess to what extent HLA alleles influence HIV-1 fitness through their ability to drive sequence variability in Gag-protease.
- Analyse the impact of Gag-protease-mediated fitness on HIV-1 disease progression.
- Identify Gag-protease mutations that significantly impact HIV-1 fitness and mutations that may compensate for viral attenuation.
- Determine whether Gag-protease contributes significantly to overall HIV-1 fitness.
- Determine whether fitness differences between HIV-1 subtype B and C may be partly explained by Gag-protease function.

To achieve these aims the following specific objectives were fulfilled:

- Construction of recombinant viruses encoding Gag-protease sequences derived from HIV-1 subtype C chronically infected patients (n=406) and recently infected patients (n=60), as well as a small subset of HIV-1 subtype B chronically infected patients (n=25).
- Analysis of *in vitro* replication capacity for Gag-protease recombinant viruses using a flow cytometry-based GFP-reporter T cell assay.

- Relation of Gag-protease replication capacity values to HLA class I profiles in the large population of HIV-1 subtype C chronically infected patients.
- Relation of Gag-protease replication capacity values to cross-sectional viral loads, CD4+ T cell counts, and subsequent rate of CD4+ T cell decline in HIV-1 subtype C chronically infected patients, and to viral set point and rate of CD4+ T cell decline in recently infected patients.
- Sequencing of Gag-protease, identification of HLA-associated mutations and relation of mutations to Gag-protease replication capacity values.
- Isolation of HIV-1 from infected PBMCs followed by replication capacity measurement and comparison to replication capacity values of recombinant viruses encoding Gag-protease from the same isolates.
- Comparison of *in vitro* replication capacities of HIV-1 subtype B and C Gag-protease recombinant viruses and identification of subtype-specific mutations linked to altered replication capacities.
- Confirmation of suspected fitness cost and/or compensation of HLA-associated or subtype-specific mutations by site-directed mutagenesis followed by assaying the replication capacities of mutant viruses.

CHAPTER 2

**MATERIALS AND
METHODS**

CHAPTER2 – MATERIALS AND METHODS

2.1 Ethics

The present study forms a component of research protocols approved by the Biomedical Research Ethics Committee of the University of KwaZulu-Natal (E028/99 and E036/06) as well as collaborative projects approved independently by the Institutional Review Boards of Botswana (PPME-13/18/1 Vol VI [17]) and the Harvard School of Public Health (USA) (P10491-136). Study materials from previously recruited patients comprised stored plasma and PBMCs. Written informed consent was obtained from these individuals at the time of recruitment and patient identity remained confidential. In addition, reverse transcriptase polymerase chain reaction (RT-PCR) products containing HIV-1 subtype B *gag-protease* sequences were obtained from collaborators at the University of British Columbia. Collection of these anonymised patient-derived materials was approved by the Research Ethics Board of the University of British Columbia/Providence Healthcare (Canada) (H07-03006).

2.2 Measurement of replication capacities of viruses encoding Gag-protease isolated from chronically and recently HIV-1 infected subjects

2.2.1 Study subjects

2.2.1.1 HIV-1 subtype C chronically infected subjects

Study subjects chronically infected with HIV-1 subtype C comprised 406 antiretroviral naïve individuals from the Sinikithemba (SK) cohort in Durban, South Africa. Plasma samples at study entry (baseline) that were stored at -80 °C were the study material used for experiments to determine Gag-protease-mediated replication capacities. In addition, viral loads, CD4+ T cell counts, and HLA class I profiles of these subjects were available [170].

Viral load measurements were previously obtained by the Roche Amplicor version 1.5 assay at baseline and approximately every 6 months thereafter. CD4+ T cell count measurements were obtained by flow cytometry at study baseline and at intervals of approximately 3 months thereafter. Baseline measurements were available for all participants and antiretroviral naïve follow-up was available for 339 individuals. For each of these 339 subjects, a rate of CD4+ T cell decline over the treatment-free follow up period was calculated by linear regression in collaboration with Bingxia Wang (Program in HIV Outcomes Research, Massachusetts General Hospital, Massachusetts). HLA types to 4-digit resolution were previously determined using molecular methods. Cohort characteristics, as well as demographic data, are presented in Table 2.1.

2.2.1.2 HIV-1 subtype B chronically infected subjects

For comparison of Gag-protease-mediated replication capacities with HIV-1 subtype C samples, *gag-protease* RT-PCR products from 25 HIV-1 subtype B infected individuals from the British Columbia HOMER cohort were donated by Dr. Zabrina Brumme (Simon Fraser University, Canada; formerly of the Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology and Harvard University). The Gag-protease-mediated replication capacities for 803 individuals from the HOMER cohort were previously measured [251]. The 25 samples donated were chosen as they covered the range of replication capacities and were a representative subset of the cohort in terms of Gag-protease-mediated replication capacity. Viral load and CD4+ T cell data on these 25 subjects was also provided. Twenty-five HIV-1 subtype C chronically infected patients matched for CD4+ T cell counts and viral loads with the HIV-1 subtype B chronically infected patients were randomly selected from the SK cohort (Table 2.2).

Table 2.1 Clinical and demographic characteristics of HIV-1 subtype C chronically infected patients

Characteristic	Median (interquartile range) or percentage (%)
Gender	79% Female, 21% Male
Age (at baseline)	31 (27-36) years
Viral load (at baseline)	4.77 (4.15-5.27) log ₁₀ copies/ml
CD4+ T cell count (at baseline)	340 (238-477) cells/mm ³
Rate of CD4+ T cell decline	-30 (-73 to -3) cells/mm ³ per year
Follow-up time	2.29 (1.21-3.02) years

HIV-1 – human immunodeficiency virus type 1; CD4 – cluster of differentiation 4.

Table 2.2 Clinical characteristics of matched HIV-1 subtype B and C chronically infected patients

Characteristic	Median (interquartile range)	
	Subtype B (n=25)	Subtype C (n=25)
Viral load	5.38 (5.17-5.61) log ₁₀ copies/ml	5.3 (4.9-5.7) log ₁₀ copies/ml
CD4+ T cell count	140 (60-320) cells/mm ³	137 (76-341) cells/mm ³

HIV-1 – human immunodeficiency virus type 1; CD4 – cluster of differentiation 4.

2.2.1.3 HIV-1 subtype C recently infected subjects

Study subjects recently infected with HIV-1 subtype C comprised 32 individuals from the HIV Pathogenesis Programme Acute Infection Study in Durban, South Africa, and 28 individuals from the Tshedimoso Study in Botswana (plasma samples donated by Vladimir Novitsky from the Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, MA).

At screening, 27 subjects from Durban had detectable HIV RNA but had not yet seroconverted (negative ELISA and negative Western blot according to Centre for Disease Control criteria) and were defined as acutely infected [252]. The date of infection was estimated to be 14 days prior to screening as previously described [253]. A further 5 subjects from Durban had an incomplete Western blot at the time of screening that later became complete and these were defined as recently infected. The date of infection was estimated for 2 of these subjects by Fiebig staging but was indeterminable for the other 3 who were screened in Fiebig stage V-VI or VI [81]. For acutely infected subjects from Botswana, days post-seroconversion were estimated as the midpoint between the last seronegative and the first seropositive test, while for recently infected subjects the calculation was based on Fiebig staging [254]. For comparison with the Durban subjects, the days post-infection were estimated for the Botswana subjects by adding 22 days to the estimated days post-seroconversion since this is the estimated window period [81]. Plasma samples from the earliest time point available for use following screening of subjects (5 at Fiebig stage IV, 45 at Fiebig stage V, 7 at Fiebig stage VI, and 3 indeterminable; median of 55 days post-infection, interquartile range [IQR]: 43-76 days) were the material used for analysis of Gag-protease-mediated replication capacities. In addition, Gag-protease-

mediated replication capacities were analysed at a median of 1 year [IQR, 0.92 to 1.13 years] later for 13 of the Durban subjects for whom plasma samples were available.

Viral loads, CD4+ T cell counts, and HLA profiles were previously determined for all subjects, and CD8+ T cell responses to Gag optimal epitopes were also available for 13 of the Durban subjects for whom replication capacity was measured longitudinally [252]. For Durban subjects, viral loads were obtained at screening and at approximate intervals of 2 weeks, 2 weeks, 2 weeks, 1 month, 1 month, and thereafter every 4-5 months by the Roche Amplicor version 1.5 assay or the Cobas Taqman HIV-1 Test. For Botswana subjects, viral loads were obtained approximately every 100 days using the Cobas Ampli-Prep/Cobas Amplicor HIV-1 Monitor Test version 1.5 [255]. CD4+ T cell count measurements were performed at the same intervals by flow cytometry for both Durban and Botswana participants [255]. Viral set points were calculated by averaging the viral loads between 3 and 12 months post-infection, since peak viremia occurs in Fiebig stage III and then there is a decline towards a steady state which is established during Fiebig stages V or VI (approximately 3 months post-infection) [81]. Rates of CD4+ T cell decline were calculated as described above (Section 2.2.1.1). HLA class I typing was previously performed at least to 2-digit resolution and to 4-digit resolution for the majority of samples by molecular methods. CD8+ T cell responses to Gag optimal epitopes were previously measured [252], at the earliest available time points and a median of 6.1 weeks (IQR: 5.9-7 weeks) later, by interferon gamma (IFN- γ) enzyme-linked immunospot assay (ELISPOT) using fresh or frozen PBMC samples. A response was defined as positive if greater than 100 spot forming cells (SFCs)/10⁶ PBMCs after subtraction of the negative control value plus 3 standard deviations. Demographics and characteristics of recently infected patients are shown in Table 2.3.

Table 2.3 Clinical and demographic characteristics of HIV-1 subtype C recently infected patients

Characteristic	Median (interquartile range) or percentage (%)
Gender	65% Female, 35% Male
Age (at sampling)	27 (25-34) years
Fiebig stage (at sampling)	8.3% stage IV, 75% stage V, 11.7% stage VI, 5% unknown
Viral load (at sampling)	5.06 (4.33-5.57) log ₁₀ copies/ml
CD4+ T cell count (at sampling)	418 (302-524) cells/mm ³
Viral set point	4.69 (4.11-5.15) log ₁₀ copies/ml
Rate of CD4+ T cell decline	-7 (-13-1) cells/mm ³ per month
Follow-up time	365 (184-457) days

HIV-1 – human immunodeficiency virus type 1; CD4 – cluster of differentiation 4.

2.2.2 Generation of Gag-protease NL4-3 recombinant virus stocks

Gag-protease NL4-3 recombinant viruses were generated by co-transfection of a CEM-derived T cell line with patient-derived *gag-protease* amplicons and *gag-protease*-deleted NL4-3 plasmid (pNL4-3 Δ *gag-protease*) based on methods previously described [225].

2.2.2.1 Amplification and sequencing of patient-derived *gag-protease*

Patient plasma samples were thawed at room temperature. HIV-1 RNA was extracted from 140 μ l plasma using the QIAamp Viral RNA Mini Kit (Qiagen, Valencia, USA) as per manufacturer's instructions. If the plasma viral load was less than 5,000 copies/ml, virus was concentrated by centrifugation (Jouan MR23i, Thermo Scientific, USA) at 14,000 rpm for 2 hours at 4 °C prior to RNA extraction. RT-PCR was performed using the Superscript III One-Step RT-PCR kit (Invitrogen, Carlsbad, USA) and the following *gag-protease* specific primers: 5' CAC TGC TTA AGC CTC AAT AAA GCT TGC C3' (HXB2 nucleotides 512-539) and 5' TTT AAC CCT GCT GGG TGT GGT ATY CCT 3' (2851-2825). The RT-PCR reaction comprised 14.4 μ l diethyl pyrocarbonate (DEPC) treated water (Invitrogen), 20 μ l 2X buffer, 0.8 μ l of each primer (10 μ M), 0.8 μ l RT/Taq enzyme mix, and 4 μ l RNA extract. The reaction mixture was incubated (GeneAmp PCR system 9700, Applied Biosystems, Foster City, USA) at 30 °C for 55 minutes and 94 °C for 2 minutes, followed by 35 cycles of 94 °C for 15 seconds, 55 °C for 30 seconds, and 68 °C for 2 minutes, and ended with a 5-minute incubation at 68 °C.

A second round of PCR was performed with 100-mer forward (5' GAC TCG GCT TGC TGA AGC GCG CAC GGC AAG AGG CGA GGG GCG GCG ACT GGT GAG TAC GCC AAA AAT TTT GAC TAG CGG AGG CTA GAA GGA GAG AGA TGG G 3') and

reverse (5' GGC CCA ATT TTT GAA ATT TTT CCT TCC TTT TCC ATT TCT GTA CAA ATT TCT ACT AAT GCT TTT ATT TTT TCT TCT GTC AAT GGC CAT TGT TTA ACT TTT G 3') primers that were exactly complementary to NL4-3 on either side of *gag-protease* using the TaKaRa Ex Taq HS enzyme kit (Takara, Shiga, Japan). Two 50 μ l PCR reactions were prepared per sample, comprising 37 μ l DEPC water, 5 μ l 10X Ex Taq buffer, 4 μ l deoxyribonucleotide triphosphate (dNTP), 0.8 μ l forward primer (10 μ M), 0.8 μ l reverse primer (10 μ M), and 0.25 μ l Ex Taq to 2 μ l RT-PCR product. Thermocycler conditions were as follows: 94 °C for 2 minutes, 40 cycles of 94 °C for 30 seconds, 60 °C for 30 seconds, and 72 °C for 2 minutes, followed by 72 °C for 7 minutes. PCR products from two 50 μ l reactions were pooled and stored at -20 °C. Approximately 80-90 μ l of PCR product was used in the generation of recombinant viruses. *Gag-protease* amplification was confirmed by running 5 μ l of the product on a 1% polyacrylamide gel, resulting in a band of \approx 1.7 kb. The remainder was diluted 1:15 in DEPC water and set aside for sequencing.

Gag-protease PCR product was population or bulk sequenced using the Big Dye ready reaction termination mix V3 (Applied Biosystems) and the following sequencing primers: 5' CTT GTC TAG GGC TTC CTT GGT 3'(1098-1078), 5' CTT CAG ACA GGA ACA GAG GA 3'(991-1010), 5' GGT TCT CTC ATC TGG CCT GG 3'(1481-1462), 5' CAA CAA GGT TTC TGT CAT CC 3'(1755-1736), 5' CCT TGC CAC AGT TGA AAC ATT T 3'(1981-1960), 5' TAG AAG AAA TGA TGA CAG 3'(1817-1834), 5' CAG CCA AGC TGA GTC AA 3'(2536-2520) and 5' GGA GCA GAT GAT ACA GTA TT 3' (2331-2350). Each sequencing reaction per primer prepared in a 96-well plate (Applied Biosystems) comprised 0.4 μ l big dye mix, 2.6 μ l sequencing primer, 2 μ l sequencing buffer, 4 μ l DEPC water, and 1 μ l diluted template. The reaction was incubated at 96 °C for 1 minute, followed by 25 cycles of 96 °C for 10 seconds, 50 °C for 5 seconds, and 60 °C for 4 minutes. A

sequencing plate clean up was performed the same day. Sequencing products were first diluted in ethylenediaminetetraacetic acid (EDTA; 125 mM, pH 8.0, 1 μ l) and a mixture of sodium acetate (3 M, pH 5.2, 1 μ l) and cold ethanol (100%, 25 μ l), followed by centrifugation (Eppendorf centrifuge 5810R, Merck, Germany) at 3,000 \times g for 20 minutes. To dry products, the plate was inverted on paper towel and centrifuged at 150 \times g for 1 minute. Pellets were immediately resuspended in cold ethanol (70%, 35 μ l) and then centrifuged at 3,000 \times g for 5 minutes. Following inversion and centrifugation at 150 \times g for 1 minute, products were dried in a thermocycler at 50 $^{\circ}$ C for 5 minutes and then stored at -20 $^{\circ}$ C until analysis. Immediately prior to analysis of sequences on the ABI 3130xl Genetic Analyzer (Applied Biosystems), products were resuspended in 10 μ l formamide, vortexed and denatured in a thermocycler at 95 $^{\circ}$ C for 3 minutes and then cooled at 4 $^{\circ}$ C for 3 minutes. Sequences were visualized and edited in Sequencher 4.8 (Gene Codes, Corp., Ann Arbor, USA). Sequence data were aligned to HIV-1 subtype B reference strain HXB2 (Genbank accession number K03455) using a modified NAP algorithm [256] and insertions with respect to HXB2 were stripped out. HIV-1 subtype was confirmed using the REGA subtyping tool [257].

These sequences can be located in Genbank under the following accession numbers: HM593106-HM593510 and HQ696791- HQ696863.

2.2.2.2 Preparation of the pNL4-3 Δ *gag-protease* backbone

The subtype B pNL4-3 Δ *gag-protease* plasmid [225], containing a BstE II restriction enzyme site in place of the viral *gag-protease* coding region, was supplied in *E. coli* STBL3 cells by Toshiyuki Miura (University of Tokyo, Japan; formerly of the Ragon Institute of

Massachusetts General Hospital, Massachusetts Institute of Technology and Harvard University). A subtype C pMJ4 Δ *gag-protease* vector was initially tested, but due to its low infectivity [258] and poor recovery of recombinant stocks, the available subtype B backbone was selected for these experiments. To amplify pNL4-3 Δ *gag-protease* stocks, Luria-Bertani (LB) broth (Sigma, St Louis, USA) containing 100 μ g/ml ampicillin was inoculated with 17.5 μ l STBL3 stock containing pNL4-3 Δ *gag-protease* per 100 ml broth and incubated overnight at 37 °C in a shaking incubator (Infors HT, Bottmingen, Switzerland). The plasmid was then purified using the Plasmid Maxi kit (Qiagen), quantified using a nanodrop spectrophotometer (Thermo Scientific, Delaware, USA) and stored in a -80 °C ultrafreezer (Snijders Scientific, Holland) until use.

2.2.2.3 Thawing, culturing, and storing of CEM-GXR25 cells

A GFP-reporter T cell line, CEM-GXR25 (GXR) [71], was donated by Dr. Mark Brockman (Simon Fraser University, Canada; formerly of the Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology and Harvard University). GXR cells express the HIV-1 CD4 receptor as well as the co-receptors CXCR4 and CCR5, and encode a Tat-inducible HIV-1 LTR-GFP expression cassette [71]. Thus, GXR cells are permissible to HIV-1 replication and cells produce GFP upon infection, thereby allowing their detection using flow cytometry.

A frozen 1 ml aliquot of GXR cells was transferred immediately from the liquid nitrogen freezer (Custom Biogenics Systems, Romeo, USA) into a 37 °C waterbath and gently agitated until the last ice crystal melted. Once thawed, cells were transferred to a T25 flask (Corning-Costar, New York, USA) with 4 ml pre-warmed R10 medium and thereafter

incubated in a humidified Heraeus incubator (Thermo Scientific) at 37 °C and 5% CO₂. R10 medium consisted of RPMI-1640 (Sigma) supplemented with 10% foetal bovine serum (FBS; Gibco, New York, USA), 2 mM L-glutamine (Sigma), 10 mM N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid (HEPES; Gibco), and 50U/ml penicillin-streptomycin (Gibco). The following day, the GXR cells growing in suspension were pelleted at 1,500 rpm for 10 minutes (Heraeus multifuge 3SR+, Thermo Scientific) and then resuspended in 10 ml fresh R10 medium to remove dimethylsulfoxide (DMSO; Sigma) present in the storing medium. When the phenol red indicator in the medium turned yellow, indicating cell growth and a need for medium replacement, GXR cells were transferred to a T75 flask (Corning) and a stock culture was maintained in 30-40 ml R10 medium. Every 2-3 days, approximately 75-90% of the culture was discarded (if not used for experiments) and replaced with fresh R10 medium. The GXR stock culture was discarded after 2-3 months and a new vial of GXR cells thawed.

Prior to initiation of experiments several vials of GXR cells were stored in liquid nitrogen for future use. This required determining the cell concentration of the GXR stock culture. A Neubauer haemocytometer (Assistant, Germany) was used for this purpose. A haemocytometer is a glass slide with a grid of 2 sets of 4 large squares each of volume 0.1 mm³ (10⁻⁴ ml) and each subdivided into 16 smaller squares for ease of counting. After thorough mixing of the culture, 10 µl cell suspension was pipetted under the coverslip placed over the grid area of the slide. The cells were then counted in 2 large squares of each set of 4 squares with a phase contrast, bright field microscope (Leica, Germany) and an average taken. The average number of cells in the large squares was multiplied by 10⁴ to determine the average number of cells per ml. The volume of culture to yield the required number of

cells for storage was removed and pelleted at 1,500 rpm for 10 minutes. This volume was calculated as follows:

volume of cell culture required (ml) = number of cells required/cell concentration (cells/ml)

Cells were resuspended in R10 medium at a concentration of 5.5 million cells/ml and 900 μ l of this suspension was pipetted into each cryovial (Greiner Bio-One, Germany). DMSO (100 μ l), a cryo-protectant [259, 260], was added drop-by-drop to each vial, with simultaneous gentle swirling, yielding a final concentration of 5 million cells/ml and a final volume of 1ml per vial. To allow sufficient time for DMSO penetration yet minimising toxicity, cells were allowed between 10 and 20 minutes exposure to DMSO [261] prior to transfer to Mr Frosty (Nalgene, Thermo Scientific) and placement in a -80 °C ultrafreezer (Snijders Scientific). Mr Frosty insulates the cells and achieves a cooling rate of approximately 1 °C per minute, which is optimal for most animal cells [259]. The following day, cells were stored in the liquid nitrogen freezer.

2.2.2.4 Co-transfection of GXR cells with *gag-protease* and pNL4-3 Δ *gag-protease*

To allow for production of recombinant viruses, GXR cells were co-transfected with PCR amplified patient *gag-protease* and the pNL4-3 Δ *gag-protease* backbone. Immediately prior to the co-transfection, pNL4-3 Δ *gag-protease* (10 μ g per sample) was digested with BstE II enzyme (Promega, Madison, USA) for 2 hours at 60 °C in a waterbath to linearise the plasmid at the site of *gag-protease* deletion and allow for insertion of amplified patient-derived *gag-protease* by recombination. Recombination of *gag-protease* with pNL4-3 Δ *gag-protease* was made possible by the use of long primers exactly complementary to NL4-3 on either side of *gag-protease* to generate *gag-protease* PCR product.

During the BstE II digestion of pNL4-3 Δ *gag-protease*, the GXR stock culture was counted and the required number of cells were removed, pelleted, and resuspended in fresh R10 medium as described in Section 2.2.2.3. Between 2 and 2.8 million GXR cells, depending on availability, in a volume of 800 μ l were electroporated (Gene Pulser II, Biorad, Hercules, USA) in 4 mm cuvettes (Biorad) at 300 V and 500 μ F with 10 μ g BstE II-digested pNL4-3 Δ *gag-protease* (\approx 10-15 μ l) and \approx 80-90 μ l *gag-protease* PCR product. Electroporation experiments were performed in batches of approximately 24 different samples. As a negative control, GXR cells were electroporated with pNL4-3 Δ *gag-protease* only. Electroporation allows for the temporary disruption of cell membranes allowing uptake of molecules and is a technique used to transfect cells with DNA [262]. Following a 1-hour incubation at room temperature to allow for cell recovery before pipetting, GXR cells were gently transferred to T25 flasks containing 4 ml pre-warmed R10 medium each and incubated at 37 °C and 5% CO₂. Five days later, 5 ml R10 medium was added to each flask followed by further incubation.

It should be noted that at an advanced stage of the present project a more efficient modified electroporation protocol was described [246]. Therefore, when the generation of 2 recombinant viruses encoding Gag-protease from recently infected patients (AS1-0703 and AS2-0802) from time points 1 year later than baseline failed using the above methods, the new modified protocol was employed. Briefly, 4 million cells, in a volume of 300 μ l, were electroporated with the linearised plasmid and PCR product at 250 V and 950 μ F. The cells were rested for 5 minutes before being transferred to a T25 flask with 10 ml medium and 1 million non-electroporated cells. Five days later, 5 ml R10 medium was added to each flask.

2.2.2.5 Monitoring of virus production by flow cytometry

Electroporated GXR cells were incubated for 12 days before virus production was monitored to allow time for recombination of *gag-protease* amplicons with the NL4-3 Δ *gag-protease* plasmid, followed by viral transcription, translation, particle production, and spread in the culture. The percentage infected cells was monitored by flow cytometry on a FACSCalibur (BD Biosciences, San Jose, USA) since GXR cells produce GFP when HIV-1 infected (Section 2.2.2.3). Flow cytometry is the measurement of cell characteristics such as size and granularity (through capture of light scatter) and fluorescence (through fluorescence capture) as single cells pass through a fluid stream [263].

In preparation for flow cytometry, 1 ml culture was removed from each flask following thorough mixing and pipetted into 96-well matrix cluster tubes (Corning Costar). The box containing these tubes was centrifuged at 1,100 rpm for 10 minutes. Following removal of supernatant, cells were resuspended in residual supernatant by vortexing. Cells were then mixed with 200 μ l of 2% paraformaldehyde (PFA) fixative and incubated for 10-15 minutes prior to flow cytometry. The fixative was prepared by dissolving PFA (Merck) in phosphate buffered saline (Gibco) at 55 °C for 30 minutes and thereafter adjusting the pH to 7.0 to complete dissolution. PFA fixative was stored at -20 °C and thawed aliquots stored at 4 °C were discarded after a week.

Matrix cluster tubes containing fixed cells were placed in re-usable FACS tubes (BD Biosciences) that were then placed in the FACSCalibur to allow entry of sample into the flow cell chamber. Flow cytometry instrument settings used were: forward scatter = E-1, side scatter = 440, FL1 = 280, FL2 = 493, and FL3 = 722. The negative control (uninfected

GXR cells) was used to distinguish (i) intact cells from debris, and (ii) GFP positive cells from GFP negative cells, as shown in Figure 2.1. These settings were saved as a template file that was minimally adjusted at the start of each flow cytometry run using the relevant experiment negative control. Fluorescence information was captured for 15,000 intact cells per sample. Results were saved and gates were finely adjusted using FlowJo [264] to accurately determine the percentage of intact cells in each culture that were infected with HIV-1.

2.2.2.6. Harvesting virus stocks

When approximately 25-30% of the GXR cells were infected, virus concentration in the culture supernatant was deemed to be sufficiently high to harvest virus stocks. The culture was centrifuged at 1,700 rpm for 5 minutes and the supernatant was stored in 10 aliquots of 1 ml each at -80 °C for use in subsequent titration and replication assays.

2.2.2.7 Validation of virus stocks

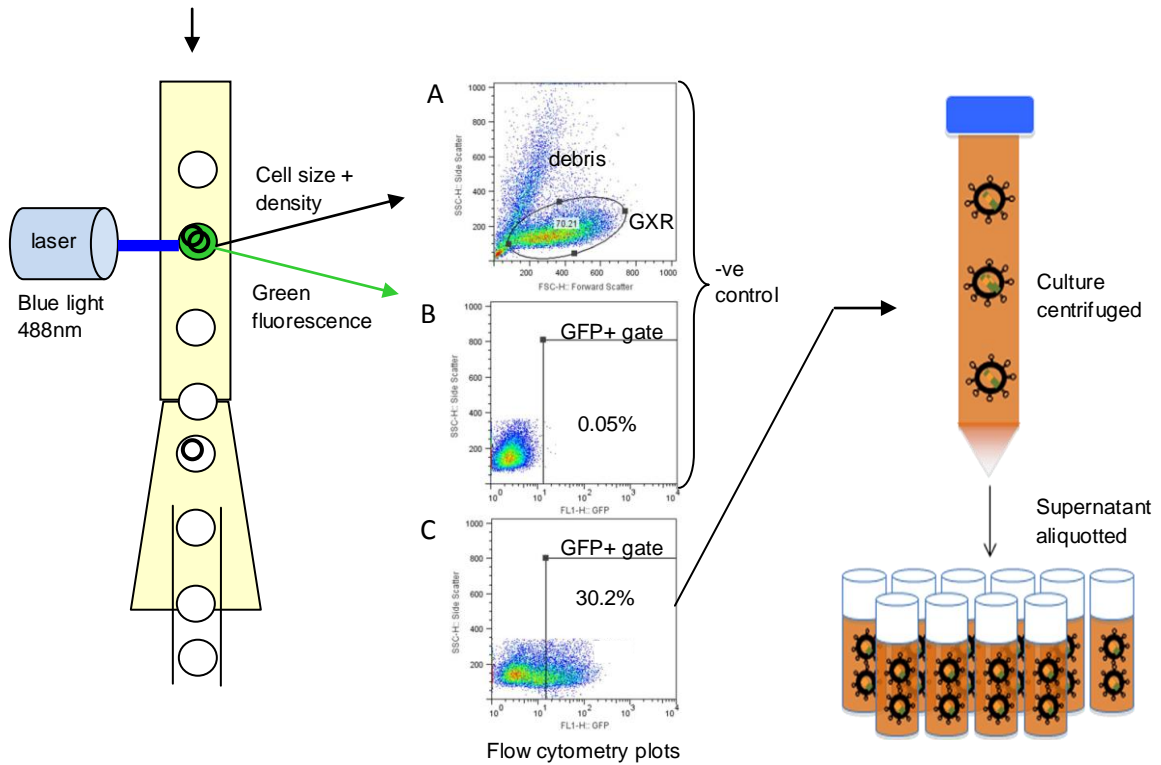
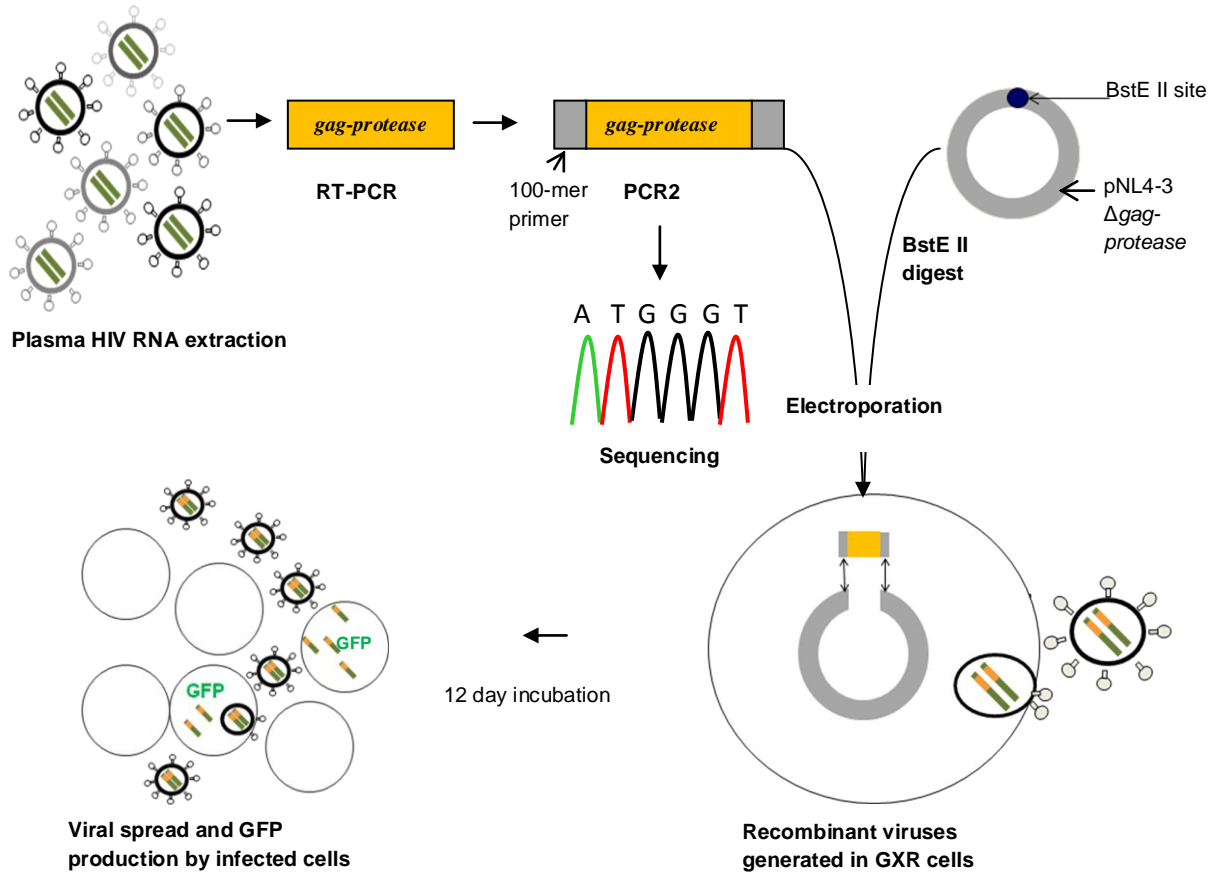
To test whether the Gag-protease contained within each recombinant virus pool was representative of the original plasma quasispecies, the *gag-protease* region was amplified and sequenced (Section 2.2.2.1) from 42 randomly-selected recombinant viruses and compared with the original plasma HIV RNA sequences.

The generation of Gag-protease NL4-3 recombinant viruses is summarised in Figure 2.1.

Figure 2.1 Generation of Gag-protease NL4-3 recombinant virus stocks

Diagram showing the incorporation of patient-derived HIV-1 *gag-protease* PCR product into an NL4-3 *gag-protease*-deleted viral backbone (pNL4-3 Δ *gag-protease*) to generate recombinant viruses. HIV-1 RNA was extracted from patient plasma followed by nested PCR amplification (RT-PCR and PCR2) of *gag-protease* and sequencing of the PCR product. In the second round of PCR, 100-mer primers exactly complementary to the pNL4-3 backbone were used to allow for recombination of PCR product and backbone. Prior to co-transfection of cells by electroporation with patient-derived *gag-protease* PCR product and pNL4-3 Δ *gag-protease*, the backbone was digested with BstE II enzyme to open the plasmid at the site of *gag-protease* deletion and allow for recombination. Cells were cultured for 12 days before monitoring viral spread in the culture. Viral spread was monitored every second day by flow cytometry since the GXR cells express green fluorescent protein when infected with HIV-1. The scattering of light, as the cells pass through the capillary of the flow cytometer, allowed distinction of cell populations (GXR cells) by cell size and density (A) and the emission of fluorescence allowed calculation of percentage HIV-1 infected GXR cells (B and C). Samples (C) were gated according to the negative control (A and B). At approximately 30 percent infection, the virus-containing supernatant was collected from the infected culture, aliquotted and stored at -80 °C until use in titration and replication assays.

Gag – group specific antigen; HIV-1 – human immunodeficiency virus type 1; RNA – ribonucleic acid; RT-PCR – reverse transcriptase polymerase chain reaction; PCR – polymerase chain reaction; pNL4-3 Δ *gag-protease* – *gag-protease*-deleted NL4-3 plasmid backbone; GFP – green fluorescent protein.



% infected cells monitored every 2 days by flow cytometry

2.2.3 Titration and replication capacity measurement of virus stocks

Titration of virus stocks and measurement of the replication capacities of these viruses was based on methods previously described [71, 225].

2.2.3.1 Titration

The aim of the replication capacity assay was to measure the exponential spread of each virus in GXR cell cultures over a short period of time. Therefore, the initial percentage of infected cells was required to be sufficiently low to observe subsequent exponential increase in infected cells, yet sufficiently high for this to occur over the course of a few days. Furthermore, the initial level of infection for the different recombinant viruses had to be within a similar range for the replication capacity measurements to be comparable. In preliminary tests, it was determined that an initial level of infection, or multiplicity of infection (MOI), of 0.3% was suitable. The infectious titre of each virus stock determines the volume required to obtain the necessary MOI. This can be precisely measured by infecting GXR cells with a known volume of virus stock in the presence of the HIV-1 protease inhibitor Indinavir, which allows infection of cells but prevents further spread of viral infection in the culture, and measuring the percentage of infected cells 2-3 days post-infection by flow cytometry [71]. However, since it was previously demonstrated that similar results could be obtained on day 2 post-infection in the absence of Indinavir [71], this simplified approach was used to determine the titre of each recombinant virus in the present study. Specifically, 0.4 ml virus stock was thawed (or 0.4 ml R10 medium for the negative control) and incubated with 1 million GXR cells in a final volume of 0.5 ml R10 medium in a 24-well culture plate (Corning Costar) for 24 hours at 37°C and 5% CO₂. This was followed by addition of 1 ml fresh R10 and incubation for a further 24 hours. Cultures were mixed by pipetting and 0.5 ml was then removed from each well and prepared for flow

cytometry as described in Section 2.2.2.5. The volume of virus stock to use to obtain a MOI of 0.3% in the replication assay was calculated as follows:

$$\text{volume of virus stock (ml)} = (0.3\%/\% \text{ cells infected}) * 0.4 \text{ ml}$$

2.2.3.2 Replication capacity assay

A previously unfrozen aliquot of virus stock was thawed and the calculated volume of virus stock (from the titration experiment, Section 2.2.3.1) was diluted to a final volume of 0.4 ml with R10 medium. GXR cells were counted, pelleted, and resuspended at a concentration of 10,000 cells/ μ l (Section 2.2.2.3). Next, 100 μ l of cell suspension (1 million GXR cells) was pipetted into each well of a 24-well plate and 0.4 ml virus stock was added. Following 24 hours of incubation at 37 °C with 5% CO₂, 1 ml R10 medium was added to each well. A negative (R10 medium) and positive (wild-type NL4-3 virus) control, generated by electroporation of GXR cells with 10 μ g NL4-3 plasmid (Section 2.4), was included in every assay. The percentage infected cells was measured by flow cytometry (Section 2.2.2.5) from days 2-6 post-infection by removal of 0.5 ml culture from each well (and replacement with fresh R10 medium).

Viral replication capacity was measured as the slope of exponential increase in percentage infected cells from days 3-6 post-infection, calculated using the semi-log method in Microsoft Excel, and expressed relative to the wild-type NL4-3 control, as follows:

$$\text{Replication capacity} = \text{Natural log (Logest[day 3 \% infection : day 6 \% infection])}$$

$$\text{Normalised replication capacity} = \text{replication capacity}_{\text{sample}} / \text{replication capacity}_{\text{NL4-3}}$$

The logest function in Excel calculates a \log_{10} exponential curve that best fits the data and returns the slope of the curve. Logest values were converted to natural log, which is more typically used for descriptions of exponential growth and decay in biology.

To ensure appropriate calculation of slope during only the exponential phase of viral spread, if the percentage infected cells exceeded a cut-off of 11% on day 5 post-infection, the slope was calculated from days 3-5 post-infection,. Replication assays were performed at least in duplicate independently and results were averaged.

The methods used to titrate and measure the replication capacity of virus stocks are depicted in Figure 2.2.

2.2.4 Variability in measuring Gag-protease-mediated replication capacity

To assess the total method variability (the sum of variability introduced at each step of the methods), RNA was extracted 5 times each from 7 different plasma samples and used to generate 5 independent Gag-protease recombinant stocks per individual. These viruses were then assayed to determine replication capacity.

2.2.5 Data Analysis

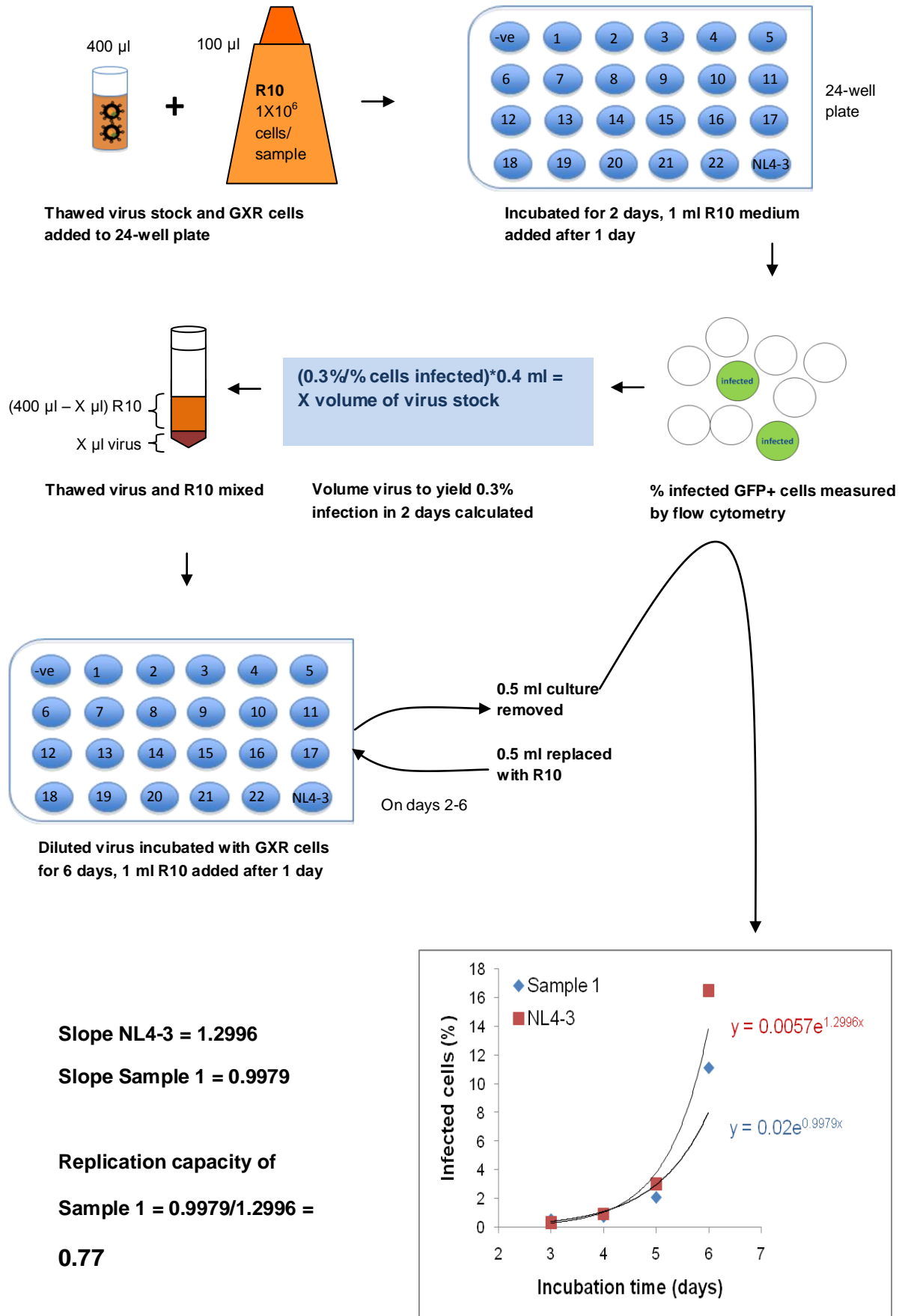
Method validation and variability

To test whether the Gag-proteases encoded by recombinant viruses were representative of the original plasma quasispecies, nucleotide differences between the plasma and

Figure 2.2 Titration and measurement of replication capacities of Gag-protease NL4-3 recombinant viruses

In this figure, the titration of virus stocks to determine the volume of viruses to yield a similar and appropriate starting percentage infection (0.3%) for the replication capacity assay is shown. This volume was determined by incubating a set volume of virus stock with GXR cells that express green fluorescent protein when infected in a 24-well plate and measuring the percentage infected cells by flow cytometry after 2 days. The replication capacities of recombinant viruses were determined by calculating the slope of exponential increase in percentage infected cells, as measured by flow cytometry, from days 3-6 following infection at a multiplicity of infection of 0.3%. These slope values were divided by the slope of growth of the wild-type NL4-3 positive control included in each assay in order to control for assay-to-assay variation.

Gag – group specific antigen; GFP – green fluorescent protein.



recombinant viral sequences were quantified using the Highlighter tool at Los Alamos [265]. In addition, to visually assess the similarity between plasma and recombinant virus sequences, a neighbour-joining tree was constructed from nucleotide sequences using Paup 4.0 and edited in Figtree [266].

To test whether variability in the percentage infection on day 3 (the first reading used in the calculation of slope of spread) of the replication assay influenced the replication capacity calculated, these 2 parameters were correlated by Pearson's correlation. The reproducibility of replication capacity measurements taken in duplicate was also tested by Pearson's correlation.

The replication capacities between different recombinant viruses generated 5 times each (*i.e.* each step of the methods was performed 5 times for each sample) were compared to assess whether differences were consistent. Analysis of variance (ANOVA) was used to determine whether there were overall significant differences between the replication capacities of viruses generated from the different samples and the Tukey post-hoc tests were used to identify which specific samples differed significantly from each other.

Replication capacities versus HLA class I alleles

Viral replication capacities were grouped according to the HLA class I alleles expressed by the host for chronically infected patients. ANOVA was used to assess whether significant differences in replication capacities were observed within expressed HLA-A, -B, and -C alleles. Then, for each individual allele ($n \geq 5$), the Student's T test (or the Mann-Whitney U-

test in cases where the assumptions of the Student's T test were not met) was used to compare replication capacities of viruses generated from persons expressing versus not expressing the allele in question. Recombinant viruses were also categorised according to the 10th and 90th percentiles of the replication capacity data into a low- and high-replication-capacity group, respectively. The proportion of protective alleles in each group was compared with Fisher's exact test. For recently infected patients, replication capacities were compared between individuals expressing protective and non-protective alleles by the Student's T test.

Replication capacities versus cross-sectional and longitudinal viral loads and CD4+ T cell counts

For chronically infected patients, the relationships between replication capacity and log viral load, CD4+ T cell count, and rate of CD4+ T cell decline were assessed using Pearson's correlation (for normally-distributed variables) or Spearman's rank correlation (for non-normally distributed variables). For recently infected patients, the relationships between replication capacity and viral set point as well as rate of CD4+ T cell decline were similarly assessed. In addition, replication capacities were also compared above and below the median viral set point and median rate of CD4+ T cell decline by the Student's T test or Mann-Whitney U test if assumptions of the Student's T test were violated.

Replication capacities versus Gag-protease sequences

To investigate the relationship between the number of polymorphisms in Gag or protease and replication capacity, the percent amino acid similarities of Gag and protease to the 2004 consensus subtype C sequence [267] were calculated using the sequence identity matrix

function in BioEdit 7.0 [268] and correlated (Pearson's or Spearman's correlation) with replication capacity.

Next, HLA-associated polymorphisms (amino acids that are significantly more likely to occur in the presence of a particular HLA allele) in HIV-1 subtype C Gag-protease were identified using the large chronically infected cohort data by Dr. David Heckerman and Jonathan Carlson (eScience Group, Microsoft Research, Washington). This was done using published methods that take into account the phylogenetic relatedness of sequences, amino acid co-variation, and HLA linkage disequilibrium effects [269]. Briefly, a maximum likelihood phylogenetic tree was constructed for each gene. For each observed amino acid at each codon and every HLA allele, the likelihood that the observed amino acid was generated by the phylogenetic tree alone or alternatively HLA-mediated pressure was tested and p values were generated with the likelihood ratio test. These tests were binary – the presence versus absence of each amino acid variant at a codon was related to the presence versus absence of each HLA allele. The HLA allele most significantly associated with the variant was removed and the analysis repeated iteratively until no HLA alleles were significantly associated with the variant, thereby taking into account overlap of HLA associations. Multiple tests were addressed using q values, which are an estimate of the expected proportion of false positive associations for each p value threshold [270]. Only associations with $q \leq 0.2$ (a false-positive proportion of 20% among identified associations) were considered.

The relationships between number or proportion of HLA-associated polymorphisms and replication capacities were then assessed using Pearson's correlation (for normally-distributed variables) or Spearman's rank correlation (for non-normally distributed variables). For recently infected patients, Pearson's or Spearman's correlation was also used

to assess the relationship between the number of HLA-associated polymorphisms in Gag and viral set point. Where the HLA-associated polymorphism variable was binary (*i.e.* 1 polymorphism or none) the Student's T test or Mann-Whitney U test was used for analysis.

In an exploratory analysis, the Mann-Whitney U test was used to identify single, specific codons in Gag-protease at which amino acid variants (occurring at a frequency of $n \geq 5$) were associated with increased or decreased replication capacity. In addition, in the larger chronically infected cohort, linear regression with a forward selection process (a multivariate method) was used to identify single amino acids associated with changes in replication capacity. Q values were computed to account for multiple comparisons. To increase power, these analyses were also repeated limited to HLA-associated positions only.

Replication capacity versus HIV-1 subtype

Replication capacities of subtype B and C Gag-protease recombinant viruses were compared with the Student's T test.

Longitudinal replication capacity measurements

For 13 of the recently infected Durban subjects for whom replication capacity was measured at 2 time points approximately 1 year apart, a paired T test was used to test for overall significant change in replication capacities over this time. The longitudinal replication capacity measurements were also related to patient viral loads, CD4+ T cell counts and CD8+ T cell responses.

Significance threshold

The significance cut off for all analyses unless otherwise indicated was $p < 0.05$.

2.3 Measurement of replication capacities of intact viral isolates

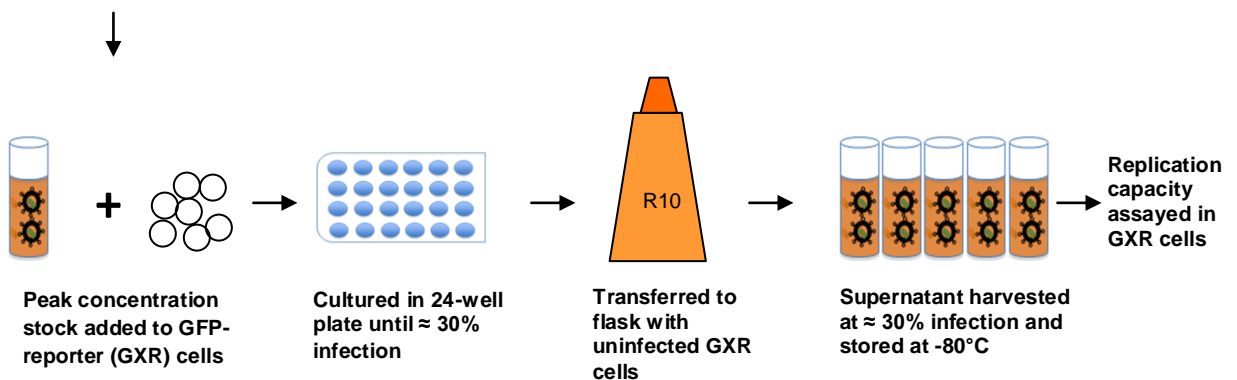
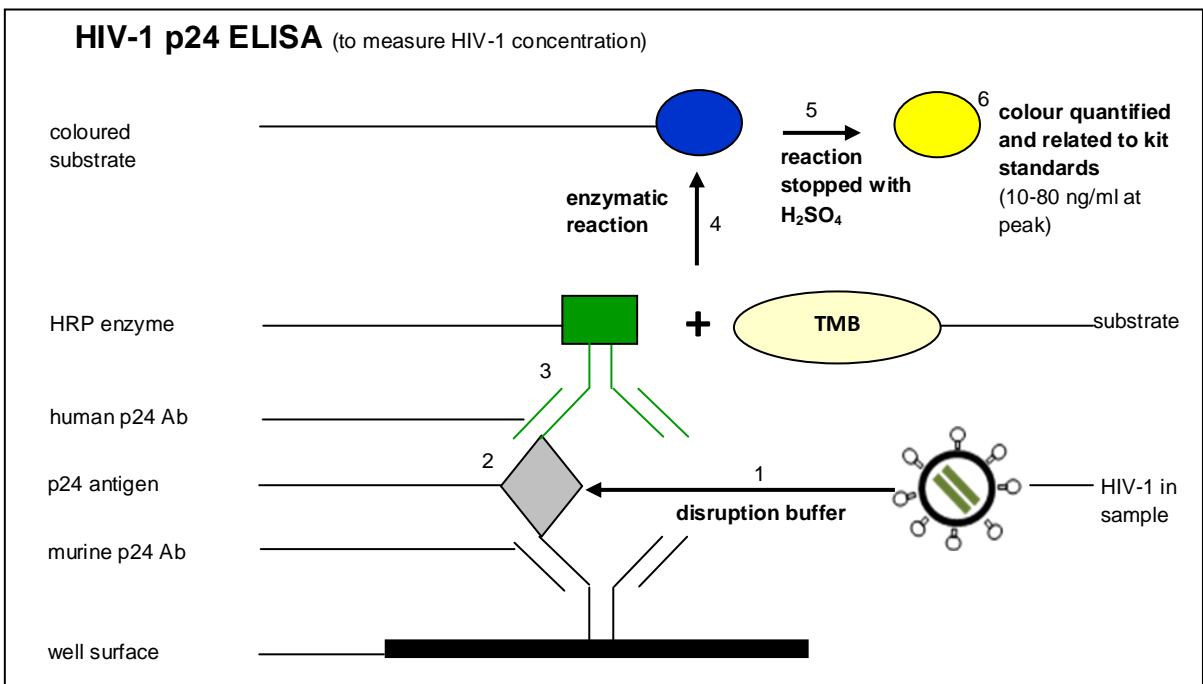
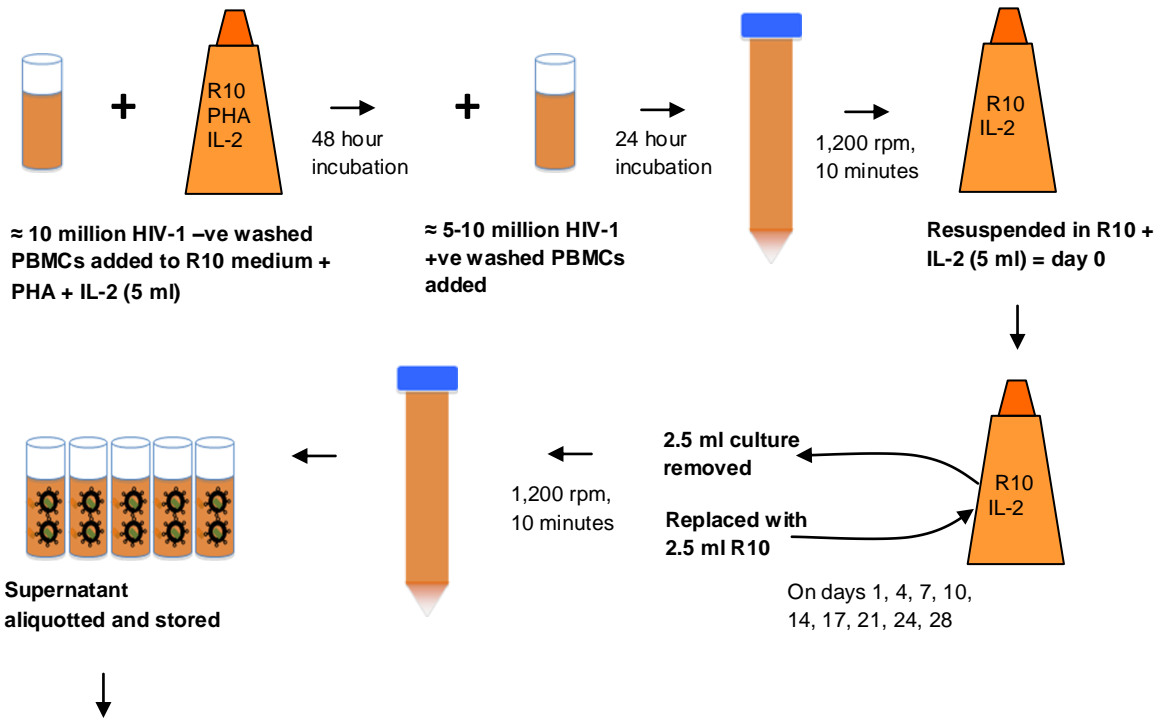
2.3.1 Isolation of HIV-1 and measurement of replication capacity

HIV-1 subtype C isolates were generated from PBMC samples obtained from 15 chronically infected patients (and from plasma for 1 patient [242]) from the Sinikithemba cohort in Durban, as shown in Figure 2.3. Uninfected PBMCs obtained from 2 anonymous HIV-negative healthy donors were thawed at 37 °C in a water bath, washed twice with 10 ml medium (RPMI-1640 containing L-glutamine and HEPES [Sigma] supplemented with 10% FBS [Gibco] and 50 U/ml penicillin-streptomycin [Gibco]) at 1,200 rpm for 10 minutes, and then combined. The uninfected PBMCs were stimulated for 72 hours at 37 °C and 5% CO₂ with 5 µg/ml phytohaemagglutinin (PHA; Sigma) and 20 U/ml interleukin-2 (IL-2; Roche Applied Science, Mannheim, Germany) in medium at a concentration of 2 million cells/ml. After 48 hours of stimulation, infected PBMCs (approximately 5-10 million) were similarly thawed, washed and then incubated with 10 million uninfected stimulated cells in a final volume of 5 ml for a further 24 hours. Thereafter cells were washed to remove PHA and maintained in R10 with 20 U/ml IL-2. Every 3-4 days, 2.5 ml supernatant was removed, stored in 500 µl aliquots at -80 °C, and replaced with fresh medium. The p24 concentration in stored supernatants was determined by an ELISA using the Vironostika HIV-1 Antigen Microelisa system (Biomérieux, Boxtel, The Netherlands) according to manufacturer's instructions as depicted in Figure 2.3. Supernatants corresponding to peak virus concentrations were further cultured in GXR cells to generate virus stocks. Virus growth was monitored by flow cytometry and supernatants were harvested when cultures were

Figure 2.3 Isolation of HIV-1 from infected PBMCs and generation of concentrated virus stocks

To isolate HIV-1 from infected PBMCs, uninfected PBMCs were first stimulated for 48 hours with R10 medium containing PHA and IL-2 before adding infected PBMCs, followed by stimulation for a further 24 hours. To remove PHA, PBMCs were washed by centrifugation followed by resuspension in R10 medium without PHA. To monitor virus concentration, at three to four day intervals supernatant was removed, stored in aliquots at -80 °C until HIV-1 p24 concentration was measured by ELISA, and replaced with fresh R10 medium. The ELISA procedure comprising the following steps is illustrated: (1) disruption of HIV-1 in supernatant, (2) binding of p24 to murine p24 antibody bound to plate surface, (3) binding of human p24 antibody with attached HRP to the p24 antigen, (4) HRP-mediated cleavage of added TMB substrate to form coloured product, (5) stopping of the reaction with H₂SO₄, and (6) quantification of the coloured product by spectrophotometry followed by calculation of p24 concentration using a curve generated from kit standards. Supernatants corresponding to peak virus concentrations (approximately 10-80 ng/ml) were cultured in GXR cells to generate concentrated virus stocks. Virus growth was monitored by flow cytometry since GXR cells express green fluorescent protein when HIV-1-infected and supernatants were harvested when cultures were approximately 30% infected. Supernatants were stored at -80 °C until assayed to determine replication capacities of HIV-1 isolates.

HIV-1 – human immunodeficiency virus type 1; PBMCs – peripheral blood mononuclear cells; PHA – phytohaemagglutinin; IL-2 – interleukin-2; ELISA – enzyme-linked immunosorbent assay; p24 – protein of 24 kDa; HRP – horse-radish peroxidase; TMB – tetramethylbenzidine; H₂SO₄ – sulphuric acid; GFP – green fluorescent protein.



approximately 30% infected (Sections 2.2.2.5-2.2.2.6). Virus stocks were stored at -80 °C until assayed in the GXR cells to determine viral titres and viral replication capacities as described in Sections 2.2.3.1-2.2.3.2.

2.3.2 Data analysis

Pearson's correlation was used to test for a significant association between replication capacities of viral isolates and recombinant viruses encoding corresponding Gag-protease from isolates. The level of significance was $p < 0.05$.

2.4 Measurement of replication capacities of mutant viruses

2.4.1 Construction of mutant plasmids

To confirm suspected fitness cost or compensation of Gag-protease mutations identified from the analysis in Section 2.2, these mutations were introduced singly and in combination into a plasmid containing HIV-1 *gag-protease* and the ampicillin-resistance gene by site-directed mutagenesis. To investigate fitness differences between HIV-1 subtype B and C, the same procedure was used to test the fitness effects of subtype B- or C-specific residues at codons that were statistically associated with alterations in viral replication capacity [251] (Section 2.2).

2.4.1.1 The site-directed mutagenesis process

Briefly, the process of site-directed mutagenesis involves PCR amplification of a plasmid using primers encoding the desired mutation [271, 272]. The PCR product is the modified plasmid containing the mutation. The original plasmid (isolated from *E. coli*) is removed by

digestion with DpnI enzyme since this enzyme is specific for methylated DNA, and DNA isolated from bacterial cells is methylated while that amplified by PCR is not. To achieve plasmid nick repair, the modified plasmid is transformed by heat-shock into bacterial cells followed by plating onto LB-ampicillin agar plates for propagation of cells transformed by the ampicillin-resistant plasmid only. Colonies containing the plasmid may be checked for the presence of the desired mutation by sequencing. Stocks of modified plasmid may then be generated through propagation of the selected colony in LB broth followed by plasmid purification. The process of site-directed mutagenesis as conducted in the present study is illustrated in Figure 2.4.

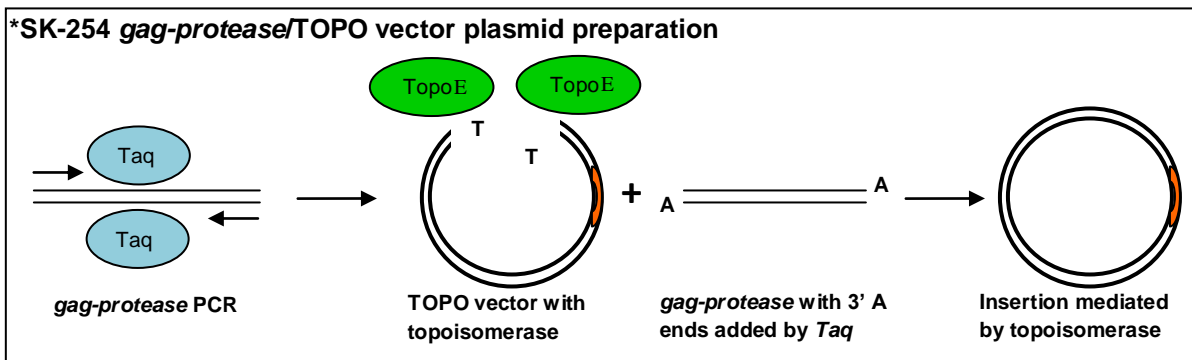
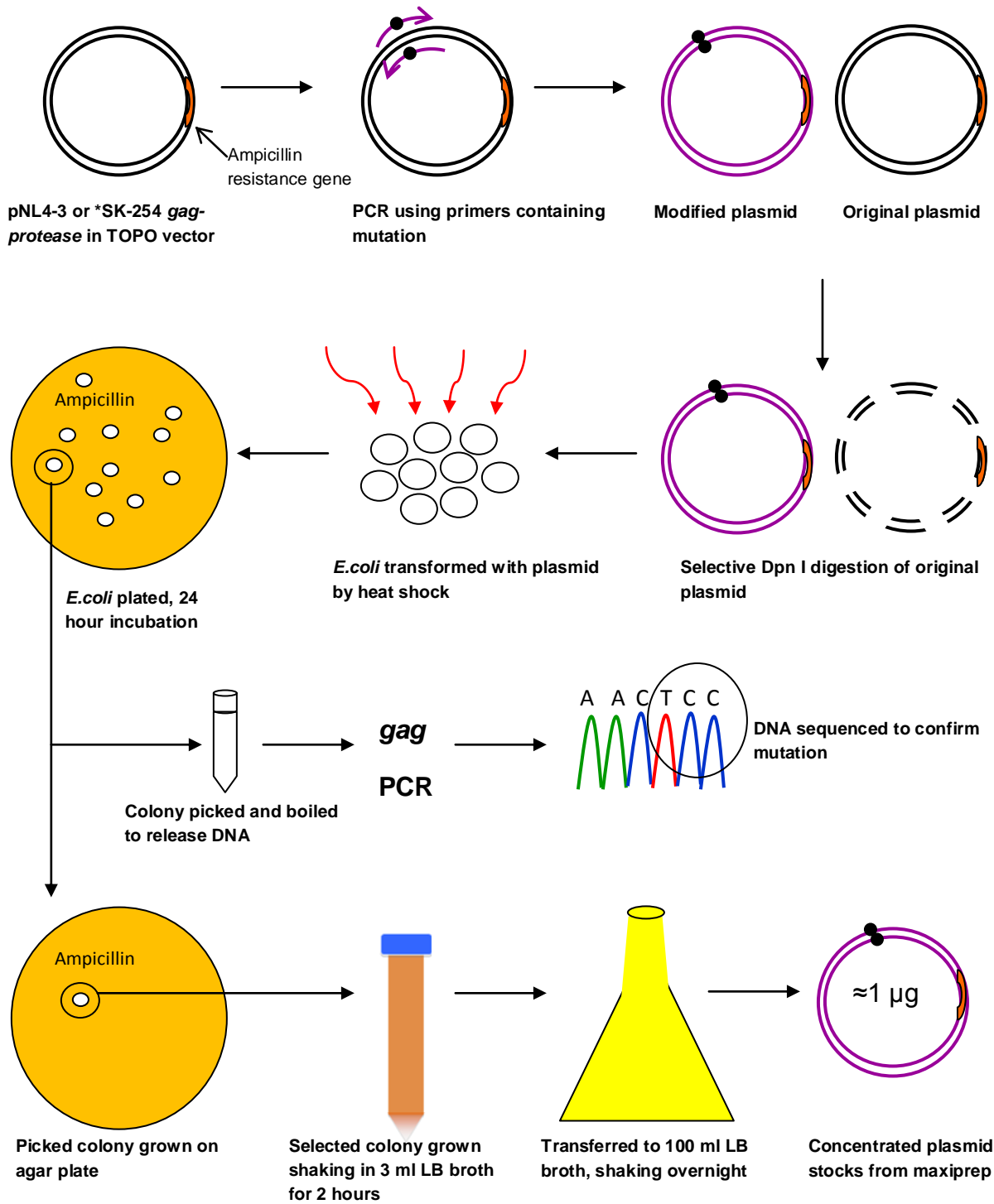
2.4.1.2 Preparation of plasmid vectors

Mutations were first introduced into the HIV-1 *gag-protease* gene isolated from patient SK-254 (Genbank accession number HM593258) since this sequence was most similar to consensus subtype C Gag-protease from all patient-derived sequences generated (Section 2.2). SK-254 *gag-protease* PCR product was first purified (PCR purification kit from Qiagen) and then cloned into a TOPO vector plasmid (pCR2.1-TOPO) using the TOPO TA cloning kit (Invitrogen) and One Shot TOP10 Chemically Competent *E. coli* (Invitrogen) as per manufacturer's instructions (Figure 2.4). Briefly, TA cloning is the insertion of a Taq-amplified PCR product with single 3' deoxyadenosine nucleotide (A) overhangs (added by the Taq enzyme) into a linearised plasmid vector with single 3' deoxythymidine nucleotide (T) overhangs. The ligation of the overhangs is mediated by topoisomerase I which is covalently bound to the TOPO vector. Following the cloning procedure, transformed colonies were analysed by PCR, followed by gel electrophoresis, for the *gag-protease* insert. The final volume of the PCR reaction was 10 µl, comprised of 6.6 µl DEPC water, 1 µl 10X Ex Taq buffer, 0.8 µl dNTP, 0.25 µl 100-mer forward primer (10 µM), 0.25 µl 100-mer

Figure 2.4 Site-directed mutagenesis procedure

Diagram showing the introduction of mutations into NL4-3 and SK-254 *gag-protease*/TOPO vector plasmids by site-directed mutagenesis, as well as the preparation of the SK-254 *gag-protease*/TOPO vector plasmid. SK-254 *gag-protease* was inserted into the TOPO vector with a TOPO cloning kit. The TOPO vector has T overhangs and an attached topoisomerase enzyme (Topo E) which mediates the ligation of the vector with PCR product containing A overhangs (introduced by *Taq* polymerase during PCR). Mutations were introduced into the plasmids by PCR with primers containing the mutation. The original plasmid was removed by digestion with DpnI which selectively recognises the original plasmid as it contains methylated sites due to its prior propagation in bacterial cells. The modified plasmid was then transformed into *E. coli* by heat shock and plated on ampicillin-containing agar plates. Due to the presence of an ampicillin resistance gene in the plasmids, only transformed cells grow on the ampicillin-containing agar. A colony was picked, touched to another agar plate to keep a record, and checked for the presence of the plasmid with the introduced mutation by sequencing. Following confirmation of the introduced mutation, the selected colony was propagated in LB broth and a concentrated stock of purified modified plasmid was isolated using a maxiprep kit.

pNL4-3 – NL4-3 plasmid; *gag* – group specific antigen; PCR – polymerase chain reaction; Topo E - topoisomerase enzyme; T – deoxythymidine nucleotide; A – deoxyadenosine nucleotide; DNA – deoxyribonucleic acid; LB - Luria-Bertani.



reverse primer (10 μ M), 0.08 μ l Ex Taq, and 1 μ l of a boiled colony suspension (1 colony in 10 μ l DEPC water). Details of 100-mer primers and thermocycler conditions used are described in Section 2.2.2.1. A colony containing the SK-254 *gag-protease* insert was grown in 3 ml LB broth overnight and then transferred to 100 ml LB broth and grown overnight. Purified plasmid stocks were generated using the Plasmid Maxi kit (Qiagen).

The same mutations were also introduced into a NL4-3 (HIV-1 subtype B) plasmid (pNL4-3) containing the full HIV-1 genome. The NL4-3 plasmid was supplied in *E. coli* STBL3 cells by Dr. Mark Brockman (Simon Fraser University, Canada; formerly of the Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology and Harvard University). Stocks were amplified as described in Section 2.2.2.2.

2.4.1.3 Introduction of mutations into plasmid vectors

The following positions in SK-254 and NL4-3 Gag were then mutated using the QuikChange II XL Site-Directed Mutagenesis kit (Stratagene, La Jolla, CA) and the relevant mutagenesis primers (high-performance liquid chromatography purified, 25-45 base pairs, melting temperature >75 $^{\circ}$ C) according to kit specifications: Q182S, T186S, T190A, T190I, Q182S/T186S, T186S/T190A, and T186S/T190I. Subtype C-specific Gag residues that were associated with altered replication capacity (67A, 473A, 483-/484-) were introduced into pNL4-3 and the corresponding clade B-specific residues (67S, 473P, 483L/484Y) were introduced into SK-254 *gag-protease* in the TOPO vector. Following colony PCR (Section 2.4.1.2) and sequencing (Section 2.2.2.1) to confirm the presence of introduced mutations, the selected colony was propagated and mutant plasmid stocks were purified (Section 2.4.1.2). The forward mutagenesis primers used to generate the various mutants are shown in Table 2.4. The reverse mutagenesis primers comprised the reverse complement of these.

Table 2.4 Forward primers used to introduce specific Gag mutations into SK-254 HIV-1 Gag-protease and pNL4-3

Template	Mutant	Primer
SK-254	186S	5'CCACAAGATTTAAACTCCATGTAAATACAGTGGGGGG3'
SK-254	182S	5'CAGAAGGAGCCACCCCA TCA GATTTAAACACCATGTAAATAC3'
SK-254	190A	5'CCACAAGATTTAAACACCATGTAAAT GCA GTGGGGGGAC3'
SK-254	190I	5'CCACAAGATTTAAACACCATGTAAAT ATA GTGGGGGGAC3'
SK-254 186S	186S/182S	5'CAGAAGGAGCCACCCCA TCA GATTTAAACTCCATGTAAATAC3'
SK-254 186S	186S/190A	5'CCACAAGATTTAAACTCCATGTAAAT GCA GTGGGGGGAC3'
SK-254 186S	186S/190I	5'CCACAAGATTTAAACTCCATGTAAAT ATA GTGGGGGGAC3'
SK-254	67S	5'CAGCTACAACCA TCC TTCAGACAGGAACAGAGGAAC3'
SK-254	473P	5'GAGGAGACAACCCCC CCT CCTCCGAAGCAGGAG3'
SK-254	483L/484Y	5'GAAAGACAGGGAA CTGTAT CCCTTAACTTCCCTC3'
NL4-3	186S	5'CCACAAGATTTAAAT TCC ATGCTAAACACAGTGGGGGG3'
NL4-3	182S	5'CAGAAGGAGCCACCCCA TCA GATTTAAATACCATGCTAAACAC3'
NL4-3	190A	5'CCACAAGATTTAAATACCATGCTAAAC GCA GTGGGGGGAC3'
NL4-3	190I	5'CCACAAGATTTAAATACCATGCTAAAC ATA GTGGGGGGAC3'
NL4-3 186S	186S/182S	5'CAGAAGGAGCCACCCCA TCA GATTTAAAT TCC ATGCTAAACAC3'
NL4-3 186S	186S/190A	5'CCACAAGATTTAAAT TCC ATGCTAAAC GCA GTGGGGGGAC3'
NL4-3 186S	186S/190I	5'CCACAAGATTTAAAT TCC ATGCTAAAC ATA GTGGGGGGAC3'
NL4-3	67A	5'CTGGGACAGCTACAACCAG GCT TTCAGACAGGATCAG3'
NL4-3	473A	5'GAGACAACAACCTCCC GCT CAGAGGCAGGAGCCG3'
NL4-3	483-/484-	5'GGAGCCGATAGACAAGGAA-----CCTTTAGCTTCCCTC3'

Mutated Gag codons (HXB2 numbered) are shown in red.

2.4.2 Generation of mutant viruses and measurement of replication capacity

Mutated SK-254 *gag-protease* was amplified using 100-mer primers complementary to NL4-3 on either side of *gag-protease*, followed by electroporation of GXR cells with PCR product and pNL4-3 Δ *gag-protease* to generate mutant viruses (Sections 2.2.2.1-2.2.2.4). Similarly, mutated NL4-3 plasmid was electroporated into GXR cells. However, the modified electroporation protocol (Section 2.2.2.4) was used to generate the mutant 186S/190I in the NL4-3 backbone (due to difficulty generating concentrated stocks of this mutant with the original protocol). Virus growth was monitored by flow cytometry and supernatants were harvested when cultures were approximately 30% infected (Sections 2.2.2.5-2.2.2.6). Virus stocks were stored at -80 °C until assayed in the GXR cells to determine viral titres and viral replication capacities as described in Sections 2.2.3.1-2.2.3.2. Replication capacities of the same mutant viruses generated by both electroporation protocols were compared to rule out the possibility of variability introduced by the use of the modified protocol.

2.4.3 Data analysis

ANOVA with Tukey post-hoc tests was used to test for significant differences in replication capacities between wild-type viruses and viruses encoding mutations. The level of significance was $p < 0.05$.

CHAPTER 3

RESULTS

CHAPTER 3 - RESULTS

3.1 Quality control

3.1.1 Validation of generated Gag-protease NL4-3 recombinant viruses

To test whether generated recombinant viruses were representative of the original plasma quasispecies, *gag-protease* was first re-sequenced from 40 randomly-selected recombinant viruses generated from HIV-1 subtype C chronically infected patients and compared with the original plasma HIV-1 RNA sequences. The median number of total nucleotide differences between recombinant virus and original plasma HIV-1 RNA sequence (when mixtures were not included as differences) was 0 (IQR: 0-1.5), resulting in an average nucleotide similarity of 99% between pairs. The average number of nucleotide mixtures in recombinant virus sequences was 21 (standard deviation [SD] =17), indicating somewhat reduced diversity (Student's T test; $p=0.0002$) when compared to the original plasma sequences (mean=35, SD=36). Since these data indicated that Gag-protease recombinant viruses were representative of the original plasma quasispecies, only 2 randomly selected recombinant viruses generated from HIV-1 subtype C recently infected patients were re-sequenced, and were found to be identical to the plasma nucleotide sequences. A phylogenetic tree incorporating these 42 sequence pairs shows that the *gag* sequences from recombinant viruses clustered closely with respective *gag* sequences from plasma (Figure 3.1). Since Gag-protease recombinant viruses were representative of the original plasma quasispecies, all further analyses are based on the original plasma HIV-1 sequences.

Figure 3.1 Neighbour-joining tree (Paup 4.0) of HIV-1 *gag* sequences

Gag-protease NL4-3 recombinant viruses were generated by co-transfection of a CEM-derived T cell line with patient-derived *gag-protease* amplicons and *gag-protease*-deleted NL4-3 plasmid. Forty-two randomly-selected recombinant viruses were re-sequenced and compared with the original plasma HIV-1 RNA sequences. *Gag* recombinant virus nucleotide sequences (pink) closely clustered with respective plasma *gag* nucleotide sequences (blue) in a tree.

HIV-1 – human immunodeficiency virus type 1; *gag* – group specific antigen; RNA – ribonucleic acid.

3.1.2 Validation of replication capacity assay

The replication capacities of Gag-protease NL4-3 recombinant viruses were assayed in duplicate independently. Duplicate measurements were highly concordant (Pearson's correlation; $r=0.88$, $p<0.0001$) (Figure 3.2i).

Accuracy of recombinant viral titres was achieved: on day 3 of the assay the median % GFP-expressing cells was 0.66% (IQR: 0.46-0.82%). Importantly, the observed variability in day 3 readings did not significantly influence viral replication capacity, as calculated from the slope of growth from days 3-6 post-infection (Spearman's correlation; $r=0.05$, $p=0.23$) (Figure 3.2ii).

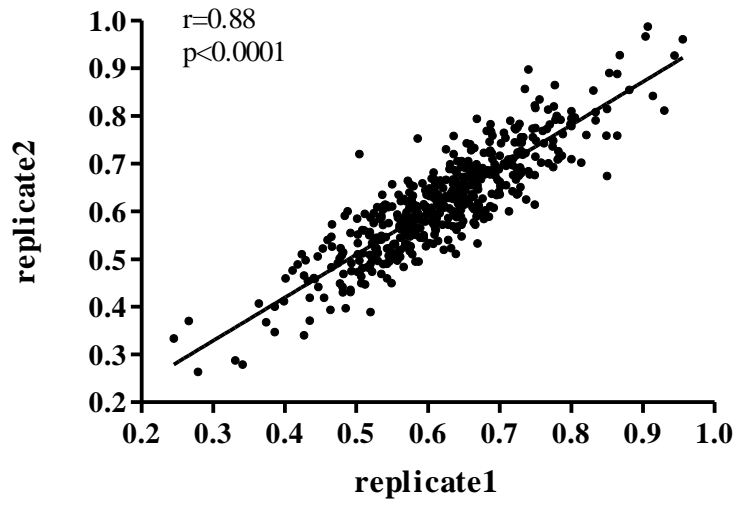
3.1.3 Overall method variability: recombinant virus generation and replication capacity assay

Firstly, 5 samples were randomly selected and methods were repeated 5 times independently on each. There was a significant difference overall between the groups of measurements (ANOVA; $p=0.0002$), indicating that the method variability was not great enough to negate differences in replication capacity between samples (Figure 3.3i). However, these 5 randomly selected samples were similar in replication capacity and thus overlapped. To test for reproducible differences between samples, 2 samples with low and high replication capacities, respectively, were selected for repeated measurements. The measurements between these 2 samples were distinctly different (Student's T test; $p<0.0001$) (Figure 3.3ii), and each replicate of 1 sample was significantly different from each replicate of the other sample (Tukey post hoc tests; $p<0.05$) (Figure 3.3iii), indicating reproducible differences.

Figure 3.2 Replication capacity assay quality control analyses

- (i) Duplicate replication capacity results displayed good concordance (Pearson's correlation).
- (ii) The range in starting percentage infection on day 3 of the assay did not influence the calculation of replication capacity measured by the slope of exponential growth from days 3-6 post-infection (Spearman's correlation).

i



ii

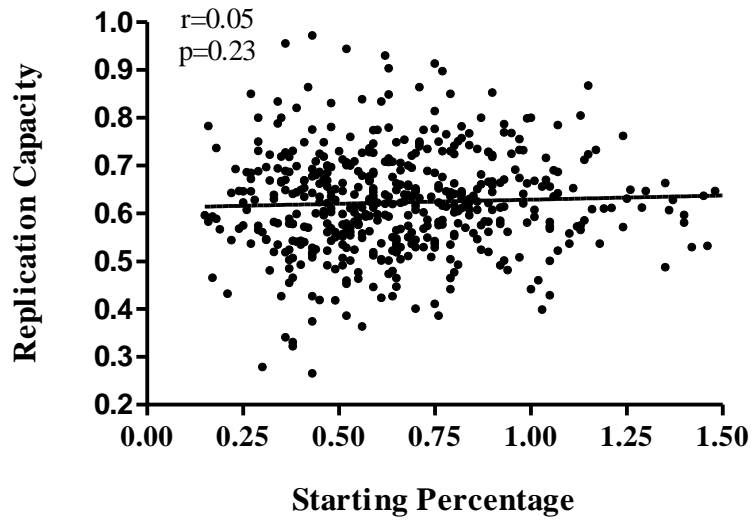


Figure 3.3 Overall method variability: recombinant virus generation and replication capacity assay

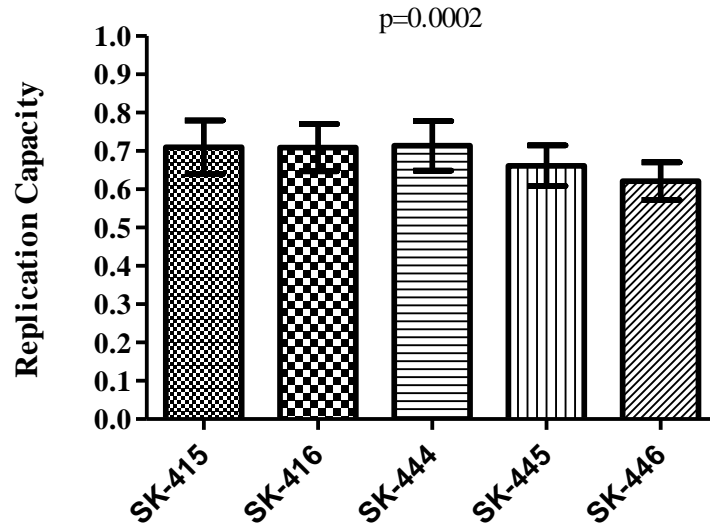
(i) Generation of recombinant viruses encoding plasma-derived HIV-1 Gag-protease followed by replication capacity measurement was repeated 5 times independently for 5 randomly selected plasma samples from the SK cohort of patients chronically infected with HIV-1 subtype C. Lines represent the mean and bars represent standard deviation. A significant difference between the groups of measurements for different patients is shown (ANOVA).

(ii) A reproducible difference in the replication capacities of independently repeated experiments from one low (SK-68) and one high (SK-276) replication capacity sample is shown (Student's T test). Lines represent the mean.

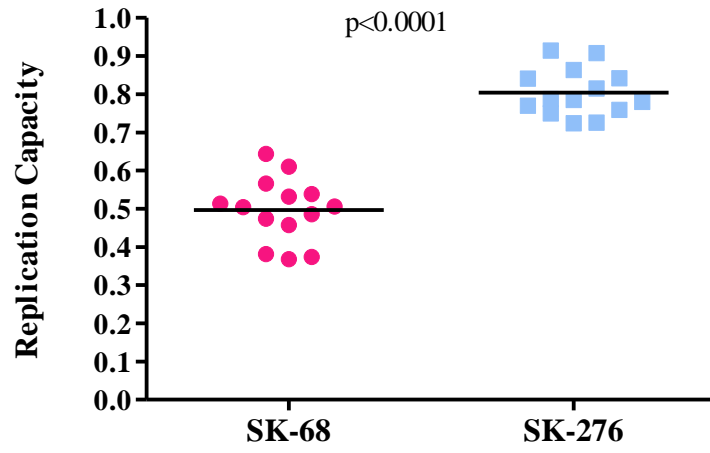
(iii) Graph showing significantly lower replication capacities of each SK-68 replicate compared to each SK-276 replicate (ANOVA, Tukey post-hoc tests). Lines represent the mean.

HIV-1 – human immunodeficiency virus type 1; Gag – group specific antigen; SK – Sinikithemba; ANOVA – analysis of variance.

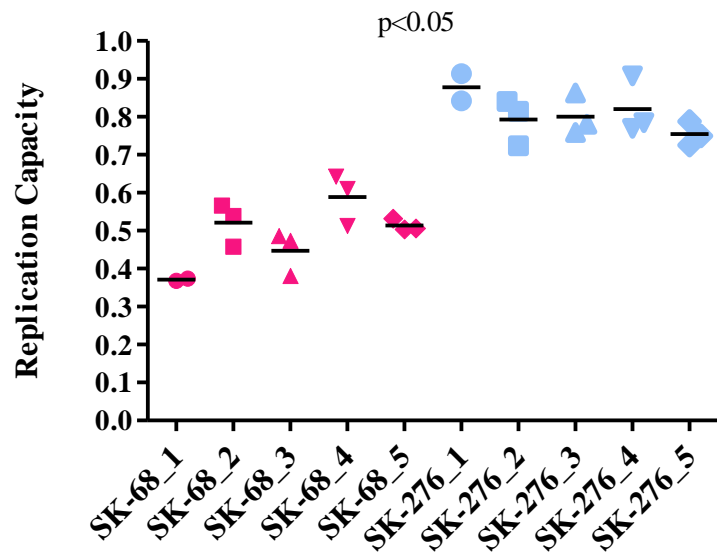
i



ii



iii



3.2 HIV-1 subtype C chronic infection

3.2.1 Distribution of replication capacities

The NL4-3-normalised replication capacities of the recombinant viruses generated from the 406 HIV-1 subtype C chronically infected patients approximated a normal distribution (mean=0.62, SD=0.1) (Figure 3.4).

3.2.2 Replication capacities versus HLA class I alleles

Replication capacities of recombinant viruses were grouped according to HLA alleles expressed by the host (Figure 3.5i-iii). Overall, replication capacities varied significantly between the different HLA-B alleles (ANOVA, $p=0.01$) but not between HLA-A or HLA-C alleles, suggesting that HLA-B alleles have the greatest impact on Gag-protease-mediated replication capacity. Relationships between specific HLA alleles and replication capacity were also observed, the strongest of which was the association of HLA-B*81 with lower replication capacities (Student's T test; $p<0.0001$). It should be noted that p values presented are uncorrected for multiple comparisons. Only the association of HLA-B*81 with lower replication capacities would remain statistically significant following Bonferroni adjustment for multiple comparisons. It should also be noted that due to the limitation of sample size, the interactions between different HLA alleles expressed within a patient were not controlled for statistically, however linkage disequilibrium between HLA alleles was taken into account in order to partially control for this.

Besides HLA-B*81, other alleles that were associated with low-replication-capacity recombinant viruses were HLA-B*5801, HLA-A*0205 (Mann-Whitney U-test; $p=0.05$ and $p=0.04$, respectively), HLA-A*3009, and HLA-A*3001 (Student's T test; $p=0.02$ and

Figure 3.4 Replication capacities of Gag-protease NL4-3 recombinant viruses encoding Gag-protease from HIV-1 subtype C chronically infected patients

Histogram showing an approximately normal distribution of replication capacities (normalised to the growth of wild-type NL4-3) with a mean of 0.62 and standard deviation of 0.1.

Gag – group specific antigen; HIV-1 – human immunodeficiency virus type 1.

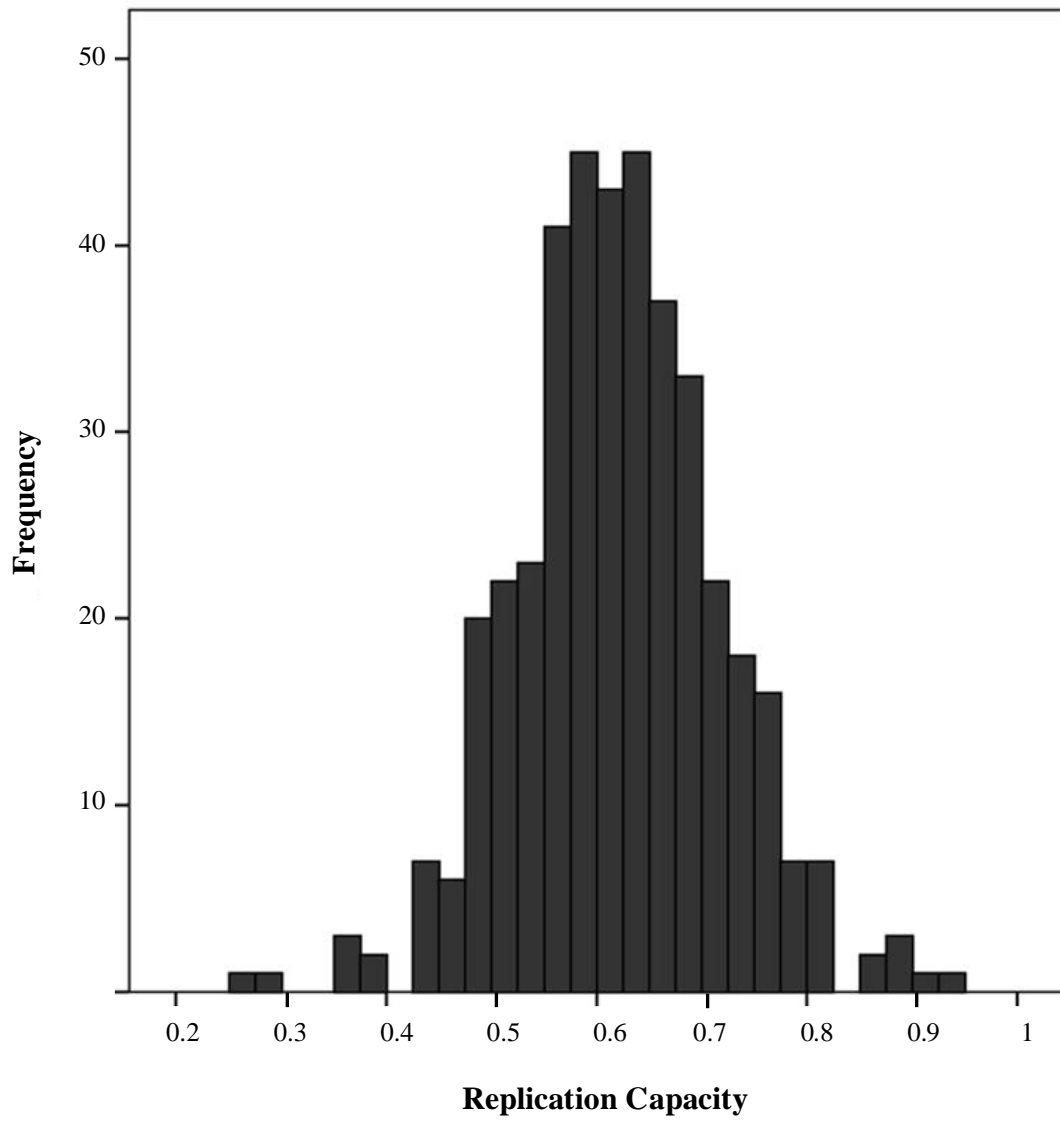


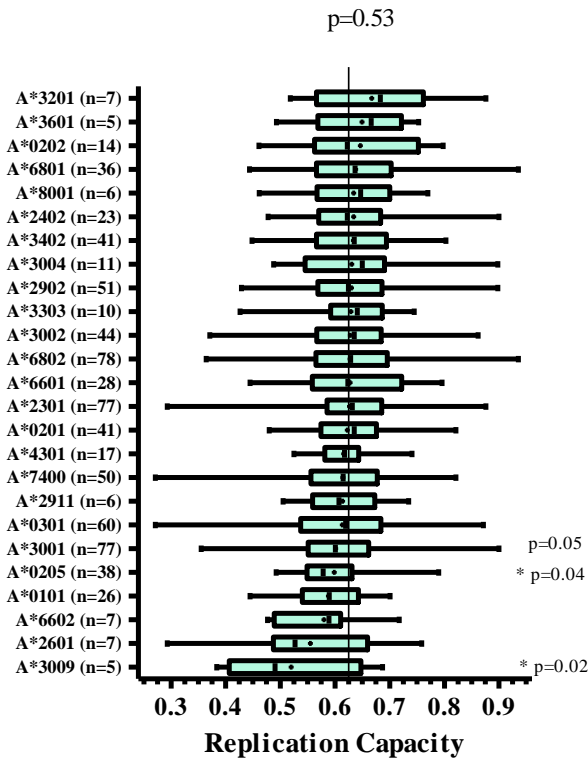
Figure 3.5 Associations between HLA alleles and replication capacities of recombinant viruses encoding Gag-protease from HIV-1 subtype C chronically infected patients

(i-iii) Graphs show the mean (dot), median (vertical line), inter-quartile range (edges of boxes), and most extreme values (edges of whiskers) of replication capacities for each different HLA allele for which $n \geq 5$. Individual significant ($p < 0.05$) associations (Student's T test) are indicated with *. Associations that survive Bonferroni correction for multiple comparisons are indicated with **. Overall p values (ANOVA) indicate significance of differences in replication capacity between alleles of each group, A, B and C.

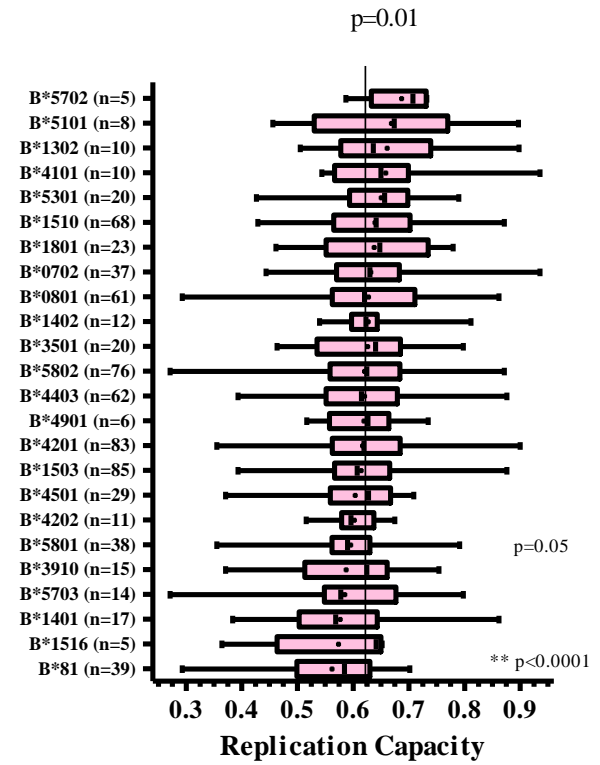
(iv) Graph showing a significantly greater proportion (Fisher's exact test) of individuals with protective HLA alleles (D= HLA-B*57, HLA-B*5801, and HLA-B*81) in the group with the least fit viruses (replication capacity ≤ 0.5) compared to those with the fittest viruses (replication capacity ≥ 0.74).

HLA – human leukocyte antigen; Gag – group specific antigen; HIV-1 – human immunodeficiency virus type 1; ANOVA – analysis of variance.

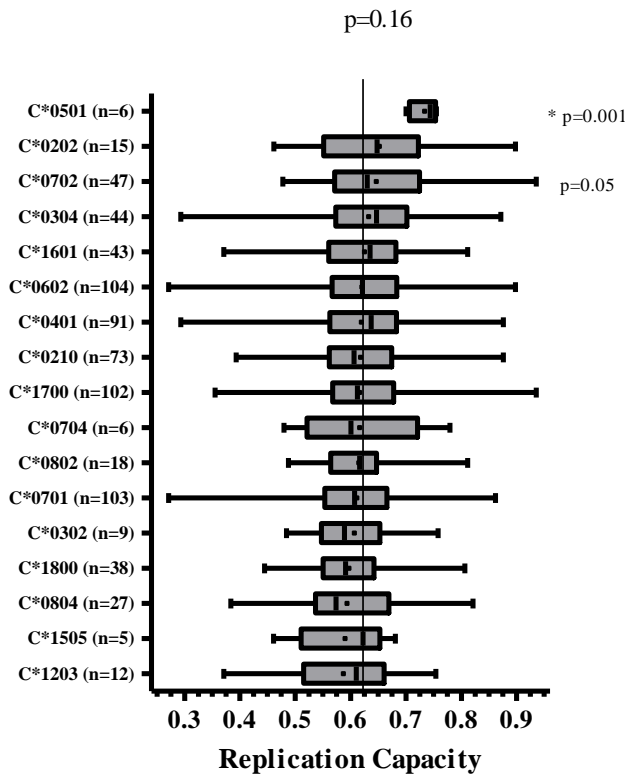
i



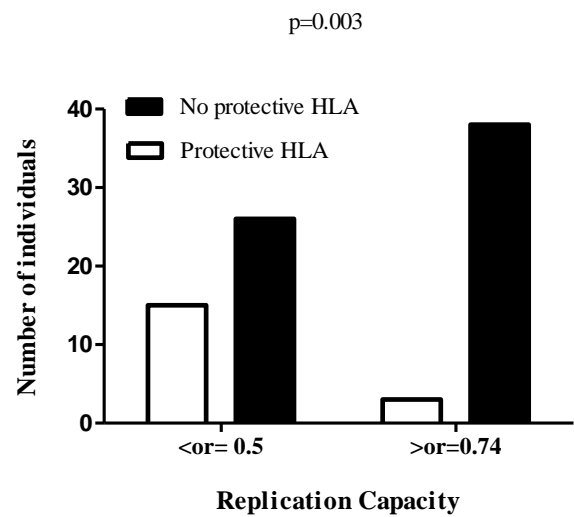
ii



iii



iv



$p=0.05$, respectively). Due to tight linkage between HLA-B*5801 and HLA-A*0205 (linkage disequilibrium $[D']=0.56$ in [175]), the allele driving the effect could not be identified. Out of 5 individuals with HLA-B*3009, 4 possessed HLA-B*81, which likely explains the association of HLA-B*3009 with lower replication capacities.

Alleles associated with higher replication capacities were HLA-C*0702 (Student's T test; $p=0.05$) and HLA-C*0501 (Mann-Whitney U-test; $p=0.001$). Only 6 individuals in this study possessed HLA-C*0501 and all were linked to replication capacities above the 80th percentile of the dataset (>0.7). HLA-C*0501 is in linkage disequilibrium with HLA-B*1801 – 5 out of 6 individuals with HLA-C*0501 also carried HLA-B*1801 - and could therefore partly contribute to the disadvantage associated with HLA-B*1801 in subtype C infection [170].

In an additional analysis, the HLA types of the individuals corresponding to the fittest recombinant viruses (≥ 90 th percentile of the dataset, *i.e.* ≥ 0.74 , $n=41$) were compared to the HLA types of individuals with the least fit recombinant viruses (10th percentile of the dataset, *i.e.* ≤ 0.5 , $n=41$). Protective alleles were defined as those that were most strongly associated with lower viral loads in HIV-1 subtype C-infected individuals, namely HLA-B*57, HLA-B*5801, and HLA-B*81 [170] (Section 1.6.4.1), which were also later found to be the most strongly associated with lower viral loads or higher CD4+ T cell counts in a cohort of over 1,000 HIV-1 subtype C-infected individuals [273]. The proportion of individuals possessing a protective allele was significantly greater in the low-replication-capacity group when compared with the high-replication-capacity group (Fisher's exact test; $p=0.003$) (Figure 3.5iv).

3.2.3 Replication capacities versus viral loads and CD4+ T cell counts

Replication capacities of recombinant viruses correlated positively with baseline log viral loads (Spearman's correlation; $r=0.24$, $p<0.0001$) and negatively with baseline CD4+ T cell counts (Spearman's correlation; $r=-0.17$, $p=0.0004$) (Figure 3.6i-ii). These effects remained after removal of protective alleles HLA-B*57, HLA-B*5801, and HLA-B*81 from analysis (Spearman's correlation; $r=0.18$ $p=0.001$ and $r=-0.14$ $p=0.01$, respectively). Of interest, analysis of the relationship between viral load/CD4+ T cell count and replication capacity among individuals expressing these protective alleles also revealed a significant positive correlation (Pearson's correlation; $r=0.33$ $p=0.001$ and $r=-0.33$ $p=0.001$, respectively).

3.2.4 Replication capacities versus rates of CD4+ T cell decline

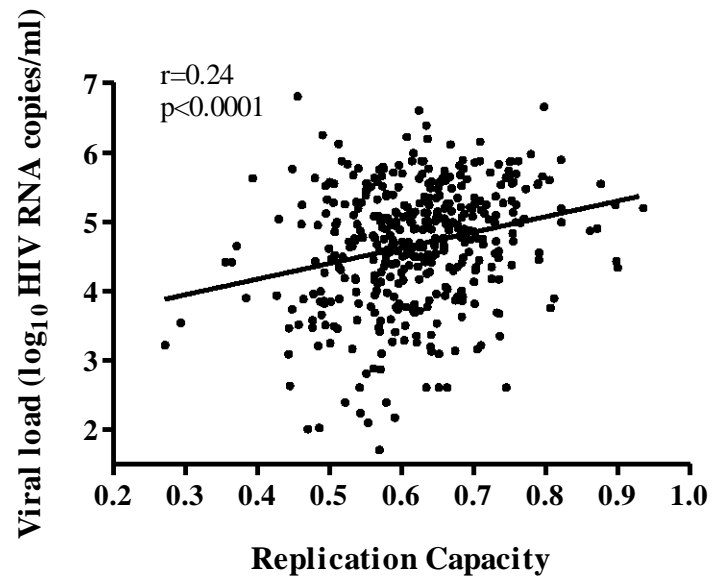
Rate of CD4+ T cell decline following baseline was calculated for 339 HIV-1 subtype C chronically infected patients. There was no statistically significant relationship between replication capacity and CD4+ T cell decline overall (Spearman's correlation; $r=-0.01$, $p=0.79$). Since baseline CD4+ T cell count is a determinant of rate of CD4+ T cell decline and individuals with lower baseline CD4+ T cell counts have a subsequently slower rate of CD4+ T cell decline [274], the analysis was repeated following stratification by baseline CD4+ T cell count in increments of 50 cells/mm³. These stratifications failed to reveal any significant correlations between replication capacity and rate of CD4+ T cell decline (data not shown). Figure 3.6iii shows a lack of correlation between CD4+ T cell decline and Gag-protease-mediated replication capacity at baseline CD4+ T cell count ≥ 200 (Spearman's correlation; $r=-0.02$, $p=0.73$).

Figure 3.6 Relationships between replication capacities of recombinant viruses encoding patient Gag-protease and markers of disease progression in HIV-1 subtype C chronically infected patients

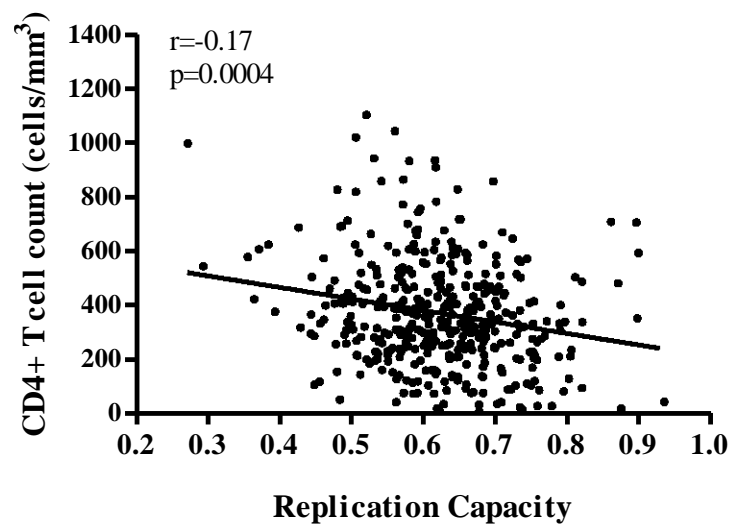
Graphs show correlations between replication capacity and baseline log viral load (i), baseline CD4+ T cell count (ii), and rate of CD4+ T cell decline (iii) (Spearman's correlation).

Gag – group specific antigen; HIV-1 – human immunodeficiency virus type 1; RNA – ribonucleic acid; CD4 – cluster of differentiation 4.

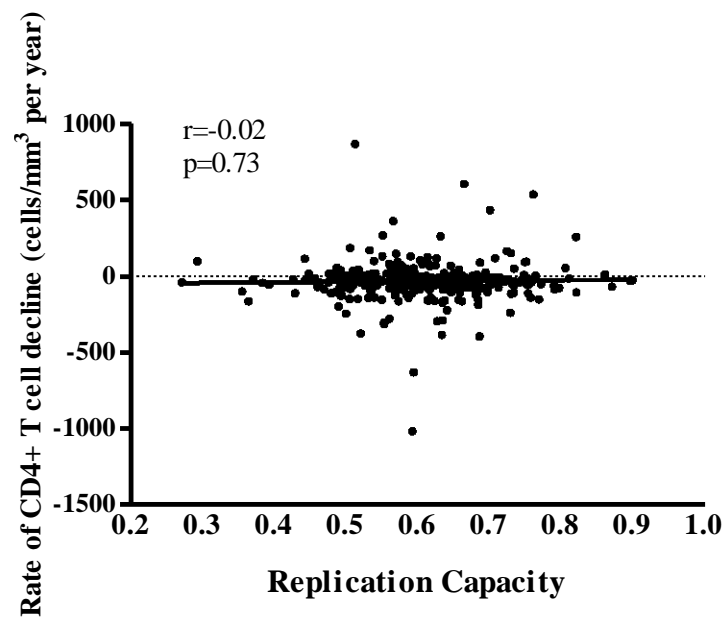
i



ii



iii



3.2.5 Replication capacities versus Gag-protease sequences

3.2.5.1 Gag-protease polymorphisms

To investigate whether an increasing number of polymorphisms in Gag or protease would tend to reduce replication capacity, the percentage similarities of both Gag and protease sequences to the consensus C sequence were calculated. Unexpectedly, the calculated Gag percentage similarity correlated negatively, although weakly, with replication capacity (Pearson's correlation; $r=-0.18$, $p=0.0004$), *i.e.* the fittest viruses were generally least like the consensus sequence (Figure 3.7i). This analysis was repeated separately for each region of Gag, namely p17, p24, p7, and p6, to see whether this relationship differed between regions. There remained an inverse relationship between percentage similarity to consensus and replication capacity in every region of Gag except p24, although this was only statistically significant for p17 and p7 (Figure 3.7ii). There was no correlation between percentage similarity to Gag p24 consensus C and replication capacity. In contrast, the majority of non-consensus residues in p17/p7 increased replication capacity. It should be noted that divergence from the consensus C Gag sequence did not represent convergence to the consensus B Gag sequence, which would have indicated that divergence from consensus C resulted in a better compatibility with the subtype B NL4-3 backbone, and therefore fitter viruses. The calculated protease percentage similarity to consensus did not correlate with replication capacity (Spearman's correlation; $r=-0.006$, $p=0.91$) (Figure 3.7iii).

3.2.5.2 HLA-associated polymorphisms

HLA-associated amino acids in Gag and protease were identified ($p<0.05$, $q\leq 0.2$). There were 131 amino acids either positively or negatively associated with an HLA allele at 66 different codons in Gag and only 3 amino acids at 2 different codons in protease (Appendix, Table A1). Thus, HLA-associated polymorphism analyses were focussed on Gag.

Figure 3.7 Associations between Gag-protease polymorphisms and replication capacities of recombinant viruses encoding Gag-protease from HIV-1 subtype C chronically infected patients

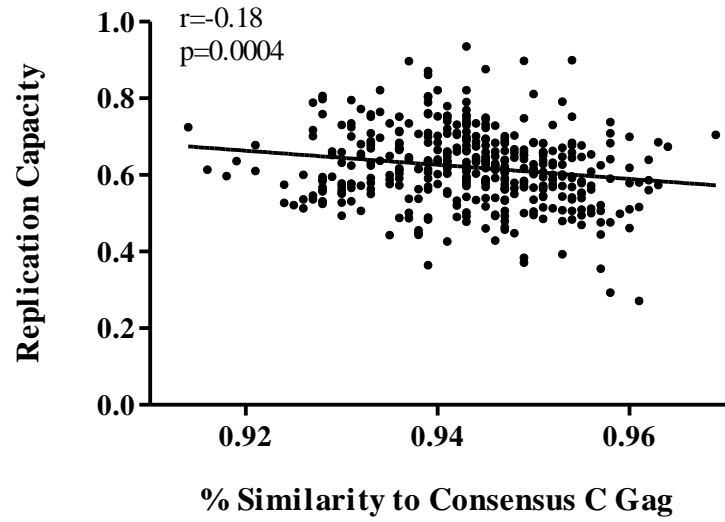
(i) A significant negative correlation between percentage similarity of Gag sequences to the consensus C Gag sequence and replication capacity (Pearson's correlation).

(ii) Gag p17, p24, p7, and p6 percentage similarity to consensus C sequence versus replication capacity (Spearman's correlation).

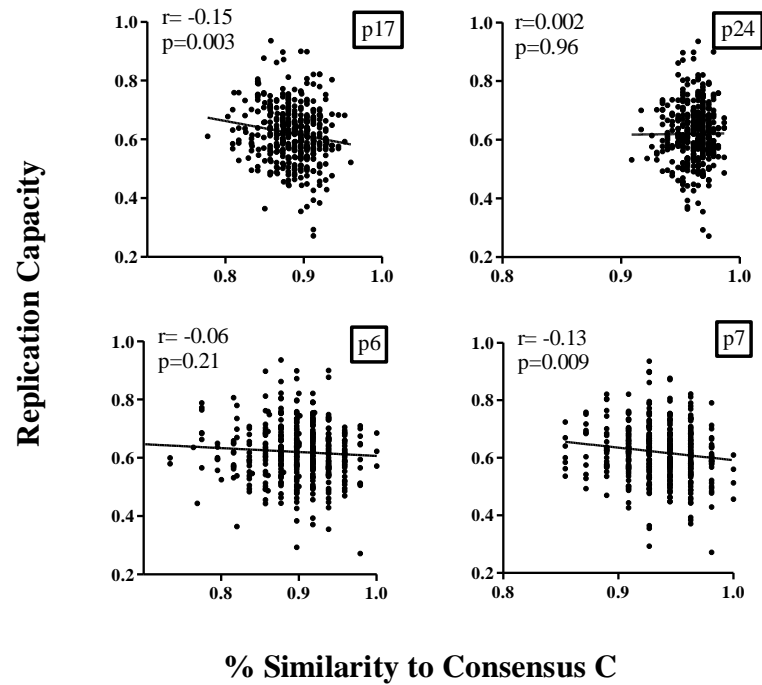
(iii) No relationship between the percentage similarity of protease sequences to the consensus C sequence (Spearman's correlation).

Gag – group specific antigen; HIV-1 – human immunodeficiency virus type 1; p17 – protein of 17 kDa; p24 – protein of 24 kDa; p7 – protein of 7 kDa; p6 – protein of 6 kDa.

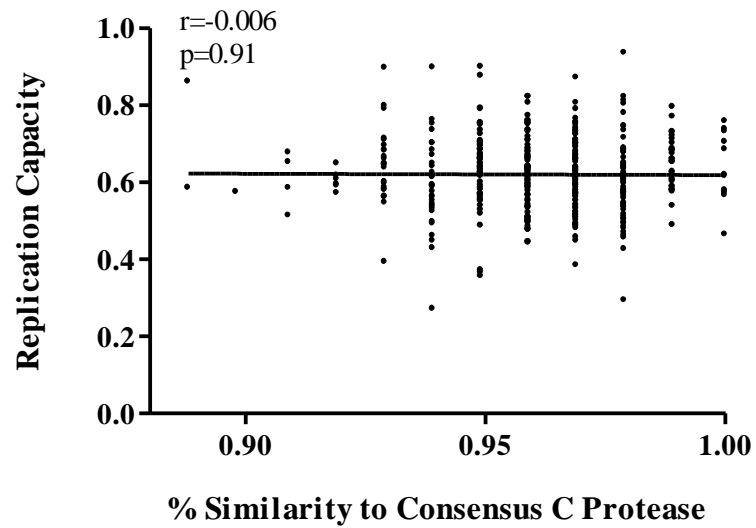
i



ii



iii



The number of HLA-associated polymorphisms specific to the host HLA alleles was computed for each sequence. To further analyse the influence of HLA alleles on Gag-protease replication capacity the computed polymorphisms were then correlated with replication capacity. The total number of HLA-associated polymorphisms in each sequence did not correlate significantly with replication capacity overall (Spearman's correlation; $r=-0.07$, $p=0.13$) (Figure 3.8i). Similarly, the proportion or number of polymorphisms stratified by HLA-A, -B, or -C alleles did not correlate significantly with replication capacity. Likewise, when the relationship between the number of HLA-associated polymorphisms and replication capacity was investigated irrespective of patient HLA class I profile, *i.e.* also taking into account inherited polymorphisms, no significant associations were found. Therefore, the sum of HLA-selected polymorphisms irrespective of location in Gag was not associated with replication capacity in these chronically infected patients.

There was, however, irrespective of host HLA profile, a weak trend (Spearman's correlation; $r=-0.09$, $p=0.08$) towards lower replication capacities with increasing number of HLA-associated polymorphisms (88% HLA-B-associated) in optimal epitopes or within 5 amino acids adjacent to epitopes restricted by the HLA allele associated with the relevant polymorphism. (The A-list of optimal epitopes from the Los Alamos HIV database [275] was used [Appendix, Table A2].) These polymorphisms are more likely to represent escape mutations and not secondarily arising compensatory mutations [207]. Similarly, the number of HLA-B-associated polymorphisms in or within 5 amino acids of Gag epitopes was also negatively correlated with fitness (Spearman's correlation; $r=-0.11$, $p=0.03$) (Figure 3.8ii). Therefore, there was a relatively weak relationship between the number of HLA-associated polymorphisms (irrespective of host HLA profile) in or immediately adjacent to Gag epitopes and replication capacity in the present chronic infection cohort.

Figure 3.8 Associations between HLA-associated polymorphisms and replication capacities of recombinant viruses encoding Gag-protease from HIV-1 subtype C chronically infected patients

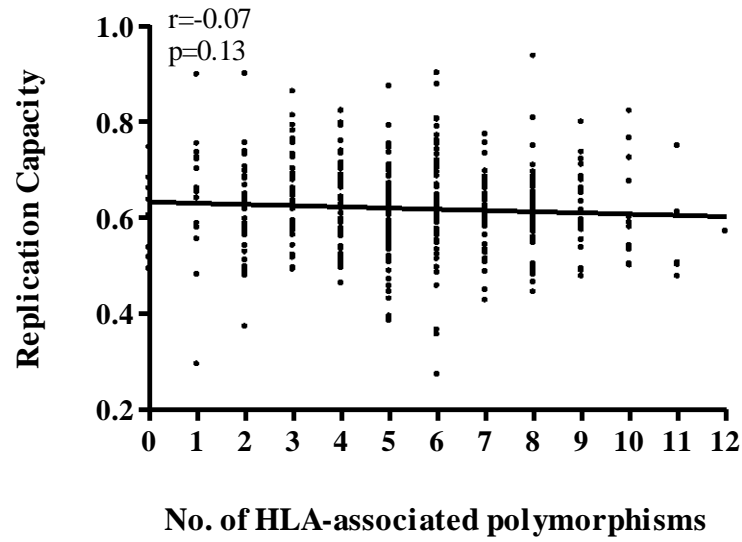
(i) No correlation between the total number of HLA-associated polymorphisms in Gag and replication capacity (Spearman's correlation).

(ii) A weak negative correlation between the number of HLA-B-associated polymorphisms in or within 5 amino acids of Gag epitopes and replication capacity (Spearman's correlation).

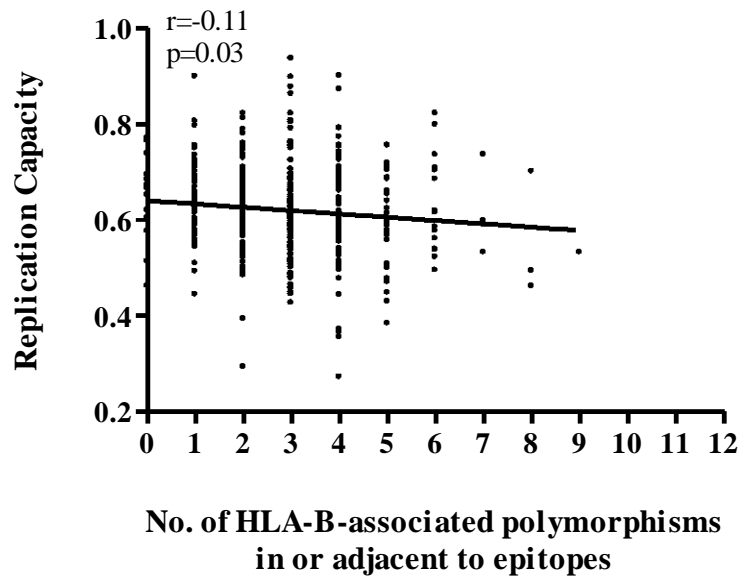
(iii) A significant positive correlation between entropy of HLA-associated sites in or within 5 amino acids of Gag epitopes and the average replication capacity of viruses with HLA-associated polymorphisms at these sites (Pearson's correlation).

HLA – human leukocyte antigen; Gag – group specific antigen; HIV-1 – human immunodeficiency virus type 1; No. - number.

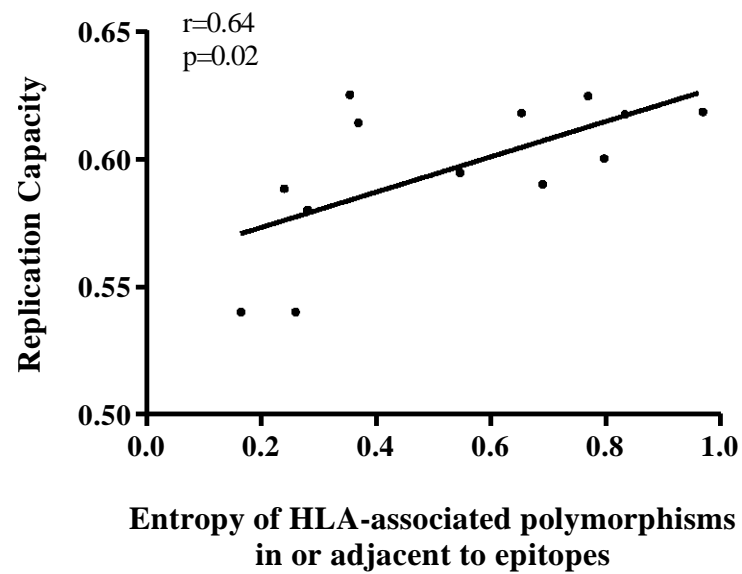
i



ii



iii



Since there is some evidence that HLA-associated escape mutations occurring in conserved sites of HIV carry a greater fitness cost than those occurring in regions of high variability [104] (Section 1.6.5.1), we compared the average replication capacity of viruses possessing each HLA-associated polymorphism with the corresponding entropy (degree of variability) at those positions. A trend towards significant correlation between these parameters was found (Pearson's correlation; $r=0.24$, $p=0.06$). When the analysis was restricted to those polymorphisms in epitopes or within 5 amino acids adjacent to epitopes restricted by the associated HLA allele, the correlation was much stronger (Pearson's correlation; $r=0.64$, $p=0.02$) (Figure 3.8iii). Thus, HLA-associated polymorphisms (likely to be escape mutations) at more conserved sites in Gag were associated with greater fitness costs.

3.2.5.3 Amino acid variants

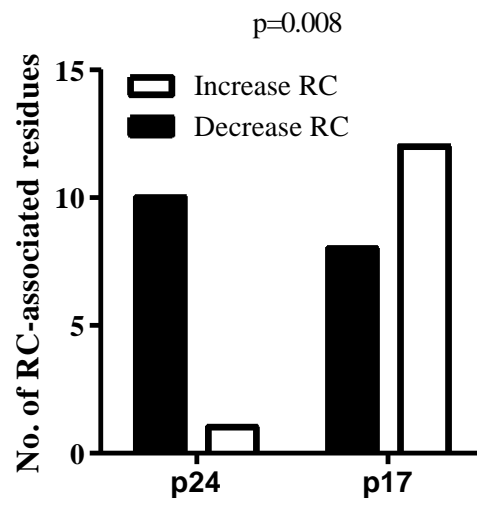
Specific codons at which amino acid variants (occurring at a frequency of $n \geq 5$) associated with significantly increased or decreased replication capacity were identified by univariate analysis (Appendix, Table A3). Although none of the comparisons yielded $q \leq 0.2$, at $p < 0.05$, 68 amino acid variants (including consensus amino acids) at 51 codons were identified in Gag and 12 amino acid variants at 9 codons were identified in protease. Of the 51 codons in Gag, 20 occurred in p17, 11 in p24, 3 in p2, 7 in p7, and 10 in p6.

Considering amino acids in Gag associated with alterations in viral replication capacity, at most of codons in p24 the absence of the consensus amino acid was associated with lower replication capacity (10/11), while in p17 the absence of the consensus amino acid was associated with decreased replication capacity in only 8 out of 20 cases. This difference was statistically significant (Fisher's exact test; $p=0.008$) (Figure 3.9). Nineteen out of 51

Figure 3.9 Association between Gag-protease amino acid variants and altered replication capacities of recombinant viruses encoding Gag-protease from HIV-1 subtype C chronically infected patients

Graph showing a significant difference in the proportion of non-consensus amino acids associated with decreased replication capacity in p24 versus p17 (Fisher's exact test).

Gag – group specific antigen; HIV-1 – human immunodeficiency virus type 1; No. – number; RC – replication capacity; p24 – protein of 24 kDa; p17 – protein of 17 kDa.



codons in Gag corresponded to an HLA-association at that position and 22 of the 51 Gag codons occurred in optimal epitopes. In combination, 29 out of the 51 Gag codons were either HLA-associated or in optimal epitopes.

Multivariate analysis (linear regression with forward selection) was also undertaken. Seventeen out of 51 codons in Gag and 4 out of 9 codons in protease identified by univariate analysis were also $p < 0.05$ (although $q > 0.2$) in the multivariate model and the strongest of these associations was the consensus T at position 186 with increased replication capacity (Appendix, Table A3). The associations with $p < 0.001$ and $q < 0.25$ are shown in Table 3.1.

To increase statistical power, the univariate analysis was repeated limited to HLA-associated positions only. Seven HLA-associated amino acid variants at 6 codons in Gag only were identified at $p < 0.05$ and $q \leq 0.2$. One variant was associated with decreased replication capacity: 186S. Six amino acids were associated with increased replication capacity: 28Q, 186T (consensus), 309A (consensus), 312D (consensus), 381G, and 467E (consensus).

3.2.5.4 HLA-B*81-associated 186S

Since the consensus 186T was strongly associated with increased replication capacity, changes away from consensus at this position were compared between the least fit and fittest viruses. In the least fit group, 9 sequences had an S and 1 had an A at position 186 while only 1 had an S at this position in the fittest group (Fisher's exact test; $p = 0.01$) (Figure 3.10i). In the entire cohort, with the exception of 1 sequence with 186A and another sequence which had a mixture of T and S at this position, the non-consensus amino acid at codon 186 was S. Overall, 186S was associated with a decrease in replication capacity when

Table 3.1. Amino acids in HIV-1 subtype C Gag-protease from chronically infected subjects associated with altered replication capacities by multivariate analysis (p<0.001, q<0.25)

Protein	Codon	AA	Con.	RC	p	q	Optimal epitopes with HLA restriction	HLA association
p24	186	T	Yes	+	0.0001	0.225	TL9 – B*0702, B*3910, B*4201, B*8101, C*0802	B*81 ^a
p17	118	V		-	0.0002	0.225		
p17	62	S		+	0.0002	0.225		B*1402 ^a
p7	389	M		+	0.0003	0.225		
p17	61	I	Yes	+	0.0003	0.225		
p17	28	Q		+	0.0004	0.225	RK9/RLY10 – A*0301, GK9 – B*0801, KW9 – A*2402	B*13 ^a , B*42 ^a , C*17 ^a
pro	64	M		+	0.0007	0.236	RI10 – B*13	
p24	323	V	Yes	-	0.0008	0.236		
p7	411	K	Yes	+	0.0008	0.236	CC9 – B*14	A*0301 ^a

HIV-1 – human immunodeficiency virus type 1; Gag – group specific antigen; p24 – protein of 24 kDa; p17 – protein of 17 kDa; p7 – protein of 7 kDa; pro – protease; AA – amino acid; Con. – consensus; RC – replication capacity; HLA – human leukocyte antigen.

^aHLA allele is negatively associated with the consensus amino acid or positively associated with a polymorphism at that codon.

Figure 3.10 Single HLA-B*81-associated Gag amino acid associations with altered replication capacities of recombinant viruses encoding Gag-protease from HIV-1 subtype C chronically infected patients

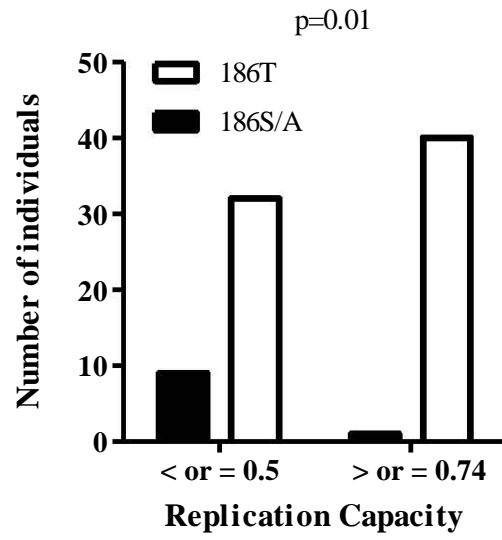
(i) A greater proportion of 186S/A in the least fit viruses (replication capacity ≤ 0.5) compared to the fittest viruses (replication capacity ≥ 0.74) (Fisher's exact test).

(ii) Lower replication capacities of viruses with 186S compared to those with consensus 186T (Student's T-test). Lines represent the mean.

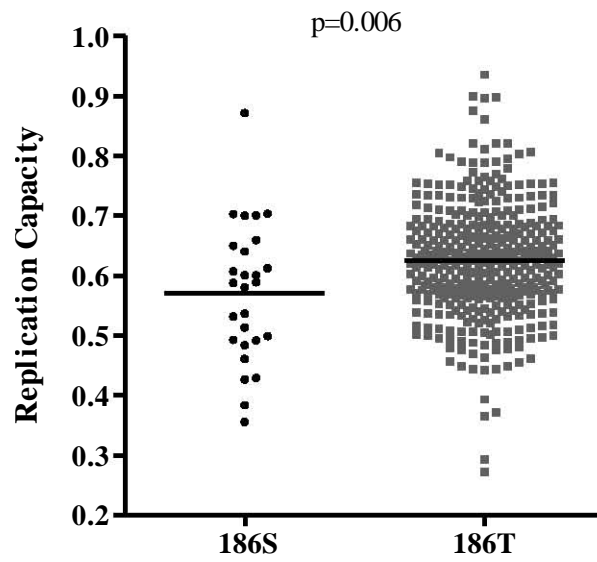
(iii) Significantly higher replication capacities of viruses with 186T and 186S with co-varying mutations 190X and 182S compared to viruses with 186S alone (Student's T-test). Lines represent the mean.

HLA – human leukocyte antigen; Gag – group specific antigen; HIV-1 – human immunodeficiency virus type 1.

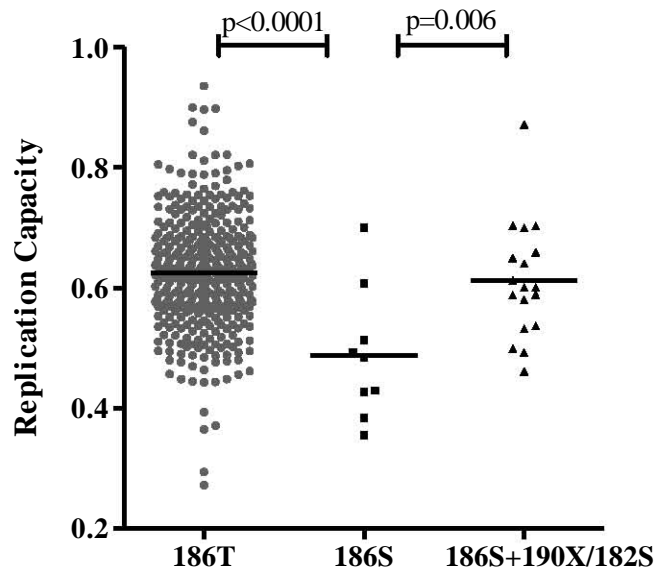
i



ii



iii



compared with consensus 186T (Student's T test; $p=0.006$) (Figure 3.10ii). This polymorphism is associated with HLA-B*81 and occurs in the HLA-B*81-restricted epitope, TL9. However, when only HLA-B*81 positive individuals were considered, the replication capacities of viruses with 186S and 186T were both below average and were not significantly different from one another (Student's T test; $p=0.82$), suggesting that other mutations may also be responsible for the lower fitness of viruses from these individuals and/or that the fitness cost of 186S was compensated for in some cases.

3.2.5.5 Residues co-varying with 186S

Amino acid co-variation lists were generated from the current dataset when testing for HLA-associated amino acids [269] (Section 2.2.5). Amino acids positively associated with 186S and/or negatively associated with 186T include 177D, 182S, 190A, 190I, 256I, and 343I ($p<0.05$, $q\leq 0.2$). Amino acids negatively associated with 186S and/or positively associated with 186T include 65Q, 177E, 190T, 256V, and 343L ($p<0.05$, $q<0.2$). Replication capacities of viruses with 186S and various numbers of associated residues (Q65X, E177X, Q182S, T190X, V256X, and L343X) were compared to assess whether these might be functioning as compensatory mutations. The number of co-varying residues present correlated positively but not significantly with replication capacity (Pearson's correlation; $r=0.26$, $p=0.19$). However, on closer examination of sequences with 186S, a greater occurrence of mutations at positions 182 and 190 (but not at other co-varying positions) was noted in the fitter viruses (Figure 3.10iii). This was statistically significant (Student's T test; $p=0.006$), suggesting that 190X and 182S, which occur parallel to and on either side of residue 186 in a helix structure [276, 277], might be compensatory mutations. Results of experiments testing the fitness effects of these mutations are presented in Section 3.6.

3.3 HIV-1 subtype C recent infection

3.3.1 Distribution of replication capacities

To assess the impact of HIV-1 Gag-protease function on disease progression, recombinant viruses encoding Gag-protease from patients in early HIV-1 subtype C infection were constructed and the replication capacities of these were measured. Replication capacities of recombinant viruses encoding Gag-protease from patients in early infection were normally distributed (Figure 3.11). The mean replication capacity expressed relative to wild-type NL4-3 was 0.63 (SD=0.12), which is comparable to replication capacities of recombinant viruses encoding Gag-protease from chronically infected patients (Section 3.2.1).

3.3.2 Replication capacities versus markers of disease progression

Viral set point and rate of CD4+ T cell decline were used as markers of HIV-1 disease progression in the present study (Section 1.6.1). The average viral load between 3 and 12 months post-infection was the measure of viral set point used in this study, and could be calculated for 56 of 60 (93%) study subjects. No significant correlation between the viral set points and replication capacities of the Gag-protease NL4-3 recombinant viruses was observed (Pearson's correlation; $r=0.12$ and $p=0.37$) (Figure 3.12i). Next, the replication capacities of recombinant viruses encoding Gag-protease from subjects with below median versus above median viral set points were next compared. Patients with below median viral set points tended to have viruses with an attenuated Gag-protease function and this was statistically significant (Student's T test; $p=0.03$) (Figure 3.12ii).

Figure 3.11 Replication capacities of Gag-protease NL4-3 recombinant viruses encoding Gag-protease from HIV-1 subtype C recently infected patients

Histogram showing an approximately normal distribution of replication capacities (normalised to the growth of wild-type NL4-3) with a mean of 0.63 and standard deviation of 0.12.

Gag – group specific antigen; HIV-1 – human immunodeficiency virus type 1.

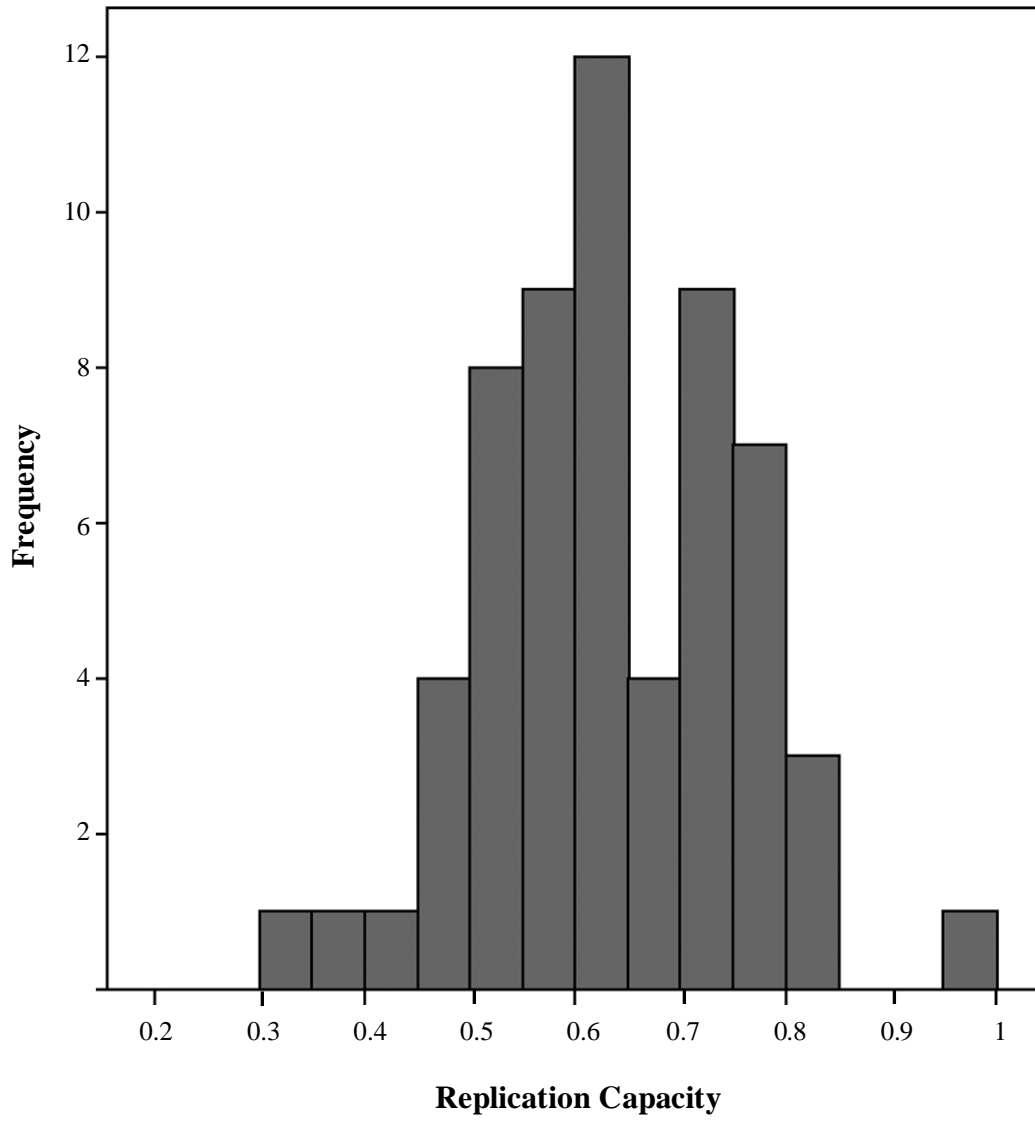


Figure 3.12 Relationships between replication capacities of recombinant viruses encoding patient Gag-protease and markers of disease progression in HIV-1 subtype C recently infected patients

(i) A positive non-significant correlation between replication capacity and viral set point (Pearson's correlation).

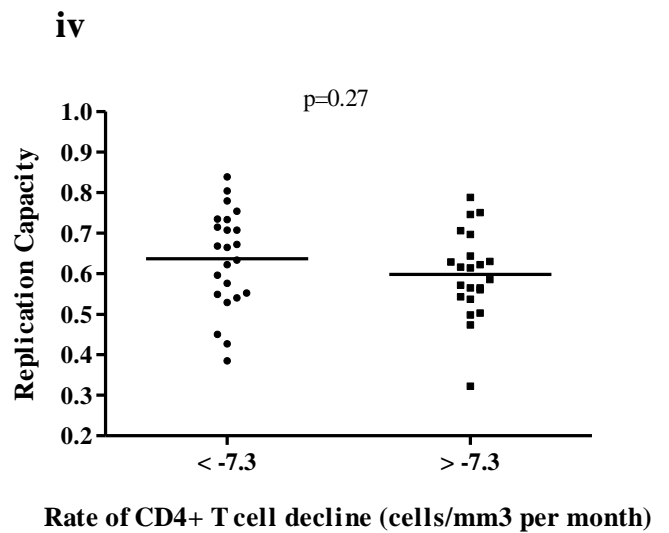
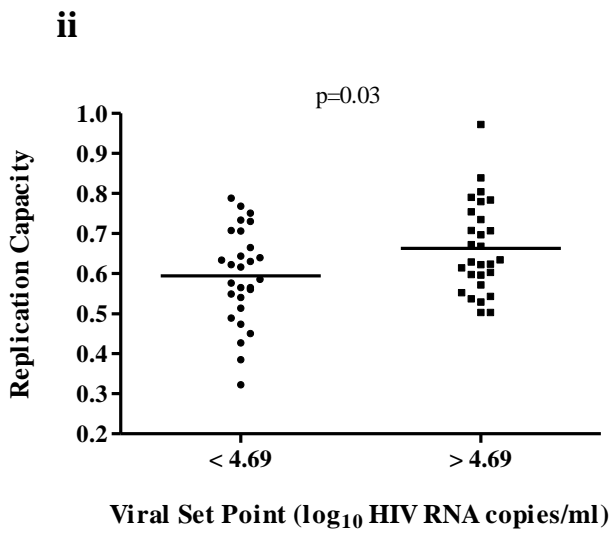
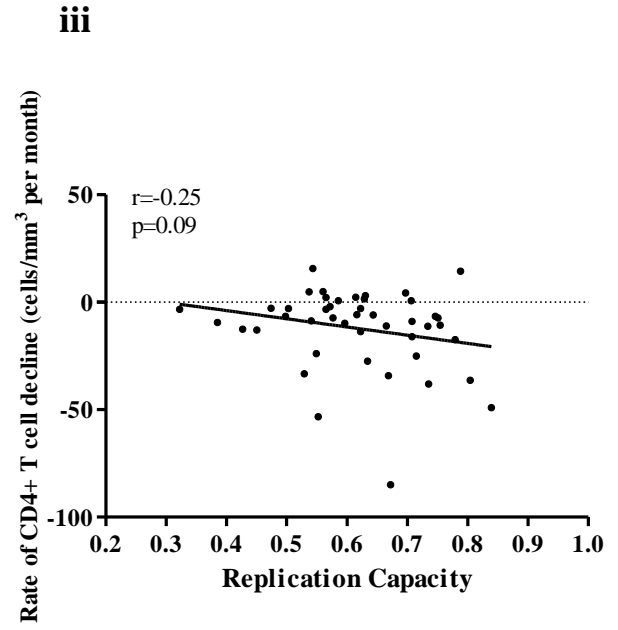
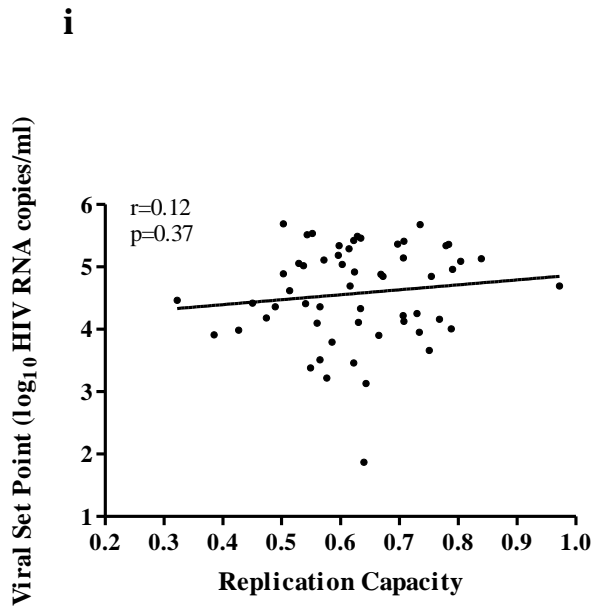
(ii) Significantly lower replication capacities of viruses encoding Gag-protease from patients with below median viral set points versus above median viral load set points (Student's T test). Lines represent the mean.

(iii) A negative non-significant correlation between replication capacity and rate of CD4+ T cell decline for subjects with baseline CD4+ T cell counts >300 cells/mm³ (Spearman's correlation).

(iv) Lower replication capacities of viruses encoding Gag-protease from patients with above median (slower) rates of CD4+ T cell decline versus below median (faster) rates of CD4+ T cell decline for subjects with baseline CD4+ T cell counts >300 cells/mm³ (Student's T test).

Lines represent the mean.

Gag – group specific antigen; HIV-1 – human immunodeficiency virus type 1; RNA – ribonucleic acid; CD4 – cluster of differentiation 4.



A rate of CD4+ T cell decline was then calculated for each subject by linear regression. Since baseline CD4+ T cell count influences rate of decline, when analysing the relationship between rate of CD4+ T cell decline and Gag-protease-mediated replication capacity, subjects with a reasonably moderate to high CD4+ T cell count (>300 cells/mm³, $n=45$) at baseline were considered separately from those with already substantially low CD4+ T cell counts at baseline (<300 cells/mm³, $n=15$). At higher CD4+ T cell counts there was a weak negative trend of association between rates of CD4+ T cell decline and replication capacities, *i.e.* at higher replication capacities the rate of CD4+ T cell decline was increasingly negative or faster, however this trend was not statistically significant (Spearman's correlation; $r=-0.25$ and $p=0.09$) (Figure 3.12iii). The same trend was not shown for subjects with low baseline CD4+ T cell counts which could be due to limited sample size ($n=15$) in this stratum (Pearson's correlation; $r=-0.18$ and $p=0.53$). When replication capacities were grouped according to the median rate of CD4+ T cell decline for subjects with baseline CD4+ T cell count >300 cells/mm³, although the mean replication capacity was lower for subjects with slower rates of decline this was not statistically significant (Student's T test; $p=0.27$) (Figure 3.12iv).

3.3.3 Replication capacities versus protective HLA alleles

Previous studies have reported few sequence changes in Gag in at least the first 6 months of infection [187, 207]. However, for HLA-B*57 positive individuals, associated escape in Gag has been reported as early as 45 days following screening in acute infection [106], although the median reported time in another study is 13 months [187]. Although it is expected that the Gag-protease recombinant viruses were largely representative of transmitted viruses, the possibility of some early selection, particularly by protective alleles, could not be excluded. Therefore replication capacities were compared between protective

and non-protective HLA alleles expressed by the host. There was a trend towards lower replication capacities of recombinant viruses encoding Gag-protease from patients with protective HLA alleles (Student's T test; $p=0.08$) (Figure 3.13), suggestive of some early selection pressure.

It should be noted that for the patients analysed here, viral set points did not differ significantly between individuals with and without protective HLA alleles (4.4 versus 4.61 \log_{10} copies/ml, Student's T test; $p=0.46$). Here a study limitation should also be noted: 12 patients were excluded from this study due to *gag-protease* amplification failure (likely as a consequence of limited available plasma); of these 7 had protective HLA alleles and the median viral set point of excluded patients was significantly lower than those patients included in the study (3 versus 4.69 \log_{10} copies/ml; Student's T test; $p<0.0001$).

3.3.4 Replication capacities/viral set points versus Gag-protease sequences

3.3.4.1 Amino acid variants

Although statistical power was limited by the number of patients, in an exploratory analysis, the Mann-Whitney U test was used to identify specific codons in Gag-protease at which amino acid variants (occurring at a frequency of $n \geq 5$) were associated with increased or decreased replication capacity in early infection. Eighteen codons in Gag (24 associations) and 2 codons in protease (3 associations) with $p < 0.05$ were identified, however, calculated q values were >0.2 in all cases (Table 3.2). The number of patients was not sufficient to perform a multivariate analysis.

Figure 3.13 Relationship between protective HLA alleles and replication capacities of recombinant viruses encoding Gag-protease from HIV-1 subtype C recently infected patients

Lower replication capacities of viruses encoding Gag-protease from patients with protective HLA alleles versus non-protective HLA alleles (Student's T test). Lines represent the mean.

HLA – human leukocyte antigen; Gag – group specific antigen; HIV-1 – human immunodeficiency virus type 1.

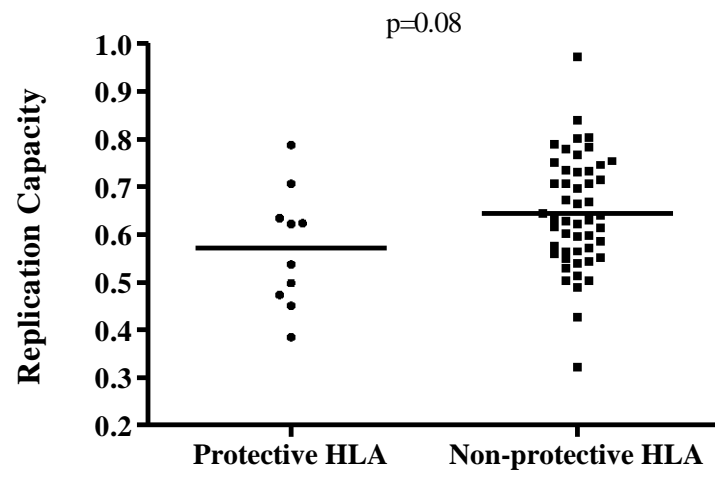


Table 3.2. Amino acids in HIV-1 subtype C Gag-protease from recently infected patients associated with alterations in replication capacity (p<0.05)

Protein	Codon	AA	Con.	No. of samples		Replication capacity		p	q	Optimal epitopes with HLA restriction	HLA association
				-AA	+AA	- AA	+AA				
p17	15	K	Yes	31	29	0.58	0.67	0.012	0.80	GI9 – B*4002	
p17	15	T		49	11	0.63	0.56	0.043	0.80	GI9 – B*4002	
p17	28	H	Yes	37	23	0.67	0.61	0.029	0.80	RK9/RLY10 – A*0301, GK9 – B*0801, KW9 – A*2402	B*13 ^a , B*42 ^a , C*17 ^a
p17	82	V	Yes	6	54	0.53	0.63	0.009	0.80	EV9 – B*0801, RY11 – A*3002, B*58, B*63, SL9 – A*0201, A*0202, A*0205, LL8 – C*14, LY9 – A*2902, B*4403	
p17	82	I		53	7	0.63	0.56	0.010	0.80	EV9 – B*0801, RY11 – A*3002, B*58, B*63, SL9 – A*0201, A*0202, A*0205, LL8 – C*14, LY9 – A*2902, B*4403	
p17	90	K		50	10	0.64	0.54	0.035	0.80	TK8 – A*1101	A34 ^b
p17	95	R	Yes	14	46	0.57	0.64	0.023	0.80	IL10- B*4001	
p17	109	K		53	7	0.62	0.75	0.017	0.80		
p17	110	K		5	55	0.74	0.62	0.017	0.80		
p17	115	T		10	50	0.72	0.62	0.005	0.79		
p17	122	-	Yes	9	51	0.70	0.62	0.035	0.80		
p24	163	G		52	8	0.63	0.54	0.038	0.80	VF9 – B*1503, EV9 – B*4415, KI8 – B*5703, KF11 – B*5701, B*5703, B*63	B*5703 ^a
p24	223	I	Yes	41	19	0.67	0.57	0.011	0.80	HA9 – B*07	A*2911 ^b
p24	242	T	Yes	18	42	0.59	0.67	0.014	0.80	TW10 – B*5701, B*5801	B57 ^a , B*5801 ^a
p24	242	N		43	17	0.67	0.60	0.030	0.80	TW10 – B*5701, B*5801	B57 ^a , B*5801 ^a
p24	248	T		53	7	0.62	0.75	0.044	0.80	TW10 – B*5701, B*5801	
p24	256	V		42	18	0.64	0.58	0.050	0.80	PY9 – B*3501	C18 ^a
p24	256	I	Yes	18	42	0.58	0.64	0.050	0.80	PY9 – B*3501	C18 ^a
p24	323	V	Yes	6	54	0.76	0.62	0.038	0.80		
p2	374	T	Yes	43	17	0.63	0.56	0.031	0.80		B*1503 ^a
p2	377	L		50	10	0.64	0.53	0.001	0.42		B*5802 ^b
p2	377	M	Yes	9	51	0.53	0.64	0.001	0.42		B*5802 ^b
p7	389	I	Yes	25	35	0.67	0.59	0.019	0.80		
p7	389	T		47	13	0.62	0.71	0.038	0.80		
Pro	15	I		51	9	0.64	0.54	0.005	0.28		
Pro	15	V	Yes	8	52	0.55	0.64	0.011	0.28		
Pro	19	T		49	11	0.64	0.55	0.015	0.28		

Gag – group specific antigen; p17/24/7/6 – protein of 17/24/7/6 kDa; Pro – protease; AA – amino acid; con. – consensus; No. – number. HLA – human leukocyte antigen; ^aHLA allele is negatively associated with the consensus amino acid or positively associated with a polymorphism at that codon. ^bHLA allele is positively associated with the consensus amino acid or negatively associated with a polymorphism at that codon.

Eight of the associations identified (all in Gag) corresponded to HLA-associated codons (28H, 90K, 163G, 223I, 242N, 256V, 374T, and 377L) and a further 4, although not HLA-associated, were situated in Gag epitopes (15T, 82I, 95R, and 248T) (Table 3.2). Limiting the analysis to HLA-associated positions only, one association achieved statistical significance at $q < 0.2$, namely 377L was significantly associated with lower replication capacities (Mann-Whitney U Test; $p = 0.001$ and $q = 0.07$). This polymorphism was negatively associated with HLA-B*5802 in the subtype C chronically infected cohort - it did not occur in any of the 68 HLA-B*5802 positive individuals but was present in 32 out of 316 HLA-B*5802 negative individuals. The selection or maintenance of the consensus 377M by HLA-B*5802 which is associated with higher replication capacities could possibly be one of the contributing factors to the detrimental effect of this HLA allele in HIV-1-infected individuals [170, 176].

Two polymorphisms associated with protective HLA alleles, namely 242N (HLA-B*5801/B*57) and 163G (HLA-B*57), were associated with lower replication capacities in the present study, which is consistent with previous reports [33, 187, 211] (Section 1.6.5.1). Interestingly, only 2 of the 18 individuals that had a virus with 242X (17=242N and 1=242S) were HLA-B*5801 positive and none of the 8 with 163G were HLA-B*57 positive, indicating that these mutations were transmitted to, rather than selected in, the study subjects in nearly all cases. Consistent with previous reports [251], 242X-expressing viruses harbouring 2 or more 242X-associated compensatory mutations – 219Q, 223V, 228I, and 248T [33] - displayed higher replication capacities (0.67, $n = 3$) than those with 1 or 0 (0.55, $n = 15$; Mann-Whitney U test; $p = 0.04$) and 248T was individually associated with higher replication capacities (Table 3.2).

3.3.4.2 Number of HLA-associated polymorphisms in Gag

3.3.4.2.1 Replication capacity

Next the relationship between the number of HLA-associated polymorphisms in Gag and replication capacity was investigated. Each sequence was analysed firstly in the context of the patient's HLA alleles (the absolute number as well as the proportion of polymorphisms associated with host alleles were computed) and secondly irrespective of the host HLA profile. The overall number of HLA-associated polymorphisms in Gag did not correlate with replication capacity. However, when limited to HLA-associated polymorphisms occurring in or within 5 amino acids of optimal epitopes (likely representing escape), there was a significant inverse correlation between the total proportion of host-specific HLA-A, B and C polymorphisms and replication capacity (Spearman's correlation; $r=-0.32$ and $p=0.04$) (Figure 3.14i), which appeared to be driven mainly by HLA-B-associated polymorphisms. Similarly, when all polymorphisms were included irrespective of host HLA profile, there was also a trend towards lower replication capacities with an increasing absolute number of HLA-B-associated Gag polymorphisms in or adjacent to epitopes (Spearman's correlation; $r=-0.25$ and $p=0.06$, one extreme outlier removed). In summary, overall there was a negative relationship between the number of Gag polymorphisms and viral replication capacity.

Figure 3.14 Associations between HLA-associated Gag polymorphisms, replication capacities of recombinant viruses encoding patient Gag-protease, and viral set points in HIV-1 subtype C recently infected patients

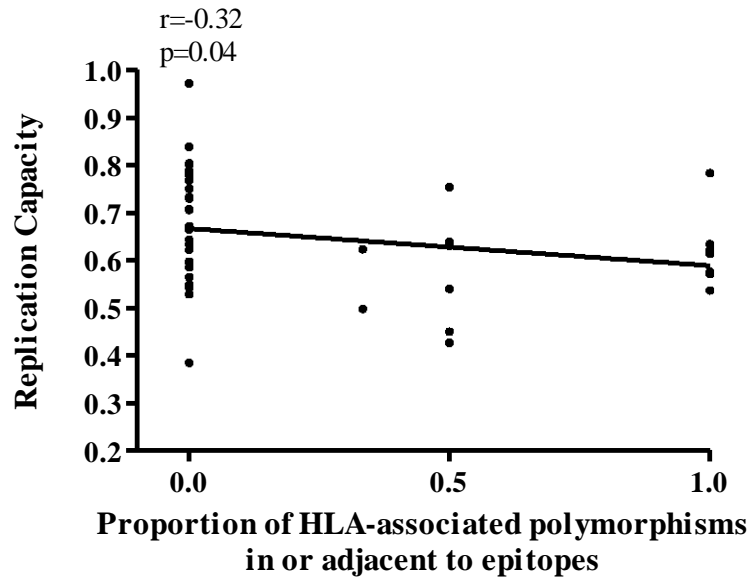
(i) A significant negative correlation between the proportion of host-specific HLA-associated polymorphisms in or adjacent to Gag epitopes and the replication capacities of recombinant viruses encoding Gag-protease.

(ii) A significant positive correlation between the proportion of host-specific HLA-B-associated polymorphisms present in Gag and viral set point.

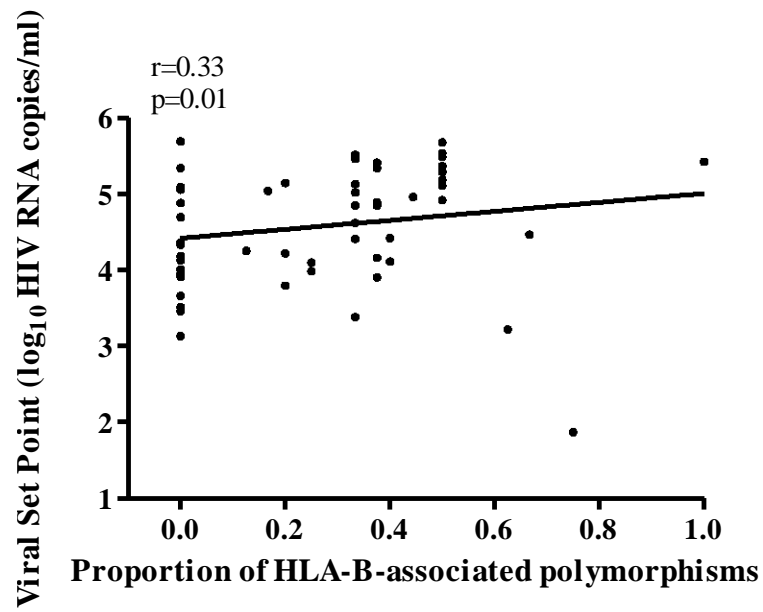
(iii) Significantly higher viral set points in patients with protective HLA alleles that have a polymorphism associated with those alleles in Gag versus those that do not (Mann-Whitney U test). Lines represent the mean.

HLA – human leukocyte antigen; Gag – group specific antigen; HIV-1 – human immunodeficiency virus type 1.

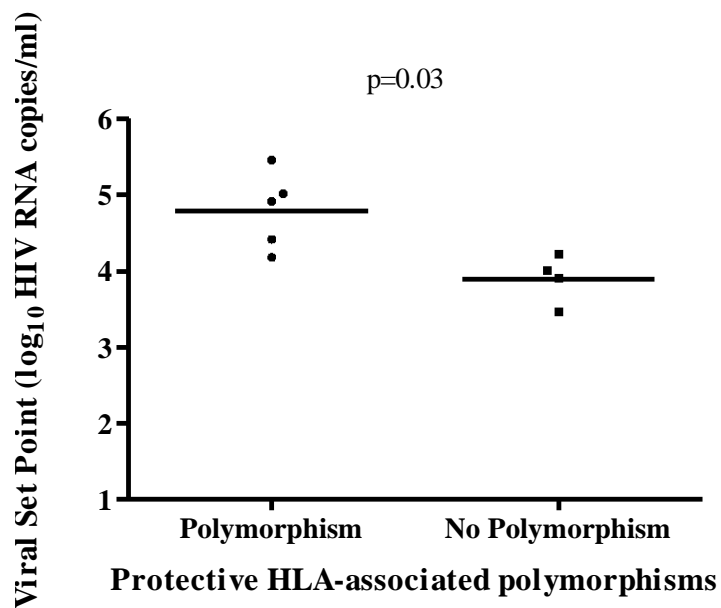
i



ii



iii



3.3.4.2.2 Viral set point

Next sequences were analysed for a correlation between Gag polymorphisms and viral set point, as has been previously described [207]. Unexpectedly, the proportion of total host-specific HLA-B-associated polymorphisms and presence of host-specific protective HLA-associated polymorphisms were both positively associated with viral set point (Spearman's correlation; $r=0.33$ and $p=0.01$; and Mann-Whitney U test; $p=0.03$) (Figure 3.14ii,iii). Similar results were calculated for absolute numbers of host-specific HLA-B-associated polymorphisms. Thus, the presence of Gag polymorphisms that could be selected by host alleles was generally associated with higher viral set points.

Since Goepfert *et al.* (2008) showed that the negative association between the number of HLA-B-associated polymorphisms in transmitted viruses and viral set point in recipients was stronger when host-specific HLA-associated polymorphisms were excluded (*i.e.* there was benefit associated with inheriting polymorphisms that were not associated with HLA alleles of the recipient) [207], the analysis was repeated the same way here using the same published list of polymorphisms. This analysis showed no significant relationship between transmitted HLA-B-associated polymorphisms and viral set points (Pearson's correlation; $r=-0.1$ and $p=0.43$) in contrast to positive significant relationships observed in earlier analyses. Also of interest, 5 of 12 subjects for whom *gag-protease* amplification failed here and who had very low viral set points overall, had viruses that harboured protective HLA-associated polymorphisms that they could not have selected based on their HLA profiles; and only 1 of the 7 with protective HLA alleles had an associated polymorphism. (Note that extraction and sequencing was performed on these subjects previously [278].) These data are not inconsistent with a benefit to inheriting Gag polymorphisms that could not be selected by host HLA alleles.

3.3.5 Longitudinal analyses of replication capacities, CD8+ T cell responses and clinical parameters

For 13 recently infected subjects, Gag-protease NL4-3 recombinant viruses were additionally generated and assayed from time points approximately 1 year later than the first available time points following screening. Overall, paired replication capacity measurements were not found to be significantly different from one another (paired T test; $p=0.77$). In some samples where the replication capacity changed more than 0.1, sequence changes in the Gag protein could partly explain the alteration. For example, in patient AS2-0945 (HLA-B*5703 positive) fitness-reducing mutations 242X (mixture of T, S, and N at this residue) and 146P were selected by 1 year later with a substantial decrease in replication capacity. Also, in patient AS2-0802 a CD8+ T cell response to the B*08-restricted epitope EV9 was present at the early time point and the mutation 75I had developed in this epitope 1 year later. The mutation 75I was significantly associated with higher replication capacities in the subtype C chronically infected cohort (Appendix, Table A3) and accordingly viral replication capacity increased for this patient.

Viral replication capacities and CD8+ T cell responses to Gag epitopes were next related to clinical information. Of interest, in patient AS2-0341, replication capacity remained low but viral set point and viral load at the later time point were high (4.89 \log_{10} copies/ml and 5.44 \log_{10} copies/ml respectively). This patient made no detectable CD8+ T cell responses to Gag epitopes. In contrast, replication capacity remained low in AS2-0174 but this patient mounted CD8+ T cell responses to 9 Gag epitopes and viral load remained below 3.15 \log_{10} copies/ml from 8 months post-infection onwards (*i.e.* patient AS2-0174 was a controller). In patient AS2-0945, in whom viral set point was low (3.46 \log_{10} copies/ml), a detectable CD8+ T cell response was made to a key epitope in Gag (TW10) and mutations were

selected that decreased replication capacity. However, it should be noted that there was a rise in viremia in this patient coinciding with selection of these mutations. In patients AS2-0358 and AS2-1037, although CD8⁺ T cell responses were made to Gag epitopes, viremia was not well-controlled and Gag-protease-mediated replication capacities were also high in these patients.

The viral replication capacities, together with CD8⁺ T cell responses to Gag epitopes, clinical parameters and patient HLA types, are presented in Table 3.3.

3.4 Contribution of Gag-protease to overall replication capacity

To test whether HIV-1 Gag-protease function significantly contributes to viral replication capacity, HIV-1 was isolated from patients and recombinant viruses encoding Gag-protease from these patients were also generated. The replication capacities of the isolates and the Gag-protease recombinant viruses were then compared. A moderate statistically significant correlation was found between the replication capacities of the isolates and their corresponding recombinant viruses (Pearson's correlation; $r=0.51$ and $p=0.04$) (Figure 3.15). This indicates that Gag-protease is a significant determinant of viral fitness.

Table 3.3. Longitudinal measurements of Gag-protease-mediated viral replication capacities, Gag CD8+ T cell responses, and clinical measures from patients recently infected with HIV-1 subtype C (including patient HLA profiles)

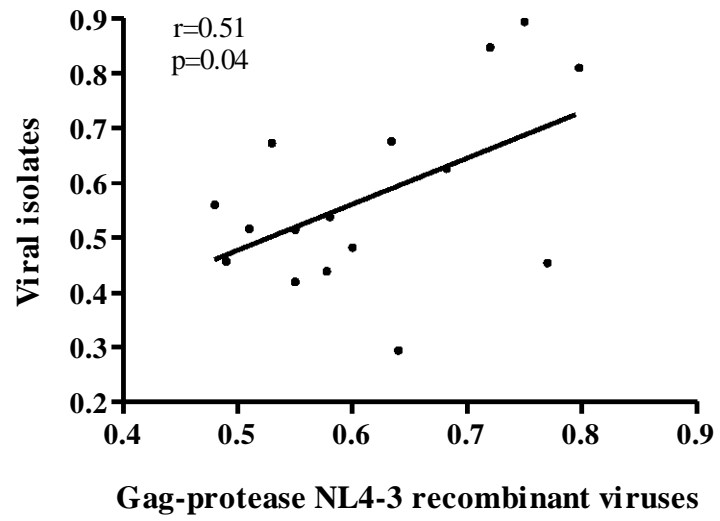
ID	RC		VL		CD4		HLA type	CD8+ T cell responses	
	A	B	A	B	A	B		Early	Median of 6 weeks later
AS1-0703	0.6	0.58	5.15	3.88	625	549	A*2601 A*3004 B*4403 B*5802 C*0210 C*0602	AW11 – B*44	AW11 – B*44, QW9 – B*57
AS1-0919	0.75	0.71	4.65	4.43	665	383	A*2902 A*7400 B*3501 B*4201 C*0401 C*1700	TL9 – B*42	
AS2-0016	0.61	0.7	4.55	3.96	528	537	A*2601 A*3002 B*1518 B*4201 C*1700 C*1800	RLY10 – A*30	RLY10 – A*30, EL9 – A*26, TL9 – B*42
AS2-0174	0.57	0.58	5.47	2.01	361	416	A*2301 A*3002 B*0801 B*1402 C*0304 C*0802		RLY10 – A*30, RY11 – A*30, AF8 – A*30, DI8 – B*08, EL8 – B*08, DA9 – B*14, CC9 – B*14, YL9 – C*03, TL9 – C*08
AS2-0341	0.5	0.43	5.86	5.44	553	346	A*3402 A*6801 B*1503 B*5802 C*0210 C*0602		
AS2-0358	0.97	0.82	5.18	5.34	331	553	A*2402 A*3001 B*4202 B*5301 C*0404 C*1700	HW9 – A*23	HW9 – A*23, SW11 – A*30, QW9 – B*53
AS2-0483	0.7	0.54	5.54	4.65	359	440	A*3001 A*3402 B*1503 B*4201 C*0210 C*1700		RY11 – A*30, TL9 – B*42
AS2-0802	0.32	0.65	5.11	4.27	316	338	A*3002 A*4301 B*0801 B*5802 C*0602 C*0701	RY11 – A*30, EV9 – B*08, KF11 – B*57	RY11 – A*30, DI8 – B*08, EL8 – B*08
AS2-0945	0.62	0.46	3.66	4.51	686	648	A*2301 A*7400 B*1510 B*5703 C*0701 C*1601	TW10 – B*57	QW9 – B*57, GI9 – B*1510
AS2-1037	0.71	0.73	5.93	4.82	403	219	A*03 A*3402 B*1503 B*1510 C*0210 C*0304	VL10 – B*1510, YL9 – C*03	YL9 – C*03
AS3-0268	0.54	0.56	5.88	5.64	483	419	A*2902 A*8001 B*1503 B*1801 C*0202 C*0210	RY11 – A*29, LY9 – A*29, VF9 – B*1503	LY9 – A*29, VF9 – B*1503
AS3-0369	0.59	0.63	4.55	4.4	485	436	A*6602 A*6802 B*1510 B*4403 C*0304 C*0701	RL11 – B*44, AW11 – B*44, YL9 – C*03	YL9 – C*03, VL10 – B*1510
AS3-0458	0.43	0.69	4.27	4.21	488	455	A*2902 A*3402 B*4403 B*5802 C*0401 C*0602	AW11 – B*44	AW11 – B*44, RY11 – A*29

Gag – group specific antigen; CD8 – cluster of differentiation 8; HIV-1 – human immunodeficiency virus type 1; ID – patient code; A – early time point; B – time point a median of 1 year later than A; RC – replication capacity; VL – viral load; CD4 – cluster of differentiation (CD)4+ T cell count; HLA – human leukocyte antigen.

Figure 3.15 Contribution of Gag-protease to overall HIV-1 fitness

A significant positive correlation between replication capacities of HIV-1 subtype C isolated from patients and corresponding NL4-3 recombinant viruses encoding Gag-protease derived from the same patients (Pearson's correlation).

Gag – group specific antigen; HIV-1 – human immunodeficiency virus type 1.



3.5 HIV-1 subtype B versus subtype C

3.5.1 Gag-protease NL4-3 recombinant viruses

The replication capacities of the 25 recombinant viruses encoding Gag-protease from patients chronically infected with HIV-1 subtype B were normally distributed, with a mean of 0.95 (SD=0.14) (Figure 3.16i). The replication capacities of the 25 subtype B recombinants were significantly greater than that of the 406 subtype C recombinant viruses constructed from chronically infected patients (Student's T test; $p < 0.0001$) (Figure 3.16ii). Consistent with this difference, the median time to generate high titre virus stocks of subtype B recombinant viruses was 12 days (IQR: 12-13 days) compared with a median of 27 days (IQR: 23-32 days) to generate subtype C recombinant viruses.

Since viral replication capacity may be related to stage of disease [247, 251], a subset ($n=25$) of HIV-1 subtype C-infected patients were matched with subtype B-infected patients in terms of CD4+ T cell counts and viral loads and the analysis repeated. The replication capacities of this subset (average replication capacity=0.65; SD=0.08) were still significantly lower than the replication capacities of the subtype B viruses (Mann-Whitney U-test; $p < 0.0001$) (Figure 3.16iii).

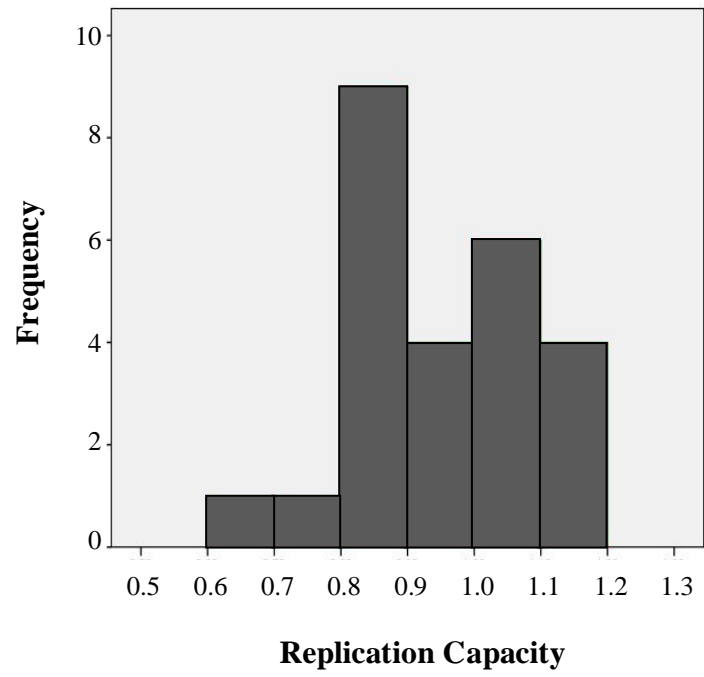
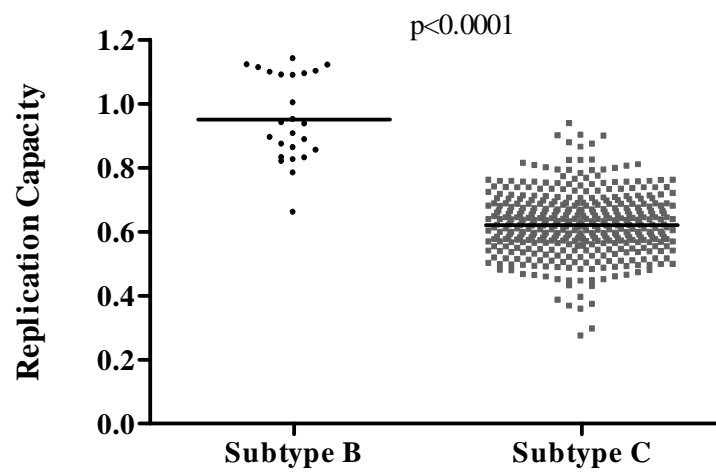
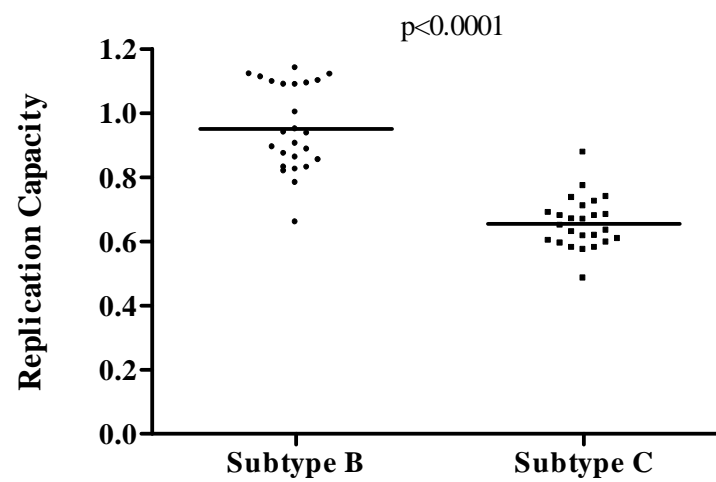
Figure 3.16 Comparison of replication capacities of recombinant viruses encoding subtype B and C Gag-protease from patients chronically infected with HIV-1

(i) Histogram showing an approximately normal distribution of subtype B Gag-protease recombinant virus replication capacities (normalised to the growth of wild-type NL4-3) with a mean of 0.95 and standard deviation of 0.14.

(ii) Significant difference between HIV-1 subtype B and C Gag-protease-mediated replication capacity (Student's T test). Lines represent the mean.

(iii) Significant difference in HIV-1 Gag-Protease-mediated replication capacity between subtype B and subtype C samples matched for viral load and CD4+ T cell count (Student's T test). Lines represent the mean.

Gag – group specific antigen; HIV-1 – human immunodeficiency virus type 1; CD4 – cluster of differentiation 4.

i**ii****iii**

3.5.2 Subtype-specific Gag mutations

Specific single amino acids in subtype B [251] and C (Appendix, Table A3) Gag-protease significantly associated with viral fitness differences have been identified. Residues 67A, 473P and 483L in subtype B Gag were strongly and significantly associated with altered replication capacities ($p < 0.0001$ and $q < 0.0001$) [251]. Similarly, 473P in subtype C Gag was significantly associated with increased replication capacity (Appendix, Table A3). The consensus amino acids at these codons differ between subtype B and C, which therefore might partly explain the observed Gag-protease-mediated fitness difference between the subtypes. In HIV-1 subtype C, 67A, which is strongly associated with decreased fitness in subtype B viruses, is the consensus amino acid while in subtype B viruses 67S is the consensus. In subtype B viruses, 483L, which is strongly associated with increased fitness, is the consensus amino acid, while there is a deletion at positions 483 and 484 in subtype C viruses. In both subtype B and C viruses, 473P is associated with increased fitness however the frequency of this amino acid at position 473 is 1.7% in the subtype C chronically infected cohort of the present study and 39% in a subtype B-infected cohort [251].

The fitness effects of these amino acid differences between subtypes B and C were evaluated by site-directed mutagenesis. The insertion of amino acids L and Y at positions 483 and 484 in subtype C Gag-protease resulted in significantly increased fitness when compared to the wild-type or 67S and 473P mutants (ANOVA, Tukey post-hoc tests; $p < 0.01$) (Figure 3.17i). Accordingly, the deletion of these amino acids in the subtype B virus NL4-3 resulted in decreased replication capacity compared with other mutants and the wild-type (ANOVA, Tukey post-hoc tests; $p < 0.05$) (Figure 3.17ii). However, alterations at residues 67 and 473 did not appreciably alter fitness in either backbone (Figure 3.17i,ii).

Figure 3.17 Replication capacities of HIV-1 encoding subtype-specific mutations in Gag

Differences in the replication capacities of:

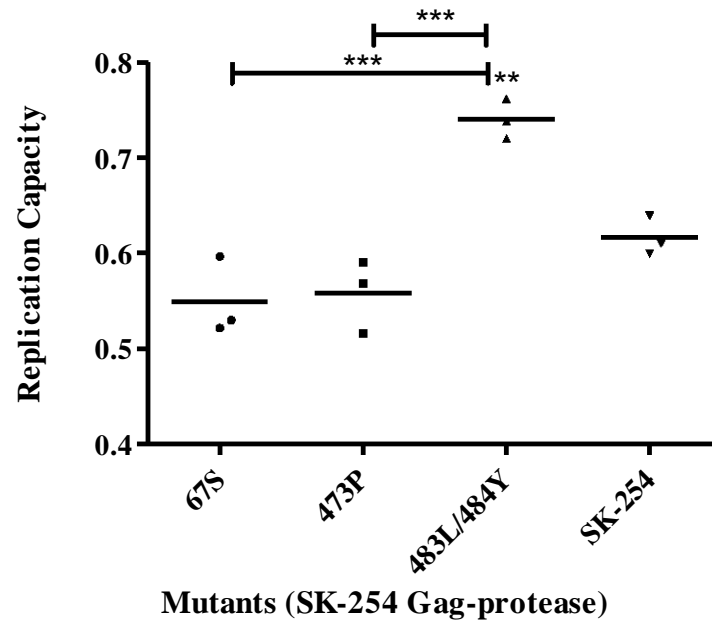
(i) NL4-3 recombinant viruses encoding wild-type HIV-1 subtype C Gag-protease derived from patient SK-254 and SK-254 Gag-protease with introduced mutations 67S, 473P, and 483L/484Y; and

(ii) NL4-3 wild-type virus and NL4-3 viruses encoding Gag mutations 67S, 473A, and 483-/484-.

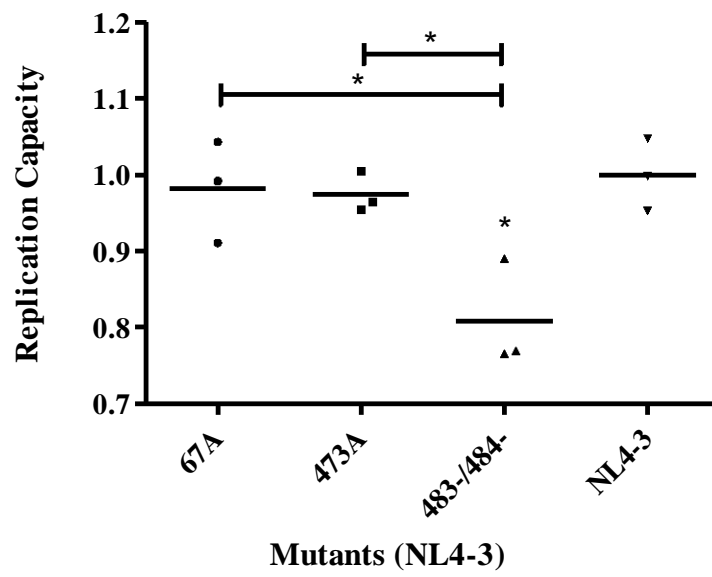
ANOVA with Tukey post-hoc tests was used to test for significant differences in replication capacities between wild-type and mutant viruses. Asterisks above mutant groups indicate significantly altered replication capacity when compared with the wild-type group, while asterisks with a bar indicate significant differences between various mutant groups. The number of asterisks denotes the level of significance, namely, $p < 0.05$ (*), $p < 0.01$ (**) and $p < 0.001$ (**).

HIV-1 – human immunodeficiency virus type 1; Gag – group specific antigen; SK – Sinikithemba; ANOVA – analysis of variance.

i



ii



3.6 HLA-B*81-associated mutations

Since recombinant viruses encoding HIV-1 subtype C Gag-protease from individuals with the protective allele HLA-B*81 had significantly reduced fitness (Section 3.2.2), and the B*81-associated mutation T186S in the TL9 epitope was strongly associated with reduced replication capacity (Section 3.2.5.4), mutagenesis was performed to confirm the fitness cost of T186S. In addition, since viruses with T186S and either Q182S or T190X (also HLA-B*81-associated mutations) were significantly fitter than those with T186S alone (Section 3.2.5.5), suggesting that they might function as compensatory mutations, these mutations were introduced alone and in combination into a subtype C Gag-protease (SK-254). The location of these residues in the Gag capsid is shown in Figure 3.18. The generation of the following mutants failed: 186S, 186S/182S, and 186S/190A. The remaining mutants – 182S, 190A, 190I, and 186S/190I – were successfully generated and their replication capacities were lower than that of the wild-type virus, although only the 182S and 186S/190I mutants were significantly less fit than the wild-type virus (ANOVA, Tukey post-hoc tests; $p < 0.01$) (Figure 3.19i).

The failure to generate mutant viruses containing 186S, with the exception of 186S/190I, was not inconsistent with the hypothesised fitness cost of 186S. However, to further clarify the role of these mutations in viral fitness, they were introduced into wild-type NL4-3 since NL4-3 is a highly fit laboratory strain of HIV-1, thereby presumably increasing the chances of successful generation of all mutant viruses. All mutants were successfully generated, however, the mutant 186S/190I virus was obtained by a modified electroporation protocol (Section 2.2.2.4). All mutants, except 182S, were significantly less fit than wild-type NL4-3 and the mutant 182S was significantly fitter than mutants 186S, 190I, 186S/182S, and 186S/190I (ANOVA, Tukey post-hoc tests; $p < 0.05$) (Figure 3.19ii). In contrast, the 182S

Figure 3.18 Structure of the HIV-1 Gag p24 capsid

Gag residues 182, 186, and 190 are highlighted in yellow. Helices are represented by cylinders in green, strands are represented by arrows in brown and coils are shown in blue.

Modified from structure 1GWP (Tang (2002) [276]) viewed in Cn3D (<http://www.ncbi.nlm.nih.gov/Structure/CN3D> [277]).

HIV-1 – human immunodeficiency virus type 1; Gag – group specific antigen; p24 – protein of 24 kDa.

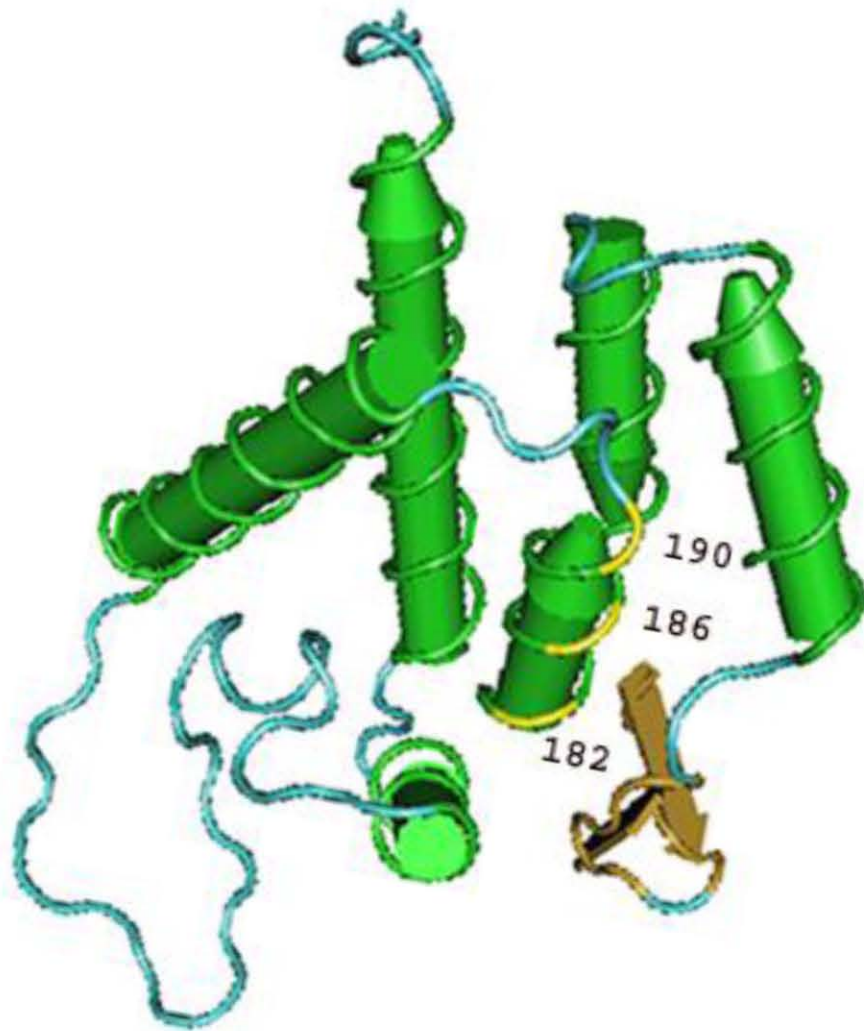


Figure 3.19 Replication capacities of HIV-1 encoding HLA-B*81-associated mutations in Gag

Differences in the replication capacities of:

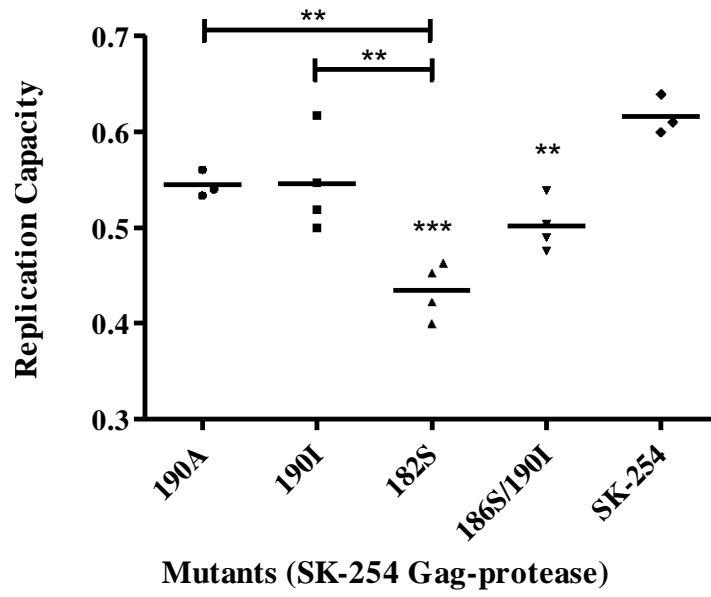
(i) NL4-3 recombinant viruses encoding wild-type HIV-1 subtype C Gag-protease derived from patient SK-254 and SK-254 Gag-protease with introduced mutations 190A, 190I, 182S, and 186S/190I; and

(ii) NL4-3 wild-type virus and NL4-3 viruses encoding Gag mutations 186S, 182S, 190A, 190I, 186S/182S, 186S/190A, and 186S/190I.

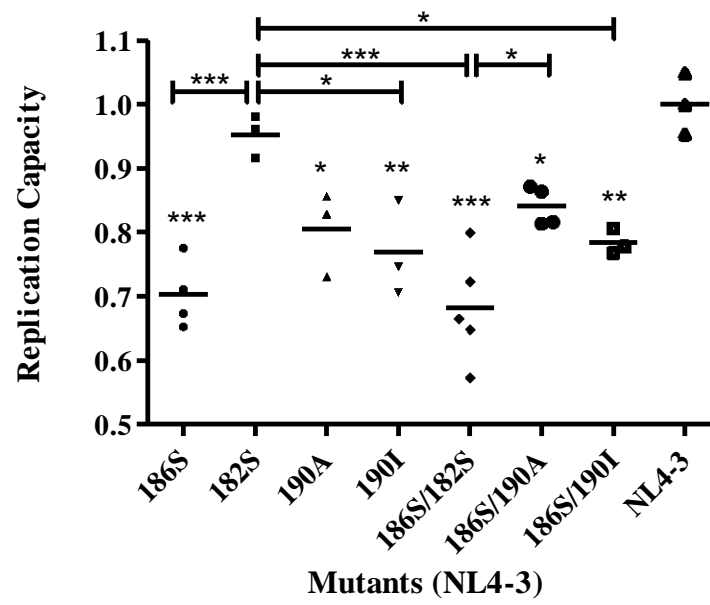
ANOVA with Tukey post-hoc tests was used to test for significant differences in replication capacities between wild-type and mutant viruses. Asterisks above mutant groups indicate significantly altered replication capacity when compared with the wild-type group, while asterisks with a bar indicate significant differences between various mutant groups. The number of asterisks denotes the level of significance, namely, $p < 0.05$ (*), $p < 0.01$ (**) and $p < 0.001$ (***).

HIV-1 – human immunodeficiency virus type 1; HLA – human leukocyte antigen; Gag – group specific antigen; SK – Sinikithemba; ANOVA – analysis of variance.

i



ii

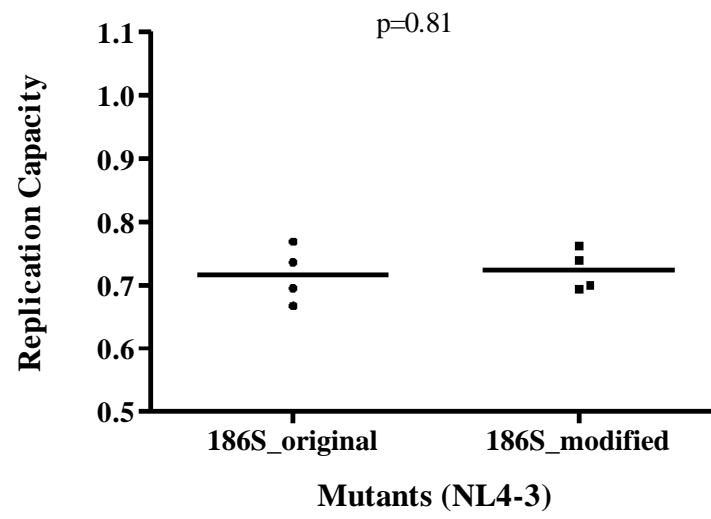


mutation in SK-254 Gag-protease significantly decreased fitness (Figure 3.19i). NL4-3 mutants 186S and 186S/182S were less fit than mutants 190I, 190A, 186S/190A, and 186S/190I, although this was only statistically significant when comparing mutants 186S/190A and 186S/182S (ANOVA, Tukey post-hoc tests; $p < 0.05$) (Figure 3.19ii). This supports a fitness cost associated with 186S and suggests that 190A rather than 182S may partially, although weakly, compensate for the fitness cost of 186S. Overall, none of the mutations, co-varying with 186S, that were tested here significantly or strongly compensated for the fitness cost of 186S.

Since a modified electroporation protocol was used to make a small minority (3) of the mutants/recombinant viruses in the latter stages of the present project, the fitness of 186S mutants that were made by both the original and modified protocols were compared, and found to be the same (Student's T test; $p = 0.81$) (Figure 3.20). This indicates that the use of the modified protocol did not introduce significant variability into the results.

Figure 3.20 Replication capacities of 186S mutants generated by the original and modified electroporation protocols

Graph showing no significant difference between the replication capacities of 186S mutant viruses generated by two different electroporation protocols (Student's T test).



CHAPTER 4

DISCUSSION

CHAPTER 4 – DISCUSSION

4.1 Impact of HLA on HIV-1 replication capacity

HLA class I profile is the strongest genetic correlate of differential rates of disease progression in HIV-1-infected individuals; however the mechanisms underlying HIV-1 control by protective HLA alleles are not fully understood. These mechanisms could involve the targeting of functionally important epitopes in HIV-1 Gag, resulting in selection of escape mutations that significantly attenuate HIV-1 and slow disease progression. Therefore, the impact at the population level of HLA-mediated immune pressure in Gag on HIV-1 replication capacity and pathogenesis was investigated in the present study.

4.1.1 Protective HLA alleles: HLA-B*81

The results showed an association between protective HLA alleles (HLA-B*57, HLA-B*5801, and HLA-B*81), especially HLA-B*81, and lower Gag-protease-mediated replication capacities in HIV-1 subtype C chronic infection. Since, (i) protective HLA alleles were associated with lower viral loads, and (ii) Gag-protease-mediated replication capacity correlated with viral loads even on removal of protective HLA alleles from the analysis and within individuals with protective alleles, the possibility that HLA alleles and replication capacity were indirectly related to each other through association with viral load cannot be excluded. Arguing against this possibility, HLA-B*81 was by far the most strongly associated with lower replication capacity even though HLA-B*5703 positive individuals had a lower average viral load than HLA-B*81 positive individuals. Furthermore, several observations from the present study as well as previous studies [33, 187, 211-213, 229, 251] are consistent with a direct influence of HLA alleles on viral replication capacity. First, in the present study, Gag-protease-mediated replication capacity varied significantly between the different HLA-B but not HLA-A or HLA-C alleles, supporting a direct effect of HLA-

driven sequence variability on replication capacity since HLA-B alleles have been shown to exert the greatest selection pressure on HIV-1 [170]. Second, increasing numbers of HLA-associated mutations, particularly those associated with HLA-B alleles, in or flanking Gag epitopes (likely HLA-selected escape mutations) generally correlated with decreased HIV-1 replication capacities in the present project (in both recent and chronic infection), and similar correlations were observed in an independent study [251]. Third, specific mutations in Gag epitopes that are selected by protective alleles have been demonstrated to significantly decrease HIV-1 replication capacity [33, 187, 211-213]. Two of these mutations, namely 242N (associated with HLA-B*57/B*5801) and 163G (associated with HLA-B*57), were associated with lower replication capacities in the early HIV-1 infection cohort of the present study. In addition, in the chronic infection cohort studied here, the 186S mutation present in the HLA-B*81-restricted epitope TL9 (residues 180-188; previously described as one of the key Gag epitopes under strong selection pressure by protective HLA alleles with variance mainly at residues 182 and 186 [174]), was identified as most strongly associated with lower replication capacities, thereby providing a possible mechanism for HLA-B*81 influence on HIV-1 replication capacity. These findings, namely the association of HLA-B*81 and the 186S mutation with lower Gag-protease-mediated fitness, were recently (in 2011) reproduced in a cohort of HIV-1 subtype C infected patients from Bloemfontein, South Africa [246]. Interestingly, in the present study, one common HLA allele (HLA-A*3001), not typically associated with HIV-1 control, was associated with lower Gag-protease-mediated replication capacities. Taking the results together, it seems likely that HLA alleles, to a greater extent protective HLA alleles and in particular HLA-B*81, may influence Gag-protease-mediated replication capacity through HLA-restricted CD8⁺ T cell-mediated selection pressure.

4.1.2 HLA-B*81 associated mutations: 186S

Since the HLA-B*81-associated 186S mutation in the TL9 epitope was strongly associated with reduced viral replication capacity, this mutation was introduced by site-directed mutagenesis into a laboratory-adapted strain of HIV-1 to directly test the consequences of this mutation for HIV-1 replication. A significant deleterious effect of 186S on HIV-1 replication was confirmed. Consistent with this data, residue 186 in HIV-1 Gag has previously been classified as a reverting mutation site, at which mutations revert on transmission to a host lacking the HLA allele that selected them, presumably due to a fitness cost [209]. Furthermore, introduction of 186A, a rare naturally occurring variant at this codon (observed at a frequency of 0.5% in the SK cohort), into HIV-1 previously resulted in defective virus replication [70, 279].

The mechanisms involved in decreased replication capacity of the 186S mutant are unknown. There are however previous detailed reports of defective virus replication following introduction of 186A (or capsid residue 54A) into HIV-1 [70, 279] and similar mechanisms may apply to 186S. Although residue 186 occurs outside the CypA binding loop, the defective phenotype (186A) is dependent on the presence of CypA and replication of the mutant is significantly enhanced by addition of the CypA inhibitor, cyclosporin A (CsA) [279]. It has been proposed that CypA protects HIV-1 from an unknown host restriction factor in human cells, possibly through modulating capsid structure following binding [34, 279] (Section 1.3.3). Thus these studies may suggest that the 186A mutant is inhibited by an unknown host restriction factor in a CypA-dependent manner through alteration of the CypA-capsid interaction [279]. However, since CsA did not fully restore replication capacity of the mutant to wild-type levels, there may be other mechanisms

involved. Consistent with this, 186A reduced the efficiency of capsid formation [70], resulted in unstable capsids, and also modestly reduced reverse transcription [279].

In the present study, it was observed that viruses harbouring 186S in the presence of other HLA-B*81-associated mutations, 182S and 190X, replicated better than those with 186S alone (although 182X was also individually associated with lower viral replication capacities), suggesting that 182S and 190X may partially compensate for the fitness cost of 186S. However, site-directed mutagenesis experiments in a laboratory-adapted HIV-1 strain revealed that 182S did not increase fitness in the presence of 186S, and while mutations at position 190 tended to increase the fitness of 186S mutants, this was not statistically significant. In support of a compensatory role (albeit very weak) for 190X and not 182S, in the cohort analysed here, mutations at codon 190 occurred in conjunction with 186S in 13 out of 18 cases (72%) and only in 2 cases (out of 25 cases) without 186S in HLA-B*81 positive individuals, while 182S only occurred in conjunction with 186S in 7 out of 21 cases (33%) and mutations at residue 182 in the absence of 186S were common in HLA-B*81 positive individuals (occurring in 52% of sequences). The close proximity of residues 186 and 190, which occur parallel to one another in a helix structure (Figure 3.18), may explain the slight replication advantage observed when both positions, rather than position 186 alone, were mutated.

Overall, no evidence was found that mutations co-varying with 186S significantly and substantially compensate for the fitness cost of this mutation. Although, since there were a few instances in which Gag-protease recombinant viruses harbouring 186S in the natural Gag-protease background had average to high replication capacities, the possibility that mutations at other residues not identified as co-varying specifically with 186S might be an

important part of the interaction with HLA-B*81-associated mutations cannot be excluded. However, consistent with the hypothesis that the 186S mutation is difficult to compensate, Huang *et al.* (2011) showed an association of 186S with higher CD4⁺ T cell counts and evidence for its reversion (probably concurrent with immune relaxation) at extremely low CD4⁺ T cell counts (<100 cells/mm³) in individuals with advanced disease [246]. In contrast, that study showed preservation of the HLA-B*57/*5801-associated fitness-reducing 242N mutation, likely due to the development of compensatory mutations that resulted in increased Gag-protease-mediated replication capacity at extremely low CD4⁺ T cell counts.

The apparent lack of effective compensation for HLA-B*81-associated 186S mutation and efficient eventual compensation for HLA-B*57/B*5801-associated 242N mutation may also explain the particularly strong association between HLA-B*81 and lower replication capacity, relative to HLA-B*57/B*5801, in HIV-1 subtype C chronic infection. A parallel, companion study on individuals recently and chronically infected with HIV-1 subtype B [251] lends support to this idea. That study showed an association between protective HLA alleles (namely HLA-B*13, HLA-B*27, HLA-B*57, and HLA-B*5801), as well as increasing numbers of host HLA-associated polymorphisms, and lower Gag-protease-mediated replication capacities in early infection but not in the very late chronic stage of infection. This suggests that the development of compensatory mutations over the course of infection diminished associations between HLA-mediated immune pressure and replication capacity. The lack of such significant associations in the chronic infection cohort of that study while the presence of such associations in the present study may be at least partly explained by the notably earlier stage of chronic infection of the present study cohort, as evidenced by considerably higher average CD4⁺ T cell counts.

Interestingly, in a recent study, the number of public T cell clonotypes specific for SIV Gag CM9 (residues 181-189), which occurs in the same region as HLA-B*81-restricted TL9 (residues 180-188) and is restricted by the protective allele Mamu-A*01, correlated strongly and negatively ($r^2 = -0.71$) with viral set point in rhesus macaques [280]. This suggests that the CD8+ T cell targeting of CM9 may result in slower SIV disease progression, which might be partly explained by fitness constraints since an escape mutation in CM9 was previously shown to significantly reduce SIV fitness [281]. It should be noted though that the 186S mutation in the HIV-1 TL9 epitope was not associated with lower viral loads in the present study of HIV-1 chronic infection (data not shown). In fact, in HLA-B*81 positive individuals only, there was a trend towards higher viral loads in individuals harbouring the 186S mutation. It should also be noted that, although 186S was strongly associated with reduced replication capacity, there was no difference in the replication capacities of viruses with or without 186S in HLA-B*81 positive individuals, and since there was no strong evidence for compensation of the 186S mutation, this may suggest that other mutations selected by HLA-B*81 are also responsible for the lower fitness associated with this allele. In fact, mutations at the HLA-B*81-associated codon 182 were also significantly linked to lower replication capacity, although this was not as pronounced as for 186S. It may be that, while decreased viral fitness due to HLA-driven mutations could be beneficial in HLA-B*81 positive individuals, the development of the 186S mutation may decrease the effectiveness of CD8+ T cell responses to TL9 that may be important in maintaining low viral loads, and that a balance between fitness costs and effective CD8+ T cell responses is important in determining clinical outcome (discussed in Section 4.2). Huang *et al.* (2011) suggest that viral fitness influences HIV-1 disease progression: they found that reversion of 186S, compensation of fitness-reducing mutations, and generally increased viral fitness were associated with advanced disease, and suggest that increased viral fitness is a mechanism of progression to AIDS [246]. The relationship between HIV-1 Gag-protease-mediated fitness,

in particular HLA-driven fitness costs in Gag, and HIV-1 disease progression is considered next.

4.2 Impact of HIV-1 replication capacity on disease progression

4.2.1 Chronic infection: cross-sectional and longitudinal analyses

The cross-sectional correlation of lower Gag-protease-mediated replication capacities with lower baseline viral loads and higher CD4+ T cell counts in the chronically infected cohort of the present study is in accordance with Huang and co-workers' (2011) recent demonstration of increased Gag-protease fitness at extremely low CD4+ T cell counts, also in cross-section [246]. A parallel study on individuals chronically infected with HIV-1 subtype B yielded similar findings [251]. These results suggest that Gag-protease-mediated replication capacity increases with HIV-1 disease progression. This finding is also consistent with a previous study that showed increasing fitness of HIV-1 isolates with disease progression [247].

Since the correlation between markers of disease progression and replication capacity in the chronically infected cohort was cross-sectional, it could not be determined from this result whether increasing HIV-1 fitness results in further disease progression or if it is merely a consequence of disease progression. Therefore, the relationship between replication capacity at baseline and subsequent rate of CD4+ T cell decline was analysed. No correlation between these two parameters was observed. A potential explanation for the lack of correlation may be that replication capacity is not static and the development of compensatory mutations at a time point later than that measured may influence the subsequent rate of CD4+ T cell decline. Data from previous studies suggest that mutations

with a fitness cost are readily compensated [33, 212, 213, 251]. The lack of correlation between replication capacity and rate of CD4⁺ T cell decline could therefore indicate that Gag-protease-mediated replication capacity measured in chronic infection does not have an enduring impact on the rate of disease progression. An alternative explanation for the lack of correlation could be that replication capacity significantly influences disease progression but that a balance between replication capacity and Gag CD8⁺ T cell responses influences clinical outcome *i.e.* the relationship is not unidirectional. Accumulation of escape mutations in HIV-1 carries a fitness cost to the virus, but at a certain level of accumulation, the disadvantage to the virus is offset by the advantage of escaping effective CD8⁺ T cell responses that were holding replication in check, resulting in increased viral loads and accelerated disease progression despite a replication-deficient virus [187, 228]. It should also be noted that rate of CD4⁺ T cell decline was previously described to explain only 3% of the variability in time to AIDS, probably due to the large variance in the CD4⁺ T cell count slopes [282] (Section 1.6.1). Therefore, factors influencing disease progression may not correlate as strongly with this measurement as with viral set point, for example, which is a strong correlate of time to AIDS [283] (Section 1.6.1).

4.2.2 Recent infection: longitudinal analyses

To better assess the impact of Gag-protease-mediated replication capacity on the rate of disease progression, the replication capacity of the early virus in recent infection was related to the subsequent viral set point as well as rate of CD4⁺ T cell decline. A weak but significant association between reduced Gag-protease-mediated replication capacities in early infection and lower viral set points, as well as a trend of correlation between lower replication capacities and slower rates of CD4⁺ T cell decline, was found. A recent study (in 2010), comparing controllers and non-controllers infected with HIV-1 subtype B,

demonstrated considerably stronger correlations between Gag-protease-mediated replication capacity in early infection and disease outcome than were found in the present study [226]. This difference could possibly be partly explained by the lack of extreme phenotypes at the lower end of the clinical spectrum in the present study. The observation of a trend of association between Gag-protease-mediated replication capacity and rate of CD4⁺ T cell decline in early infection in the present study, while the absence of such an association in chronic infection, might suggest that replication capacity in early infection has more far-reaching effects than that measured in chronic infection.

To analyse the impact of HLA-driven fitness costs in Gag on disease progression, the number of HLA-associated Gag polymorphisms was related to replication capacity and viral set point in recently infected patients. The number of HLA-associated polymorphisms (mainly attributed to HLA-B) in or adjacent to Gag epitopes was inversely associated with replication capacity, however, unexpectedly the presence of host-specific HLA-B-associated polymorphisms was positively associated with viral set point. One contributing factor could be the presence of compensatory mutations since no significant relationship with viral set point was observed when the analysis was limited to polymorphisms in or adjacent to epitopes. Further, although the host expressed the required HLA alleles to select these polymorphisms, it is unlikely that many were selected by the host at such an early stage of infection and they were thus presumably largely inherited. Usually, in the Gag protein, the fastest selection occurs in epitopes restricted by the protective HLA-B*57/B*5801 within about six months of infection [106, 284] (but on average at 13 months post-infection in another study [187]) and reversion events rather than selection of new mutations in Gag predominate in early infection [187, 285]. Further, Schaefer *et al.* (2009) show that in linked transmission pairs most of the HLA-associated polymorphisms present at one year after

infection were transmitted and not selected by the host [286], and it is estimated that approximately 18% of the possible mutations in Gag are present in transmitted viruses on average [287]. Therefore, these results indicate a host disadvantage to inheriting polymorphisms which the host could have newly selected, likely due to diminished early CD8+T cell responses to the mutated epitopes, as suggested in previous studies [187, 207].

Analysis of Gag sequences derived from individuals with protective HLA alleles in recent infection yielded similar conclusions. Gag-protease recombinant viruses derived from individuals with protective HLA alleles in early infection tended to have reduced replication capacities, but viral set points were not significantly lower; therefore, a benefit to reduced Gag-protease function in these particular individuals was not overall evident. In fact, in individuals with protective HLA alleles, viral set points were overall significantly higher in those who had viruses with associated Gag mutations than in those who did not. Conversely, in six of the seven subjects with protective HLA alleles for whom Gag-protease amplification failed here (but were sequenced previously [278]), there were no Gag mutations associated with the hosts' protective HLA alleles and the viral set points were considerably low.

One possible explanation is that some early selection of fitness-reducing mutations occurred in individuals with protective alleles and was disadvantageous due to subsequent reduction in CD8+ T cell responses. The possibility of some early selection of fitness-reducing Gag mutations in individuals with protective alleles does exist, and has been reported as early as 45 days post-screening, but the average reported time of selection in another study was 13 months [187], suggesting that in the present study sequences were likely to be representative of the transmitted virus in most cases. Although a parallel study on individuals recently

infected with HIV-1 subtype B showed significantly reduced Gag-protease-mediated replication capacities in individuals with protective HLA alleles, the least fit viruses with associated fitness-reducing mutations were sampled from a later date post-infection than the fitter viruses [251]. Furthermore, although there was a trend towards lower replication capacities from individuals with protective HLA alleles in the present study, which may suggest some early selection pressure, there was an amplification bias. The viruses from individuals with protective HLA alleles who had higher viral set points were preferentially amplified, and this could suggest that there was preferential amplification of viruses with inherited polymorphisms that reduced fitness but precluded effective CD8⁺ T cell responses, leading to a trend of association between fitness and protective HLA alleles in early infection in the present study.

It is likely that in most cases here, the protective HLA-associated polymorphisms were inherited rather than selected by the hosts' protective alleles thereby resulting in weakened early CD8⁺ T cell responses to key epitopes and a disadvantage to the host. Supporting this hypothesis, 23 out of 50 individuals with non-protective HLA alleles harboured mutations associated exclusively with protective HLA alleles indicating high level transmission of these polymorphisms in sub-Saharan Africa. A notable example is the case of 242N which was present in 17 viruses out of which 15 were derived from individuals without the associated protective HLA-B*57/B*5801. Further, in one patient with HLA-B*5801 and a high viral set point (5.46 log₁₀ copies/ml), at 26 days post-infection two associated mutations were present (146L and 242N) at which point none of the viruses in other studies had shown evidence of adaptation in Gag [106, 251]. It is also notable that in a patient with HLA-B*81 and a high viral set point (4.92 log₁₀ copies/ml), the 186S mutation in the TL9 epitope was present, however, there was no detectable initial CD8⁺ T cell response to this epitope (data

not shown) suggesting it may have been inherited rather than selected. Yet in another patient (AS2-0945) with HLA-B*5703 and a low viral set point (3.46 log₁₀ copies/ml), all HLA-B*57-restricted epitopes were initially wild-type and one year later mutations 146P and 242N had been selected with a concomitant decrease in fitness, although there was also a coincident rise in viremia. A similar rise in viremia following selection of 242N was recently reported in a single patient [278], however decreased viremia following early selection of the HLA-B*57-associated mutations 242N and 147L has been reported in several patients [187]. Taken together, these data imply that the benefit in an individual with a protective allele is derived from an initial effective CD8⁺ T cell response and that the subsequent selection of fitness-reducing Gag mutations is likely to, but may not necessarily, contribute to clinical benefit. Furthermore, results suggest that CD8⁺ T cell-driven benefit appears diminished or even lost when protective HLA-associated polymorphisms in Gag are inherited by individuals with the relevant protective alleles, despite reduced viral replication capacity.

The high transmission prevalence of protective HLA-associated polymorphisms in the populations studied supports recent evidence that mutations selected by protective alleles may be accumulating over time [288]. This may be particularly evident in the present cohort since HIV infection is most prevalent in sub-Saharan Africa [10] (Section 1.1). In KwaZulu-Natal there is an explosive rate of infections in young adults probably fuelled by the high viral loads in acute infection [289] and consequently transmission may often occur early in infection before such mutations revert. The median reversion time reported for HLA-B*57-associated Gag mutations in HLA-mismatched recipients was 23 months [187] and for 242N mostly between six and 24 months [208]. The consequence of this is firstly, that HIV-1 is

being attenuated over time and secondly, that the association of current protective alleles with slower disease progression may eventually diminish [288].

Longitudinal measurements of replication capacities and CD8⁺ T cell responses to Gag epitopes in 13 recently infected patients also support the importance of both CD8⁺ T cell responses and viral replication capacity in influencing disease progression. For example, in one patient viral replication capacity was low but there was a lack of CD8⁺ T cell Gag responses and uncontrolled viremia. In others, CD8⁺ T cell Gag responses were present but viral replication capacities were high and viremia was uncontrolled. Two patients who had low viral set points showed low replication capacities and either CD8⁺ T cell responses to key epitopes (such as TW10) or a very broad Gag response, which has previously been suggested to be beneficial [195]. It should be noted that likely explanations for viral set point based on CD8⁺ T cell responses or viral replication capacity could not be made for all patients. Other HIV-1 regions, such as Env [222] and protease-reverse transcriptase [224], significantly contribute to overall viral replication capacity and may help to explain differences in viral control [290]. Indeed, Env ranks among the most important determinants of viral fitness [222, 240] and several studies provide evidence for overall HIV-1 fitness as a significant determinant of disease progression [218, 224] (Section 1.6.5.2.1).

Overall, results suggest that Gag-protease-mediated fitness has an impact on HIV-1 disease progression, but that a balance between viral fitness and effective CD8⁺ T cell responses together influence disease. Viral fitness in early infection may have a greater influence than that measured in chronic infection, and HIV-1 fitness tends to increase with disease progression (which may partly contribute to progression to AIDS). Since protective HLA alleles are associated with lower Gag-protease-mediated replication capacities, and viral

fitness influences disease progression, viral attenuation may be a mechanism by which protective HLA alleles slow disease progression. These results support the idea of a vaccine strategy in which immune responses are directed towards Gag with the aim to attenuate HIV-1 and slow disease progression (discussed further in Section 4.6).

4.3 Identification of mutations with impact on HIV-1 replication capacity

In such a vaccine strategy it would be desirable to identify Gag mutations that have significant impact on HIV-1 replication capacity, and more specifically, regions of Gag that would have the greatest fitness impact if mutated. In the present study, several Gag mutations associated with significant changes in replication capacity were identified. Site-directed mutagenesis experiments may constitute future work to confirm these associations with viral replication capacity.

Notably, several of the Gag mutations significantly associated with alterations in replication capacity were HLA-associated. As discussed earlier (Section 4.1), the 186S mutation linked to the protective HLA-B*81 allele was the most strongly associated with lower replication capacity. As shown previously [33, 187, 211, 212], Gag p24 mutations 163G and 242N, selected by the protective alleles HLA-B*5801 and/or HLA-B*57, were significantly associated with lower replication capacity and there was evidence for increased replication capacity with accumulation of previously described [33, 251] compensatory mutations, especially 248T. Previous reports suggest that the mechanism of replication defects induced by 186S and 242N may be related to altered interaction with the host factor CypA [33] (Sections 1.6.5.1 and 4.1.2), which plays an important role in the HIV-1 replication cycle [34] (Section 1.3.3). Two of the other HLA-associated codons linked to altered replication

capacities, codons 223 and 256, occur in and within 6 residues, respectively, of loops in Gag p24 that were previously described to interact with host cellular factors that determine HIV-1 host tropism [291]. A recent study included 223V and 256V in a group of eight polymorphisms suggested to be compensatory since they co-vary with polymorphisms in Gag p24 that were shown or inferred to carry fitness costs [292]. Accordingly, in the early infection cohort of the present study, V instead of the consensus amino acid I [267] was the most frequent residue at codon 223, and 223I was associated with decreased replication capacity. However, the 256V polymorphism correlated negatively with replication capacity. It is notable that several of the HLA-associated codons linked to replication capacity differences were associated with common HLA alleles or those not typically associated with HIV-1 control. A recent study has also reported attenuating mutations in Gag (specifically Gag p24) associated with the common allele HLA-C*03, which is not associated with viral control [293].

Several of the Gag codons identified in the present study as significantly associated with altered replication capacities in either chronic or recent infection (or both) have previously been implicated in specific roles in the HIV-1 replication cycle (discussed in Section 1.3). Codons 28, 30, 31, 110, 111, and 114 are part of NLSs in Gag p17 [41]. Some of these mutations could therefore disrupt the NLS and adversely affect replication capacity; however, only mutations at codons 31, 110, and 114 were associated with decreased replication capacity in the present study. Codons 15, 20, 28, 30, and 31 also form part of the basic domain involved in membrane binding, although codons 30 and 31 specifically are not basic residues [58]. Accordingly, mutation of the basic residues at positions 15 (K) and 20 (R) were associated with decreased replication capacity, while mutation of the basic residue H to another basic residue R at position 28 increased replication capacity. Mutation of the

consensus amino acid at codon 31 was associated with decreased replication capacity perhaps since this position is involved in Env glycoprotein incorporation [48]; however mutation of codon 35, also involved in Env glycoprotein incorporation [48], was associated with increased replication capacity. Mutations at codon 224 and an uncommon mutation at codon 223, both in the CypA binding loop (which plays an important role in the life cycle [34] [Section1.3.3]), were linked to decreased fitness. Codons 374, 377, 381, and 382 are in the p2/p7 cleavage site which controls the rate and order of Gag cleavage [249], and this may explain why these codons were significantly associated with changes in replication capacity. In addition, a mutation at residue 453 in the p1/p6 cleavage site was linked to lower replication capacity. Mutation of two of the basic residues (397 and 411) in Gag p7 required for budding [65] and partly responsible for RNA binding (as part of the hydrophobic plateau [53]), were associated with lower viral fitness. Codon 466 is one of the codons mediating incorporation of Vpr into virions [64] and a mutation at this residue was linked to decreased fitness. Mutation of residue 485 in the Alix binding L-domain, which mediates budding [67], was however associated with increased fitness. Thus, the disruption of specific functions mediated by codons listed above may explain some, although not all, of the mutation-fitness associations identified in the present study.

In general, the results from the present study support the hypothesis that mutations occurring at conserved residues/regions are more likely to result in a fitness cost or have the greatest fitness impact: HLA-associated escape mutations at conserved sites were associated with lower replication capacities, and most of the mutations significantly associated with altered replication capacities in conserved Gag p24 decreased replication capacity while most in the less conserved Gag p17 increased replication capacity. In agreement with these data, in an independent South African cohort, beneficial HLA alleles were associated with strong

selection at key epitopes which occurred mostly in Gag p24 [174] and there is evidence that HLA-B*57 mediates its protective effect mainly through attenuating mutations in Gag p24 [227]. Further, the breadth of Gag p24 but not p17 or p15 CD8+ T cell responses in HLA-B*13 positive individuals was significantly associated with decreasing viral load [188]. The direct relationship between HIV-1 replication capacity and the entropy of HLA-associated mutated sites in the present study is also supported by the recent finding that escape mutations in conserved Gag p24 carry significant fitness costs while most of the escape mutations in the highly variable Env protein are fitness neutral or increase fitness [104].

Another interesting result in the present study may be consistent with the finding that mutations in more conserved regions carry greater fitness costs than those in variable regions. Gag p17 and p7 were significantly more divergent from consensus than the p24 region, *i.e.* significantly more mutations occurred in variable p17 and p7 regions than in the conserved p24 region, and percentage similarity to consensus in both p17 and p7 was overall negatively correlated with fitness, while there was no correlation for p24. This resulted in an overall negative correlation between the number of amino acid differences from consensus C Gag and viral fitness, *i.e.* a greater number of changes from consensus was associated with increasing viral fitness. In summary, the majority of mutations in HIV-1 arise in variable regions and are likely to have no/little fitness cost or actually increase fitness.

4.3 Impact of Gag-protease on overall HIV-1 replication capacity

An important finding of the present study was the moderately strong and significant correlation between the fitness of 16 HIV-1 isolates and the fitness of recombinant viruses encoding only the Gag-protease of these isolates. This suggests that, in addition to Env [222] and protease-reverse transcriptase [224], variations in HIV-1 Gag-protease sequence

significantly impact overall viral fitness. This is consistent with the important role that Gag plays in all stages of the HIV-1 replication cycle (reviewed in Section 1.3). Since Gag-protease recombinant viruses generated from individuals with protective HLA alleles had significantly lower replicative capacities, this result further supports overall virus attenuation through targeting of Gag by protective HLA alleles as a mechanism of slower disease progression. These data are also consistent with the idea that, in a vaccine strategy to reduce viral fitness, targeting the Gag gene should effectively impact overall viral fitness following the selection of attenuating mutations.

4.4 Impact of Gag-protease on HIV-1 replication capacity in different subtypes

Given the importance of Gag in HIV-1 replication and as a target of effective immune responses, the contribution of Gag-protease function to the dramatically lower fitness of subtype C isolates when compared with subtype B isolates [221] was investigated in the present study. Gag-protease NL4-3 recombinant viruses encoding subtype C Gag-proteases were significantly less fit than those encoding subtype B Gag-proteases. These results suggest that the considerable fitness difference observed between subtype B and C isolates may not only be explained by Env differences, as suggested by a previous study comparing Env gp120 recombinant viruses derived from a single subtype C isolate and subtype B isolate [222], and similarly reverse transcriptase differences [294], but also by differences in the Gag-protease region. Since a recent study did not observe differences in protease activity between HIV-1 subtype B and C [294], and subtype C protease was previously reported to have a higher catalytic efficiency than subtype B protease [295], the Gag protein is likely to largely contribute to the results here. The finding that Gag-protease significantly contributes to overall HIV-1 fitness is also consistent with the explanation that Gag-protease variability partly explains lower fitness of subtype C viruses when compared to subtype B

viruses. However, an alternative explanation for this result could be that insertion of Gag-protease into the backbone of another subtype is disadvantageous, which would point to an important interaction between Gag-protease and another region of the HIV genome, for example Env gp41 [296]. In favour of the former rather than the latter explanation, greater similarity of subtype C Gag sequences to the consensus subtype B sequence was not associated with fitter subtype C Gag-protease NL4-3 recombinant viruses. However, further fitness studies using a replicating subtype C backbone may be required to exclude the latter possibility.

To further investigate the difference between subtype B and C Gag-protease-mediated fitness, residues at which amino acids differed between the subtypes and were also linked to altered replication capacities in the present study and/or previous studies [251], were introduced into a laboratory-adapted strain of HIV-1. A significant finding was that the deletion of Gag residues 483 and 484 from HIV-1 subtype B (since these residues are not present in consensus subtype C) substantially decreased replication capacity of HIV-1 subtype B, and correspondingly, insertion of the subtype B amino acids at these codons into HIV-1 subtype C Gag-protease (placed in a subtype B NL4-3 backbone) significantly increased replication capacity. Residues 483 and 484, in a L-domain at Gag residues 483-492, are essential for binding the Alix host protein [64], which acts in concert with the primary budding factor Tsg101, recruited by the L-domain at Gag residues 455-458, to mediate viral budding [68] (Section 1.3.8). It is notable that an unusual mutant with deletion at residues 482 and 483, shown to disrupt Alix binding [67], was previously associated with non-progressive HIV-1 infection [297]. Fujii and co-workers (2009) recently demonstrated significantly reduced particle production and infectivity of Alix-binding domain mutants (including a virus mutated at codon 484) with disrupted Alix binding, and this was particularly severe in the presence of Tsg101-binding domain mutations, demonstrating an

important role for Alix in HIV-1 replication [68]. Although Alix-binding mutations slowed viral replication in several cell types including PBMCs, Alix-binding was not essential for virus replication in PBMCs, and passaging of mutants increased replication capacity concomitant with the development of secondary site mutations in gp120 and gp41 [68]. The putative compensatory role of these secondary site mutations is yet to be confirmed by the investigators of that study. Overall, these results may suggest that the subtype-specific difference in the consensus amino acids at Gag codons 483 and 484 largely contributes to the fitness difference between HIV-1 subtypes B and C, and that this is mediated by altered budding efficiency. But a possibility exists that disrupted Alix binding is compensated for by secondary site mutations in other HIV-1 subtype C proteins, and that observed effects were due to incompatibility of subtype C and B components. Future experiments to test this possibility may include mutation of these residues in a replicating subtype C backbone.

4.5 Study limitations

As mentioned, a limitation of the present study was the insertion of subtype C *gag-protease* into a subtype B NL4-3 backbone to test the fitness effects of the introduction of subtype B-specific mutations into subtype C Gag-protease. Recombinant viruses constructed using clinically-derived HIV-1 subtype C Gag-protease sequences were also a mixture of HIV-1 subtype B and C genetic components. It is therefore possible that the fitness of the recombinant viruses encoding patient-derived Gag-protease could be a reflection of the compatibility of the subtype C Gag-protease with the subtype B backbone and that some Gag-protease mutations associated with altered replication capacities might represent those that interact with other components of the backbone. However, the degree of similarity of subtype C Gag to subtype B did not associate with the fitness of these subtype C/B recombinant viruses. Furthermore, the fitness of subtype C/B recombinants correlated with

viral loads, CD4⁺ T cell counts, as well as specific HLA alleles in the present study, as did subtype B/B recombinants in a parallel, companion study [251], indicating that, despite the mixing of subtypes, the assay was biologically and clinically relevant.

The introduction of HLA-B*81-associated Gag mutations (182S, 186S, and 190X), identified in clinically-derived subtype C Gag sequences, firstly into a subtype C/B recombinant virus and then into a fitter subtype B laboratory strain (due to failure in generating some mutants using the subtype C/B recombinant) is also a study limitation. Introduction of mutations out of context of the natural sequence background is a general limitation of site-directed mutagenesis experiments [236] (Section 1.7.1.1). However, particularly since variability at Gag residues 182, 186, and 190 is rare in HIV-1 subtype B viruses (likely because the HLA-B*81 allele is uncommon in Caucasian individuals in whom this subtype is most commonly found [298]), a subtype C background would more likely reflect the fitness effect of mutations at these residues as they occur naturally. The discrepant results obtained for the 182S mutation in the different backbones used (182S in the subtype C/B recombinant significantly reduced fitness with respect to wild-type while 182S in NL4-3 did not), further increase the desirability of confirmatory experiments in a purely subtype C background. Future experiments may involve identification of a subtype C backbone that will replicate adequately in the GXR cell line and may be used to repeat site-directed mutagenesis of the HLA-B*81-associated Gag mutations.

As discussed in Section 1.7.2, the possibility that viral replication capacity measurements might differ if primary cells were used cannot be excluded. In support of the relevance of results obtained in the current assay system, viral replication capacity measurements correlated with clinical outcomes and mutations previously shown to have a fitness cost,

such as 242N and 163G [33, 187, 211, 212], were also associated with lower fitness in the present study. Further, although primary cells are not used in the current assay system, it has the advantages of using a homogenous population of cells and of being high-throughput [71] (Section 1.7.4).

It should also be noted that parameters other than those measured in the present study, such as T cell activation, are important markers of disease progression [137] (Section 1.6.1) and it remains to be determined whether these additional important parameters that influence HIV-1 pathogenesis might also correlate with viral fitness.

4.6 Implications for HIV-1 vaccine design

The present study strengthens evidence that HLA alleles may drive biologically and clinically relevant fitness costs in HIV-1 through selection of immune escape mutations in Gag. This knowledge could be used to design an HIV-1 vaccine to stimulate specific CD8+ T cell responses that could result in HIV-1 attenuation, reduce viral set point, slow disease course, and reduce transmission at the population level, which may be a more realistic/feasible goal than a sterilising vaccine approach [94, 121]. Such a vaccine should stimulate CD8+ T cell responses to multiple epitopes in which immune escape mutations result in significant fitness costs and avoid including epitopes in which mutations result in no fitness cost or increased viral fitness [121, 299, 300]. The identification of Gag amino acid variants, many of which were at HLA-associated codons and/or located in CD8+ T cell epitopes, that significantly impact viral replication capacity in the present study should aid in rational design of a fitness-reducing vaccine. The fact that some of these variants were associated with common HLA alleles, supports the relevance of such a vaccine for the

general population. The data from the present study further indicate that conserved Gag epitopes should be targeted in such a vaccine, since mutations in conserved rather than variable regions more often resulted in significant fitness costs. The targeting of multiple conserved epitopes should constrain immune escape and only allow partial immune evasion (thereby allowing continued immune responses) which would then result in significant fitness costs [299, 300]. A vaccine targeting conserved elements has the added advantage of overcoming the problem of HIV-1 diversity [299]. Further in support of a conserved elements vaccine to maintain immune responses and/or reduce viral fitness, previous studies show that HIV-1-infected individuals who target conserved epitopes across the HIV-1 proteome, including those with common non-protective HLA alleles, have low viral loads [191, 192].

Recent work by Dahirel *et al.* (2011), using methods from quantum physics, has identified groups of amino acids in HIV-1 that co-evolve, termed “sectors”, and supports the concept of a fitness-reducing HIV-1 vaccine targeting structurally constrained regions [301]. This study revealed that there is a particularly vulnerable sector in Gag in which multiple mutations are likely to result in low viral fitness. The authors of that study recommend an HIV-1 vaccine including protein segments which cover epitopes (including those targeted sub-dominantly by common HLA alleles to achieve broad population coverage) containing sites in this sector. Interestingly, two suitable immunogens with broad population coverage suggested are Gag residues 160-188 and 240-277. These two immunogens are in the conserved Gag p24 region (in which mutations were more likely to result in fitness costs in the present study) and include residues which were identified, in the present study and/or previous studies [33, 187, 211, 212], to carry significant fitness costs when mutated, namely 163G, 186S, and 242N. Future work to confirm the findings of Dahirel *et al.* (2011) [301]

could involve testing whether an increasing number of mutations in the vulnerable Gag sector correlates with decreased viral fitness measured *in vitro*.

In addition, the fitness consequences of mutations not occurring naturally in combination could be considered for a vaccine strategy aiming to constrain immune escape and reduce viral fitness. A recent study identified 18 pairs of non-covarying (mutually exclusive) escape mutations in Gag p24 which resulted in severe fitness defects in combination, and could therefore be considered for simultaneous targeting by a HIV-1 vaccine [302]. Such an antagonistic pair of HLA-associated escape mutations in a conserved region of Nef also reduced replication capacity in combination as well as impairing the pathogenic activities of Nef [303]. This highlights that regions in the HIV-1 proteome other than Gag also be considered in a fitness-based vaccine strategy.

A potential consequence of a fitness-reducing vaccine strategy is the accumulation of immune-driven escape variants at the population level [300]. This may result in the loss of immunogenic CD8+ T cell epitopes and diminished vaccine-induced CD8+ T cell immune responses, and is a potential drawback of this vaccine strategy [300]. There is already evidence, from previous studies [288] and the present study (Section 4.2.2), that fitness-neutral immune escape mutations selected by common HLA alleles, or fitness-reducing escape mutations selected by protective alleles and transmitted prior to reversion or stabilised by compensatory mutations, are accumulating at the population level, and vaccines may have to be designed accordingly [288]. Another potential challenge of the fitness-reducing strategy is the development of compensatory mutations that may eventually restore or partially restore viral fitness, although the extent of this would probably be dependent on the level of viral control achieved [300]. A more thorough understanding of immune-driven

fitness costs and compensation and the consequences of this at the individual and population levels, is required for the development of an effective fitness-reducing vaccine [300].

The consequences of immune-driven fitness costs in HIV-1 for pathogenesis and the implications of this for HIV-1 vaccine design are outlined in Figure 4.1.

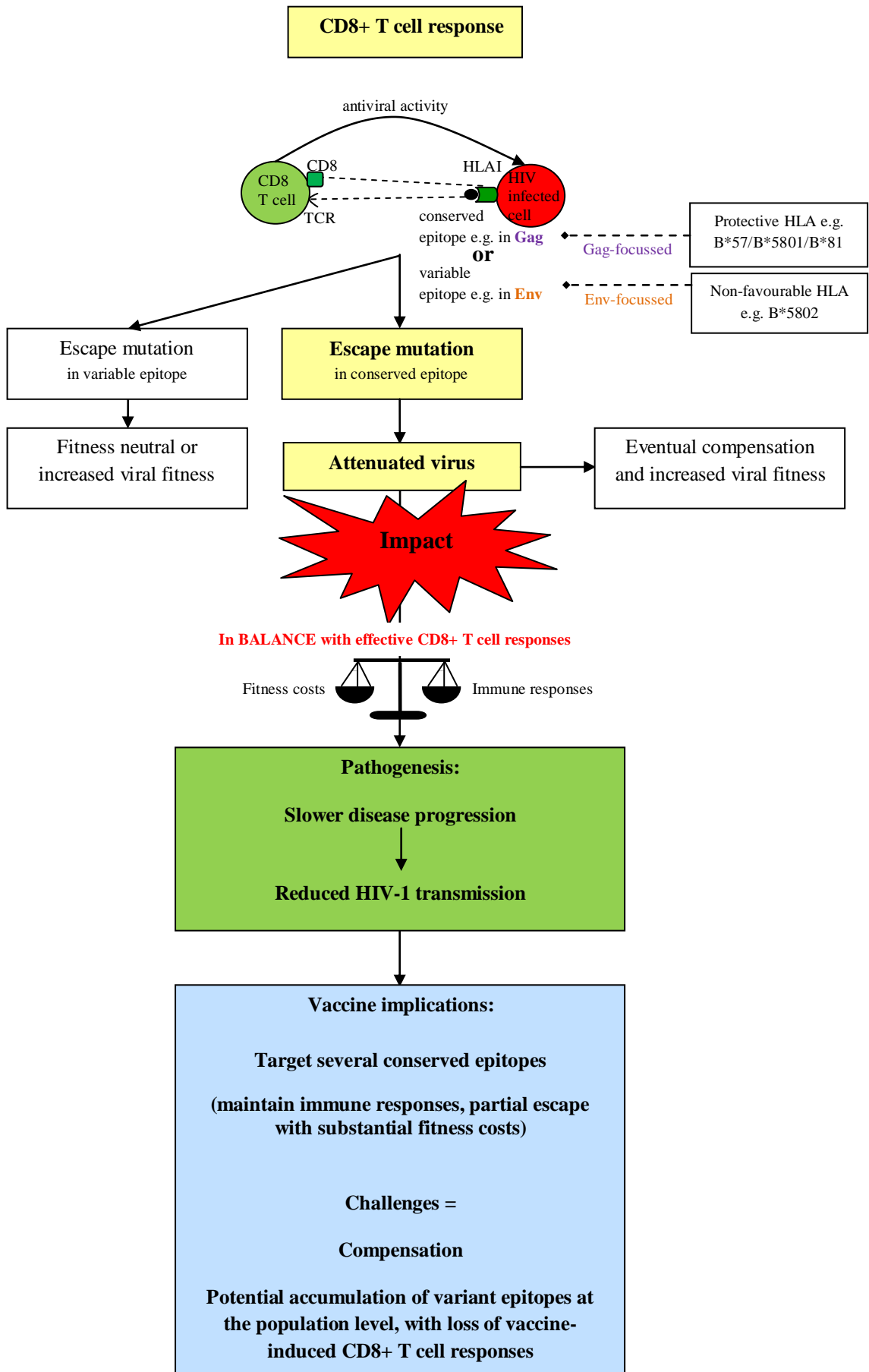
4.7 Summary and conclusions

In summary, the findings of the present study support the hypothesis that HLA alleles, in particular the protective allele HLA-B*81, may impact HIV-1 replication capacity through driving the selection of attenuating mutations in Gag, and that viral attenuation may result in clinical benefit. The data support that this may be a mechanism of slower disease progression in individuals with protective HLA alleles, but also highlight that a balance between HLA-associated viral fitness costs and effective CD8⁺ T cell responses is important in determining clinical outcome. Results also indicate that conserved regions of Gag, particularly in Gag p24, be included in a vaccine designed to constrain immune escape and reduce viral fitness in the event of escape. Results suggesting the significant contribution of Gag to overall HIV-1 fitness and as a determinant of fitness differences between subtype B and C isolates, further support that a vaccine targeting immune responses to specific Gag regions could successfully attenuate HIV-1. In conclusion, the present study showed evidence for biologically and clinically relevant fitness costs of naturally occurring HLA-driven sequence variability in Gag, particularly for protective HLA alleles, and this knowledge could possibly be exploited in the design of a fitness-based HIV-1 vaccine.

Figure 4.1 Impact of immune escape with fitness costs on HIV-1 pathogenesis and implications for HIV-1 vaccine design

The potential consequences of CD8⁺ T cell-mediated attenuation of HIV-1 for clinical outcome and HIV-1 transmission (in green) and the implications of this for HIV-1 vaccine design (in blue) are shown in this figure. A balance between fitness costs (due to escape) and effective CD8⁺ T cell responses in influencing HIV-1 pathogenesis is highlighted. The dotted lines indicate binding between the CD8 receptor and the HLA class I molecule, and between the TCR and epitope complexed with the HLA molecule on the surface of the infected cell, leading to cytotoxic and non-cytotoxic antiviral activity.

HIV-1 – human immunodeficiency virus type 1; CD – cluster of differentiation; HLA – human leukocyte antigen; TCR – T cell receptor; Gag – group specific antigen; Env – envelope glycoprotein.



APPENDIX

APPENDIX

Table A1. HLA-associated amino acids in HIV-1 subtype C Gag-protease

Protein	Association	HLA allele	Codon	Amino acid	Consensus	p value	q value
Gag	Negative	B*1503	7	I	I	<0.0005	0.106
Gag	Positive	B*1503	7	V	I	<0.0005	0.106
Gag	Positive	B*1503	11	G	G	0.001	0.176
Gag	Negative	A*74	12	K	K	<0.0005	<0.0005
Gag	Positive	A*74	12	N	K	<0.0005	<0.0005
Gag	Negative	B*5802	14	D	D	0.001	0.176
Gag	Positive	B*5802	14	E	D	0.001	0.176
Gag	Negative	A*74	20	R	R	0.001	0.165
Gag	Negative	C*06	20	R	R	<0.0005	0.106
Gag	Positive	C*06	20	K	R	<0.0005	0.001
Gag	Positive	A*6801	20	R	R	<0.0005	0.095
Gag	Positive	A*74	20	K	R	<0.0005	0.042
Gag	Negative	C*17	28	H	H	<0.0005	0.026
Gag	Positive	B*13	28	K	H	<0.0005	<0.0005
Gag	Positive	B*42	28	Q	H	<0.0005	<0.0005
Gag	Positive	C*17	28	S	H	<0.0005	0.056
Gag	Negative	C*17	30	M	M	<0.0005	<0.0005
Gag	Positive	B*42	30	K	M	<0.0005	0.022
Gag	Negative	C*0701	31	L	L	0.001	0.120
Gag	Negative	B*1503	49	G	G	<0.0005	0.043
Gag	Positive	A*3001	54	A	S	0.001	0.186
Gag	Positive	A*6802	54	T	S	0.001	0.116
Gag	Positive	B*1402	62	N	K	<0.0005	0.106
Gag	Positive	C*16	67	S	A	0.001	0.160
Gag	Positive	C*1601	67	S	A	<0.0005	0.082
Gag	Negative	A*01	73	E	E	<0.0005	0.096
Gag	Negative	A*29	79	F	Y	<0.0005	0.003
Gag	Positive	A*2902	79	Y	Y	<0.0005	0.001
Gag	Negative	A*34	90	K	E	<0.0005	0.040
Gag	Negative	A*74	91	R	K	<0.0005	0.037
Gag	Negative	C*06	91	R	K	<0.0005	0.036
Gag	Negative	B*4101	93	E	E	0.001	0.117
Gag	Positive	A*74	93	G	E	<0.0005	0.110
Gag	Positive	B*08	93	K	E	<0.0005	0.036
Gag	Positive	A*3001	103	R	K	0.001	0.165
Gag	Positive	B*5801	106	K	E	<0.0005	0.036
Gag	Negative	B*44	121	A	D	<0.0005	0.030
Gag	Positive	C*04	128	A	V	<0.0005	0.115
Gag	Negative	A*3009	138	L	L	<0.0005	0.001
Gag	Negative	C*06	146	P	A	<0.0005	<0.0005
Gag	Negative	B*1510	146	A	A	0.001	0.176
Gag	Negative	B*57	146	A	A	<0.0005	0.019
Gag	Negative	B*4201	146	P	A	<0.0005	<0.0005
Gag	Positive	B*57	146	P	A	<0.0005	0.096
Gag	Negative	A*2911	147	I	I	<0.0005	0.004
Gag	Negative	B*57	147	I	I	<0.0005	0.002
Gag	Negative	B*1503	147	L	I	0.001	0.148
Gag	Positive	A*2911	147	L	I	<0.0005	0.020
Gag	Positive	B*57	147	L	I	<0.0005	0.006
Gag	Positive	C*15	147	L	I	<0.0005	0.004

Gag	Positive	B*1503	147	I	I	<0.0005	0.042
Gag	Negative	B*5703	163	A	A	<0.0005	0.009
Gag	Positive	B*5703	163	G	A	<0.0005	<0.0005
Gag	Positive	B*5703	163	N	A	<0.0005	0.101
Gag	Negative	A*0202	165	S	S	0.001	0.174
Gag	Positive	A*0202	165	N	S	0.001	0.174
Gag	Negative	B*81	182	Q	Q	<0.0005	0.022
Gag	Negative	C*17	182	Q	Q	<0.0005	0.045
Gag	Positive	B*81	182	S	Q	<0.0005	0.042
Gag	Negative	B*81	186	T	T	<0.0005	<0.0005
Gag	Positive	B*81	186	S	T	<0.0005	<0.0005
Gag	Negative	B*81	190	T	T	<0.0005	0.004
Gag	Negative	C*18	190	T	T	<0.0005	<0.0005
Gag	Positive	C*18	190	A	T	0.001	0.129
Gag	Positive	A*43	215	M	L	0.001	0.176
Gag	Positive	C*07	215	I	L	<0.0005	0.058
Gag	Positive	B*5801	219	P	H	<0.0005	0.082
Gag	Positive	A*2911	223	I	I	0.001	0.119
Gag	Negative	B*57	242	T	T	<0.0005	<0.0005
Gag	Negative	B*5801	242	T	T	<0.0005	<0.0005
Gag	Positive	B*57	242	N	T	<0.0005	<0.0005
Gag	Positive	B*5801	242	N	T	<0.0005	<0.0005
Gag	Positive	B*4201	252	A	N	0.001	0.136
Gag	Negative	C*18	256	I	I	<0.0005	0.106
Gag	Negative	B*35	260	D	D	<0.0005	0.006
Gag	Positive	B*35	260	E	D	<0.0005	0.001
Gag	Negative	B*14	302	K	K	<0.0005	<0.0005
Gag	Positive	B*14	302	R	K	<0.0005	<0.0005
Gag	Negative	C*0304	303	T	T	<0.0005	<0.0005
Gag	Positive	C*0304	303	A	T	<0.0005	0.003
Gag	Negative	B*53	309	A	A	<0.0005	0.027
Gag	Negative	B*44	312	D	D	<0.0005	<0.0005
Gag	Positive	B*44	312	E	D	<0.0005	<0.0005
Gag	Negative	B*39	332	T	T	<0.0005	0.030
Gag	Negative	C*08	335	K	R	<0.0005	0.027
Gag	Positive	C*08	335	R	R	<0.0005	0.027
Gag	Negative	C*0804	336	A	A	<0.0005	0.001
Gag	Positive	C*0804	336	G	A	<0.0005	0.001
Gag	Negative	C*08	339	P	P	<0.0005	0.028
Gag	Negative	C*0804	339	P	P	0.001	0.120
Gag	Positive	B*15	339	A	P	<0.0005	0.028
Gag	Negative	B*1510	340	G	G	<0.0005	0.001
Gag	Positive	C*0304	340	A	G	<0.0005	0.004
Gag	Negative	B*07	357	S	S	<0.0005	<0.0005
Gag	Positive	B*07	357	G	S	<0.0005	<0.0005
Gag	Negative	C*04	370	A	A	<0.0005	0.068
Gag	Positive	B*1401	370	A	A	0.001	0.134
Gag	Positive	A*3002	371	N	N	<0.0005	0.054
Gag	Negative	B*1503	374	T	T	0.001	0.176
Gag	Negative	B*5802	377	L	M	<0.0005	0.083
Gag	Positive	B*5802	377	M	M	<0.0005	0.083
Gag	Negative	B*3501	378	M	M	<0.0005	0.020
Gag	Negative	A*43	381	G	S	0.001	0.176
Gag	Negative	C*18	382	N	N	<0.0005	0.002
Gag	Positive	A*01	382	K	N	<0.0005	0.026
Gag	Negative	B*42	386	P	P	<0.0005	0.001
Gag	Positive	B*4201	386	S	P	<0.0005	0.001
Gag	Negative	A*03	403	K	R	<0.0005	<0.0005
Gag	Negative	A*3001	403	K	R	0.001	0.197

Gag	Negative	A*74	403	R	R	<0.0005	<0.0005
Gag	Positive	A*74	403	K	R	<0.0005	0.088
Gag	Positive	A*03	403	R	R	<0.0005	<0.0005
Gag	Positive	A*3001	403	R	R	<0.0005	<0.0005
Gag	Positive	A*0301	411	R	K	0.001	0.176
Gag	Negative	A*0205	427	T	T	<0.0005	0.080
Gag	Negative	B*13	437	I	I	0.001	0.168
Gag	Positive	B*13	437	L	I	<0.0005	0.008
Gag	Negative	A*74	441	H	H	<0.0005	<0.0005
Gag	Positive	A*74	441	N	H	<0.0005	<0.0005
Gag	Negative	A*6801	464	R	R	0.001	0.119
Gag	Negative	A*6801	467	E	E	0.001	0.168
Gag	Negative	C*0304	467	E	E	<0.0005	0.106
Gag	Positive	C*0304	467	G	E	<0.0005	0.058
Gag	Positive	C*18	474	P	P	0.001	0.168
Gag	Positive	A*6802	477	D	E	0.001	0.139
Gag	Negative	C*17	478	Q	P	<0.0005	0.063
Gag	Positive	C*04	478	M	P	0.001	0.188
Gag	Negative	C*17	487	T	T	<0.0005	0.019
Gag	Positive	C*17	487	I	T	<0.0005	0.052
Gag	Negative	B*4201	488	S	S	<0.0005	0.101
Gag	Positive	B*4201	488	A	S	<0.0005	0.103
Protease	Negative	B*44	35	E	E	<0.0005	0.004
Protease	Positive	B*44	35	D	E	<0.0005	0.004
Protease	Positive	B*45	63	H	L	<0.0005	0.049

Table A2. Best-defined (A-list) optimal Gag epitopes from the Los Alamos HIV molecular immunology database (<http://www.hiv.lanl.gov/content/immunology/>) [275]

Epitope	Protein	HXB2 codons	Subtype	HLA restriction
GELDRWEKI	p17	11-19		B*4002
KIRLRPGGK	p17	18-26		A*0301
IRLRPGGKK	p17	19-27	B	B*2705
RLRPGGKKK	p17	20-28		A*0301
RLRPGGKKKY	p17	20-29	B	A*0301
GGKKKYKLK	p17	24-32	B	B*0801
KYKCLKHIVW	p17	28-36	B	A*2402
HLVWASREL	p17	33-41		C*0804
LVWASRELERF	p17	34-44		A30
WASRELERF	p17	36-44	B	B*3501
ELRSLYNTV	p17	74-82		B*0801
RSLYNTVATLY	p17	76-86	B	A*3002, B58, B63
SLYNTVATL	p17	77-85	B	A*0201, A*0202, A*0205
LYNTVATL	p17	78-85		C14
LYNTVATLY	p17	78-86		A*2902, B*4403
TLYCVHQK	p17	84-91		A*1101
IEIKDTKEAL	p17	92-101		B*4001
NSSKVSQNY	p17	124-132	B	B*3501
VQNLQGQMV	p24	135-143		B13
HQAISPRTL	p24	144-152		B*1510
QAISPRTLNAW	p24	145-155	B	A*2501
ISPRTLNAW	p24	147-155		B*5701, B63
SPRTLNAWV	p24	148-156		B*0702
VKVIEEKAF	p24	156-164		B*1503
EEKAFSPEV	p24	160-168		B*4415
KAFSPEVI	p24	162-169	B	B*5703
KAFSPEVIPMF	p24	162-172	B	B*5701, B*5703, B63
FSPEVIPMF	p24	164-172		B57
EVIPMFSAL	p24	167-175	B	A*2601
VIPMFSAL	p24	168-175	B	C*0102
SEGATPQDL	p24	176-184		B*4001
TPQDLNTML	p24	180-188	B	B*0702, B*3910, B*4201, B*8101, C*0802
GHQAAMQML	p24	193-201	B	B*1510, B*3901
KETINEEAA	p24	202-210		B*4002
ETINEEAAEW	p24	203-212		A*2501
AEWDRVHPV	p24	210-218		B*4002

HPVHAGPIA	p24	216-224		B07
GQMREPRGSDI	p24	226-236		B13
TSTLQEQIGW	p24	240-249	B	B*5701, B*5801
PPIPVGDIY	p24	254-262	B	B*3501
EIYKRWII	p24	260-267	B	B*0801
RRWIQLGLQK	p24	263-272		B*2703
KRWIILGLNK	p24	263-272	B	B*2705
GLNKIVRMY	p24	269-277	B	B*1501
VRMYSVSI	p24	274-282		C18
RMYSPTSI	p24	275-282		B*5201
FRDYVDRFF	p24	293-301		C18
FRDYVDRFYK	p24	293-302	B, D	B*1801
RDYVDRFFKTL	p24	294-304	A	A*2402
RDYVDRFYKTL	p24	294-304	B	B*4402
YVDRFYKTL	p24	296-304		A*0207
YVDRFFKTL	p24	296-304		B*1503, C*0303, C*0304
DRFYKTLRA	p24	298-306	B	B*1402
AEQASQDVKNW	p24	306-316	B	B*4402
AEQASQEVKNWM	p24	306-317		C05
QASQEVKNW	p24	308-316	B	B*5301, B*5701
DCKTILKAL	p24	329-337	B	B*0801
ACQGVGGPGHK	p24	349-359		A*1101
GPGHKARVL	p24	355-363	B	B*0702
AEAMSQVTNS	p2p7p1p6	364-373		B*4501
CRAPRKKGC	p2p7p1p6	405-413		B14
TERQANFL	p2p7p1p6	427-434		B*4002
RQANFLGKI	p2p7p1p6	429-437		B13
FLGKIWPSYK	p2p7p1p6	433-442		A*0201
KELYPLTSL	p2p7p1p6	481-489		B*4001

HIV-1 – human immunodeficiency virus type 1; HLA – human leukocyte antigen; p17 – protein of 17 kDa; p24 – protein of 24 kDa; p2 – protein of 2 kDa; p7 – protein of 7 kDa; p1 – protein of 1 kDa; p6 – protein of 6 kDa.

Table A3. Amino acids in Gag-protease from HIV-1 subtype C chronically infected subjects associated with altered replication capacities

Protein	Codon	AA	Con.	No. of samples		Replication capacity		Univariate analysis		Multivariate analysis		Optimal epitopes with HLA restriction	HLA association
				-AA	+AA	-AA	+AA	p	q	p	q		
p17	20	R	Yes	42	363	0.58	0.62	0.021	0.57	-	-	KK9/RK9/RLY10 - A*0301, IK9 - B*2705	C*06 ^a , A*74 ^a , A*6801 ^b
p17	28	Q		324	81	0.61	0.64	0.002	0.34	<0.0005	0.23	RK9/RLY10 - A*0301, GK9 - B*0801, KW9 - A*2402	B*13 ^a , B*42 ^a , C*17 ^a
p17	28	R		314	91	0.62	0.64	0.040	0.6	-	-	RK9/RLY10 - A*0301, GK9 - B*0801, KW9 - A*2402	B*13 ^a , B*42 ^a , C*17 ^a
p17	30	R		340	65	0.61	0.64	0.009	0.47	-	-	GK9 - B*0801, KW9 - A*2402	B*42 ^a , C*17 ^a
p17	31	I		355	50	0.62	0.58	0.003	0.34	0.001	0.27	GK9 - B*0801, KW9 - A*2402	C*0701 ^a
p17	35	V	Yes	44	361	0.66	0.62	0.011	0.47	-	-	KW9 - A*2402, HL9 - C*0804, LF11 - A*30	
p17	35	I		355	50	0.62	0.65	0.023	0.57	-	-	KW9 - A*2402, HL9 - C*0804, LF11 - A*30	
p17	49	S		393	12	0.62	0.56	0.036	0.58	-	-		B*1503 ^a
p17	58	Q		400	5	0.62	0.69	0.023	0.57	-	-		
p17	61	I	Yes	157	248	0.61	0.64	0.004	0.41	<0.0005	0.23		
p17	62	S		373	32	0.62	0.68	<0.0005	0.27	<0.0005	0.23		B*1402 ^a
p17	75	L	Yes	24	381	0.65	0.62	0.005	0.27	-	-	EV9 - B*0801	
p17	75	I		366	39	0.61	0.66	0.002	0.47	0.002	0.28	EV9 - B*0801	
p17	76	R	Yes	231	174	0.63	0.6	0.021	0.57	-	-	EV9 - B*0801, RY11 - A*3002, B*58, B*63	
p17	76	K		153	252	0.6	0.63	0.016	0.57	-	-	EV9 - B*0801, RY11 - A*3002, B*58, B*63	
p17	84	A		397	8	0.62	0.7	0.010	0.47	-	-	RY11 - A*3002, B*58, B*63, SL9 - A*0201, A*0202, A*0205, LL8 - C*14, LY9 - A*2902, B*4403, TK8 - A*1101	
p17	84	I		398	7	0.62	0.68	0.024	0.57	-	-	RY11 - A*3002, B*58, B*63, SL9 - A*0201, A*0202, A*0205, LL8 - C*14, LY9 - A*2902, B*4403, TK8 - A*1101	
p17	109	S		393	12	0.62	0.65	0.047	0.65	0.002	0.28		
p17	111	C		319	86	0.62	0.64	0.038	0.59	-	-		
p17	114	E		398	7	0.62	0.57	0.011	0.47	-	-		
p17	115	M		397	8	0.62	0.7	0.007	0.47	-	-		
p17	118	V		397	8	0.62	0.51	0.025	0.57	<0.0005	0.23		
p17	119	Q		398	7	0.62	0.68	0.025	0.57	-	-		
p17	119	T		392	13	0.62	0.65	0.034	0.58	-	-		
p17	120	K		394	11	0.62	0.56	0.008	0.47	-	-		
p17	123	G	Yes	91	314	0.6	0.63	0.020	0.57	-	-		
p17	123	R		391	14	0.62	0.56	0.042	0.63	-	-		

p24	182	Q	Yes	54	351	0.59	0.63	0.031	0.57	-	-	SL9 – B*4001, TL9 – B*0702, B*3910, B*4201, B*8101, C*0802	B*81 ^a , C*17 ^a
p24	182	G		398	7	0.62	0.54	0.024	0.58	-	-	SL9 – B*4001, TL9 – B*0702, B*3910, B*4201, B*8101, C*0802	B*81 ^a , C*17 ^a
p24	186	T	Yes	28	377	0.58	0.62	0.006	0.47	<0.0005	0.23	TL9 – B*0702, B*3910, B84201, B*8101, C*0802	B*81 ^a
p24	186	S		377	28	0.62	0.59	0.012	0.47	-	-	TL9 – B*0702, B*3910, B84201, B*8101, C*0802	B*81 ^a
p24	191	V	Yes	10	395	0.54	0.62	0.005	0.47	-	-		
p24	191	I		392	13	0.62	0.56	0.032	0.58	-	-		
p24	200	M	Yes	6	399	0.57	0.62	0.029	0.57	-	-	GL9 – B*1510, B*3901	
p24	200	I		399	6	0.62	0.57	0.029	0.57	-	-	GL9 – B*1510, B*3901	
p24	223	Y		398	7	0.62	0.55	0.021	0.57	0.003	0.28	HA9 – B*07	A*2911 ^b
p24	224	A	Yes	31	374	0.59	0.62	0.020	0.57	-	-	HA9 – B*07	
p24	224	P		376	29	0.62	0.59	0.046	0.64	-	-	HA9 – B*07	
p24	228	M	Yes	73	332	0.59	0.62	0.027	0.57	-	-	GI11 – B*13	
p24	302	K	Yes	5	400	0.54	0.62	0.042	0.63	-	-	FK10 – B*1801, RL11 – A*2402, B*4402, YL9 – A*0207, B*1503, C*0303, C*0304, DA9 – B*1402	B*14 ^a
p24	309	A	Yes	20	385	0.54	0.62	0.004	0.34	-	-	AW11 – B*4402, AW12 – C*05, QW9 – B*5301, B*5701	B*53 ^a
p24	309	S		387	18	0.62	0.56	0.010	0.47	-	-	AW11 – B*4402, AW12 – C*05, QW9 – B*5301, B*5701	B*53 ^a
p24	312	D	Yes	112	293	0.6	0.63	0.007	0.47	-	-	AW11 – B*4402, AW12 – C*05, QW9 – B*5301, B*5701	B*44 ^a
p24	312	E		269	136	0.62	0.6	0.032	0.58	-	-	AW11 – B*4402, AW12 – C*05, QW9 – B*5301, B*5701	B*44 ^a
p24	323	V	Yes	16	389	0.69	0.62	0.011	0.47	0.001	0.24		
p2	369	H		395	10	0.62	0.74	0.007	0.47	-	-	AS10 – B*4501	
p2	371	S		389	16	0.62	0.57	0.036	0.58	-	-	AS10 – B*4501	A*3002 ^b
p2	374	P		397	8	0.62	0.56	0.046	0.64	-	-		B*1503 ^a
p7	381	G		291	114	0.61	0.64	0.005	0.41	-	-		A*43 ^b
p7	382	K		400	5	0.62	0.49	0.015	0.52	-	-		A*01 ^a , C*18 ^a
p7	385	N		400	5	0.62	0.79	0.024	0.57	-	-		
p7	389	M		390	15	0.62	0.68	0.001	0.27	<0.0005	0.23		
p7	389	T		319	86	0.61	0.63	0.035	0.58	-	-		
p7	397	K	Yes	9	396	0.57	0.62	0.015	0.52	0.004	0.28		
p7	411	K	Yes	8	397	0.54	0.62	0.004	0.34	0.001	0.24	CC9 – B*14	A*0301 ^a
p7	425	D	Yes	11	394	0.68	0.62	0.029	0.5	-	-		
p7	425	E		381	24	0.62	0.65	0.013	0.57	-	-		
p6	453	T		394	11	0.62	0.55	0.012	0.47	-	-		

p6	466	K		400	5	0.62	0.5	0.039	0.59	0.004	0.28		
p6	467	E	Yes	34	371	0.68	0.62	0.007	0.47	-	-		C*0304 ^a , A*6801 ^a
p6	471	N		390	15	0.62	0.57	0.008	0.47	0.001	0.24		
p6	473	P		398	7	0.62	0.71	0.010	0.47	-	-		
p6	473	D		400	5	0.62	0.54	0.044	0.64	-	-		
p6	478	V		399	6	0.62	0.73	0.007	0.47	0.003	0.28		C*17 ^b , C*04 ^a
p6	479	K	Yes	79	326	0.6	0.62	0.028	0.57	0.002	0.28		
p6	481	R	Yes	28	377	0.56	0.62	0.008	0.47	-	-	KL9 – B*4001	
p6	485	H		396	9	0.62	0.68	0.023	0.57	-	-	KL9 – B*4001	
p6	498	W		394	11	0.62	0.67	0.031	0.58	-	-		
Pro	13	V		398	7	0.62	0.55	0.028	0.57	-	-		
Pro	16	G	Yes	13	392	0.56	0.62	0.001	0.27	0.002	0.28		
Pro	16	E		387	18	0.62	0.57	0.023	0.57	-	-		
Pro	20	K	Yes	73	332	0.58	0.62	0.007	0.47	-	-		
Pro	20	R		318	87	0.62	0.59	0.039	0.59	-	-		
Pro	35	E	Yes	142	263	0.61	0.62	0.042	0.62	0.003	0.28	EW9 – B*44; DL9 – A*6802	B*44 ^a
Pro	61	Q	Yes	29	376	0.65	0.62	0.018	0.41	0.002	0.28	RI10 – B*13	
Pro	61	E		381	24	0.62	0.67	0.005	0.57	-	-	RI10 – B*13	
Pro	64	M		400	5	0.62	0.73	0.024	0.57	0.001	0.24	RI10 – B*13	
Pro	70	R		391	14	0.62	0.66	0.029	0.57	-	-	GL9 – B*1503; KV8 – B*57	
Pro	74	S		361	44	0.62	0.64	0.042	0.63	-	-	GL9 – B*1503; KV8 – B*57	
Pro	89	L		351	54	0.61	0.65	0.006	0.47	-	-	TL11 – B*81	

HIV-1 – human immunodeficiency virus type 1; p17 – protein of 17 kDa; p24 – protein of 24 kDa; p2 – protein of 2 kDa; p7 – protein of 7 kDa; p6 – protein of 6 kDa; pro – protease; AA – amino acid; con. – consensus; HLA – human leukocyte antigen.

^aHLA allele is negatively associated with the consensus amino acid or positively associated with a polymorphism at that codon.

^bHLA allele is positively associated with the consensus amino acid or negatively associated with a polymorphism at that codon.

REFERENCES

1. Fauci, A.S., Lane, H.C. 2005. Human Immunodeficiency Virus Disease: AIDS and Related Disorders. Pages 1076-1139 in Kasper, D.L., Braunwald, E., Fauci, A.S., Hauser, S.L., Longo, D.L., Jameson, J.L., eds. *Harrison's Principles of Internal Medicine. 16th Edition*. New York: McGraw-Hill Professional.
2. Freed, E.O., Martin, M.A. 2007. HIVs and Their Replication. Pages 2107-2185 in Knipe, D.M., Howley, P.M., Griffin, D.E., Lamb, R.A., Martin, M.A., Roizman, B., Straus, S.E., eds. *Fields Virology. 5th Edition*. Philadelphia: Lippincott Williams & Wilkins.
3. Sharp, P.M., Hahn, B.H. 2008. Prehistory of HIV-1. *Nature* 455: 605-606.
4. Barre-Sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W., Montagnier, L. 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 220(4599): 868-871.
5. Levy, J.A., Hoffman, A.D., Kramer, S.M., Landis, J.A., Shimabukuro, J.M., Oshiro, L.S. 1984. Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. *Science* 225: 840-842.
6. Popovic, M., Sarngadharan, M.G., Read, E., Gallo, R.C. 1984. Detection, isolation and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. *Science* 224(4648): 497-500.
7. Sarngadharan, M.G., Popovic, M., Bruch, L., Schupbach, J., Gallo, R.C. 1984. Antibodies reactive with human T-lymphotropic retroviruses (HTLV-III) in the serum of patients with AIDS. *Science* 224(4648): 506-508.
8. Tebit, D.M., Arts, E.J. 2011. Tracking a century of global expansion and evolution of HIV to drive understanding and to combat disease. *The Lancet Infectious Diseases* 11: 45-56.
9. Plantier, J., Leoz, M., Dickerson, J.E., De Oliveira, F., Cordonnier, F., Lemée, V., Damond, F., Robertson, D.L., Simon, F. 2009. A new human immunodeficiency virus derived from gorillas. *Nature Medicine* 15(8): 871-872.
10. World Health Organisation. HIV/AIDS, Data and Statistics. <http://www.who.int/hiv/data/en/index.html> [Accessed 15/02/2011].
11. Nawaz, F., Cicala, C., Van Ryk, D., Block, K.E., Jelicic, K., McNally, J.P., Ogundare, O., Pascuccio, M., Patel, N., Wei, D., Fauci, A.S., Arthos, J. 2011. The Genotype of Early-Transmitting HIV gp120s Promotes $\alpha 4\beta 7$ –Reactivity, Revealing $\alpha 4\beta 7$ +/CD4+ T cells As Key Targets in Mucosal Transmission. *PLoS Pathogens* 7(2): e1001301.
12. Korber, B., Foley, B., Kuiken, C., Pillai, S., Sodroski, J. 1998. Numbering Positions in HIV Relative to HXB2CG. Pages III-102-III-111 in Korber, B., Kuiken, C., Foley, B., Hahn, B., McCutchan, F., Mellors, J., Sodroski, J., eds. *Human Retroviruses and AIDS*. Los Alamos, New Mexico: Los Alamos National Laboratory, Theoretical Biology and Biophysics.
13. Freed, E.O. 2001. HIV-1 Replication. *Somatic Cell and Molecular Genetics* 26(1-6): 13-33.
14. Bukrinskaya, A. 2007. HIV-1 matrix protein: A mysterious regulator of the viral life cycle. *Virus Research* 124: 1-11.
15. Mateu, M.G. 2009. The capsid protein of human immunodeficiency virus: intersubunit interactions during virus assembly. *The FEBS Journal* 276: 6098-6109.

16. Fiorentini, S., Giagulli, C., Caccuri, F., Magiera, A.K., Caruso, A. 2010. HIV-1 matrix protein p17: A candidate antigen for therapeutic vaccines against AIDS. *Pharmacology and Therapeutics* 128: 433-444.
17. Gamble, T.R., Vajdos, F.F., Yoo, S., Worthylake, D.K., Houseweart, M., Sundquist, W.I., Hill, C.P. 1996. Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. *Cell* 87: 1285-1294.
18. Freed, E.O. 1998. HIV-1 Gag Proteins: Diverse Functions in the Virus Life Cycle. *Virology* 251: 1-15.
19. Gamble, T.R., Yoo, S., Vajdos, F.F., von Schwedler, U.K., Worthylake, D.K., Wang, H., McCutcheon, J.P., Sundquist, W.I., Hill, C.P. 1997. Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. *Science* 278: 849-853.
20. South, T.L., Summers, M.F. 1993. Zinc- and sequence- dependent binding to nucleic acids by the N-terminal zinc finger of the HIV-1 nucleocapsid protein: NMR structure of the complex with the Psi-site analog, dACGCC. *Protein Science : a Publication of the Protein Society* 2: 3-19.
21. Muriaux, D., Darlix, J. 2010. Properties and functions of the nucleocapsid protein in virus assembly. *RNA Biology* 7(6): 744-753.
22. Gomez, C., Hope, T.J. 2005. The ins and outs of HIV replication. *Cellular Microbiology* 7(5): 621-626.
23. Markosyan, R.M., Cohen, F.S., Melikyan, G.B. 2003. HIV-1 envelope proteins complete their folding into six-helix bundles immediately after fusion pore formation. *Molecular Biology of the Cell* 14(3): 926-938.
24. Wyma, D.J., Jiang, J., Shi, J., Zhou, J., Lineberger, J.E., Miller, M.D., Aiken, C. 2004. Coupling of human immunodeficiency virus type 1 fusion to virion maturation: a novel role of the gp41 cytoplasmic tail. *Journal of Virology* 78(8): 3429-3435.
25. Murakami, T. 2008. Roles of the interactions between Env and Gag proteins in the HIV-1 replication cycle. *Microbiology and Immunology* 52: 287-295.
26. Arhel, N.J., Souquere-Besse, S., Munier, S., Souque, P., Guadagnini, S., Rutherford, S., Prévost, M., Allen, T.D., Charneau, P. 2007. HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore. *The EMBO Journal* 26: 3025-3037.
27. Mascarenhas, A.P., Musier-Forsyth, K. 2009. The capsid protein of human immunodeficiency virus: interactions of HIV-1 capsid with host protein factors. *The FEBS Journal* 276: 6118-6127.
28. Buckman, J.S., Bosche, W.J., Gorelick, R.J. 2003. Human immunodeficiency virus type 1 nucleocapsid zn(2+) fingers are required for efficient reverse transcription, initial integration processes, and protection of newly synthesized viral DNA. *Journal of Virology* 77(2): 1469-1480.
29. Berthoux, L., Péchoux, C., Ottmann, M., Morel, G., Darlix, J.L. 1997. Mutations in the N-terminal domain of human immunodeficiency virus type 1 nucleocapsid protein affect virion core structure and proviral DNA synthesis. *Journal of Virology* 71(9): 6973-6981.
30. Kiernan, R.E., Ono, A., Englund, G., Freed, E.O. 1998. Role of matrix in an early postentry step in the human immunodeficiency virus type 1 life cycle. *Journal of Virology* 72(5): 4116-4126.
31. Strebel, K., Luban, J., Jeang, K. 2009. Human cellular restriction factors that target HIV-1 replication. *BMC Medicine* 7:48.
32. Forshey, B.M., von Schwedler, U., Sundquist, W.I., Aiken, C. 2002. Formation of a HIV-1 core of optimal stability is crucial for viral replication. *Journal of Virology* 76: 5667-5677.

33. Brockman, M.A., Schneidewind, A., Lahaie, M., Schmidt, A., Miura, T., DeSouza, I., Ryvkin, F., Derdeyn, C.A., Allen, S., Hunter, E., Mulenga, J., Goepfert, P.A., Walker, B.D., Allen, T.M. 2007. Escape and Compensation from Early HLA-B57-Mediated Cytotoxic T-Lymphocyte Pressure on Human Immunodeficiency Virus Type 1 Gag Alter Capsid Interactions with Cyclophilin A. *Journal of Virology* 81(22): 12608–12618.
34. Luban, J. 2007. Cyclophilin A, TRIM5, and Resistance to Human Immunodeficiency Virus Type 1 Infection. *Journal of Virology* 81(3): 1054-1061.
35. Stremlau, M., Owens, C.M., Perron, M.J., Kiessling, M., Autissier, P., Sodroski, J. 2004. The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. *Nature* 427: 848-853.
36. Kirchhoff, F. 2010. Immune Evasion and Counteraction of Restriction Factors by HIV-1 and Other Primate Lentiviruses. *Cell Host and Microbe* 8(1): 55-67.
37. Towers, G.J., Hatzioannou, T., Cowan, S., Goff, S.P., Luban, J., Bieniasz, P.D. 2003. Cyclophilin A modulates the sensitivity of HIV-1 to host restriction factors. *Nature Medicine* 9: 1138-1143.
38. Stremlau, M., Song, B., Javanbakht, H., Perron, M., Sodroski, J. 2006. Cyclophilin A: an auxiliary but not necessary cofactor for TRIM5alpha restriction of HIV-1. *Virology* 351: 112-120.
39. Ikeda, Y., Ylinen, L.M.J., Kahar-Bador, M., Towers, G.J. 2004. Influence of gag on Human Immunodeficiency Virus Type 1 Species-Specific Tropism. *Journal of Virology* 78(21): 11816-11822.
40. Sorin, M., Kalpana, G.V. 2006. Dynamics of Virus-Host Interplay in HIV-1 Replication. *Current HIV Research* 4: 117-130.
41. Haffar, O.K., Popov, S., Dubrovsky, L., Agostini, I., Tang, H., Pushkarsky, T., Nadler, S.G., Bukrinsky, M. 2000. Two Nuclear Localization Signals in the HIV-1 Matrix Protein Regulate Nuclear Import of the HIV-1 Pre-integration Complex. *Journal of Molecular Biology* 299: 359-368.
42. Mannioui, A., Nelson, E., Schiffer, C., Felix, N., Le Rouzic, E., Benichou, S., Gluckman, J.C., Canque, B. 2005. Human immunodeficiency virus type 1 KK26-27 matrix mutants display impaired infectivity, circularization and integration but not nuclear import. *Virology* 339: 21-30.
43. Felzien, L.K., Woffendin, C., Hottiger, M.O., Subbramanian, R.A., Cohen, E.A., Nabel, G.J. 1998. HIV transcriptional activation by the accessory protein, VPR, is mediated by the p300 co-activator. *Proceedings of the National Academy of Sciences of the United States of America* 95(9): 5281-5286.
44. Dupont, S., Sharova, N., DéHoratius, C., Virbasius, C.M., Zhu, X., Bukrinskaya, A.G., Stevenson, M., Green, M.R. 1999. A novel nuclear export activity in HIV-1 matrix protein required for viral replication. *Nature* 402(6762): 681-685.
45. Leung, K., Kim, J.O., Ganesh, L., Kabat, J., Schwartz, O., Nabel, G.J. 2008. HIV-1 assembly: viral glycoproteins segregate quantally to lipid rafts that associate individually with HIV-1 capsids and virions. *Cell Host and Microbe* 3(5): 285-292.
46. Dorfman, T., Mammano, F., Haseltine, W.A., Göttlinger, H.G. 1994. Role of the matrix protein in the virion association of the human immunodeficiency virus type 1 envelope glycoprotein. *Journal of Virology* 68(3): 1689-1696.
47. Freed, E.O., Martin, M.A. 1995. Virion incorporation of envelope glycoproteins with long but not short cytoplasmic tails is blocked by specific, single amino acid substitutions in the human immunodeficiency virus type 1 matrix. *Journal of Virology* 69: 1984-1989.

48. Freed, E.O., Martin, M.A. 1996. Domains of the human immunodeficiency virus type 1 matrix and gp41 cytoplasmic tail required for envelope incorporation into virions. *Journal of Virology* 70: 341-351.
49. Ono, A., Huang, M., Freed, E.O. 1997. Characterization of human immunodeficiency virus type 1 matrix revertants: effects on virus assembly, Gag processing, and Env incorporation into virions. *Journal of Virology* 71(6): 4409-4418.
50. Lopez-Vergès, S., Camus, G., Blot, G., Beauvoir, R., Benarous, R., Berlioz-Torrent, C. 2006. Tail-interacting protein TIP47 is a connector between Gag and Env and is required for Env incorporation into HIV-1 virions. *Proceedings of the National Academy of Sciences of the United States of America* 103(40): 14947-14952.
51. Lehmann, M., Milev, M.P., Abrahamyan, L., Yao, X.J., Pante, N., Mouland, A.J. 2009. Intracellular transport of human immunodeficiency virus type 1 genomic RNA and viral production are dependent on dynein motor function and late endosome positioning *The Journal of Biological Chemistry* 284: 14572-14585.
52. Molle, D., Segura-Morales, C., Camus, G., Berlioz-Torrent, C., Kjemis, J., Basyuk, E., Bertrand, E. 2009. Endosomal trafficking of HIV-1 gag and genomic RNAs regulates viral egress. *The Journal of Biological Chemistry* 284: 19727-19743.
53. Scarlata, S., Carter, C. 2003. Role of HIV-1 Gag domains in viral assembly. *Biochimica et Biophysica Acta* 1614: 62-72.
54. Hill, C.P., Worthylake, D., Bancroft, D.P., Christensen, A.M., Sundquist, W.I. 1996. Crystal structures of the trimeric human immunodeficiency virus type 1 matrix protein: Implications for membrane association and assembly. *Proceedings of the National Academy of Sciences of the United States of America* 93: 3099-3104.
55. Dong, X., Li, H., Derdowski, A., Ding, L., Burnett, A., Chen, X., Peters, T.R., Dermody, T.S., Woodruff, E., Wang, J., Spearman, P. 2005. AP-3 Directs the Intracellular Trafficking of HIV-1 Gag and Plays a Key Role in Particle Assembly. *Cell* 120: 663-674.
56. Batonick, M., Favre, M., Boge, M., Spearman, P., Honing, S., Thali, M. 2005. Interaction of HIV-1 Gag with the clathrin-associated adaptor AP-2. *Virology* 342: 190-200.
57. Freed, E.O., Orenstein, J.M., Buckler-White, A.J., Martin, M.A. 1994. Single amino acid changes in the human immunodeficiency virus type 1 matrix protein block virus particle production. *Journal of Virology* 68(8): 5311-5320.
58. Zhou, W., Parent, L.J., Wills, J.W., Resh, M.D. 1994. Identification of a membrane-binding domain within the amino-terminal region of human immunodeficiency virus type 1 Gag protein which interacts with acidic phospholipids. *Journal of Virology* 68(4): 2556-2569.
59. Ono, A. 2009. HIV-1 Assembly at the Plasma Membrane: Gag Trafficking and Localization. *Future Virology* 4(3): 241-257.
60. Fledderman, E.L., Fujii, K., Ghanam, R.H., Waki, K., Prevelige, P.E., Freed, E.O., Saad, J.S. 2010. Myristate Exposure in the Human Immunodeficiency Virus Type 1 Matrix Protein is Modulated by pH. *Biochemistry* 49: 9551-9562.
61. Ono, A., Freed, E.O. 1999. Binding of Human Immunodeficiency Virus Type 1 Gag to Membrane: Role of the Matrix Amino Terminus. *Journal of Virology* 73(5): 4136-4144.
62. Bukrinskaya, A.G., Ghorpade, A., Heinzinger, N.K., Smithgall, T.E., Lewis, R.E., Stevenson, M. 1996. Phosphorylation-dependent human immunodeficiency virus type 1 infection and nuclear targeting of viral DNA. *Proceedings of the National Academy of Sciences of the United States of America* 93(1): 367-371.

63. Kaushik, R., Ratner, L. 2004. Role of human immunodeficiency virus type 1 matrix phosphorylation in an early postentry step of virus replication. *Virology* 78(5): 2319-2326.
64. Votteler, J., Neumann, L., Hahn, S., Hahn, F., Rauch, P., Schmidt, K., Studtrucker, N., Solbak, S.M., Fossen, T., Henklein, P., Ott, D.E., Holland, G., Bannert, N., Schubert, U. 2011. Highly conserved serine residue 40 in HIV-1 p6 regulates capsid processing and virus core assembly. *Retrovirology* 8:11.
65. Dussupt, V., Sette, P., Bello, N.F., Javid, M.P., Nagashima, K., Bouamr, F. 2011. Basic Residues in the Nucleocapsid Domain of Gag Are Critical for Late Events of HIV-1 Budding. *Journal of Virology* 85(5): 2304-2315.
66. Garrus, J.E., von Schwedler, U.K., Pornillos, O.W., Morham, S.G., Zavitz, K.H., Wang, H.E., Wettstein, D.A., Stray, K.M., Côté, M., Rich, R.L., Myszka, D.G., Sundquist, W.I. 2001. Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell* 107(1): 55-65.
67. Strack, B., Calistri, A., Craig, S., Popova, E., Göttlinger, H.G. 2003. AIP1/ALIX Is a Binding Partner for HIV-1 p6 and EIAV p9 Functioning in Virus Budding. *Cell* 114: 689-699.
68. Fujii, K., Munshi, U.M., Ablan, S.D., Demirov, D.G., Soheilian, F., Nagashima, K., Stephen, A.G., Fisher, R.J., Freed, E.O. 2009. Functional role of Alix in HIV-1 replication. *Virology* 391(2): 284-292.
69. Dorfman, T., Bukovsky, A., Öhagen, A., Höglund, S., Göttlinger, H.G. 1994. Functional Domains of the Capsid Protein of Human Immunodeficiency Virus Type 1. *Journal of Virology* 68(12): 8180-8187.
70. von Schwedler, U.K., Stray, K.M., Garrus, J.E., Sundquist, W.I. 2003. Functional Surfaces of the Human Immunodeficiency Virus Type 1 Capsid Protein. *Journal of Virology* 77: 5439-5450.
71. Brockman, M.A., Tanzi, G.O., Walker, B.D., Allen, T.M. 2006. Use of a novel GFP reporter cell line to examine replication capacity of CXCR4- and CCR5-tropic HIV-1 by flow cytometry. *Journal of Virological Methods* 131 134-142.
72. McMichael, A.J., Borrow, P., Tomaras, G.D., Goonetilleke, N., Haynes, B.F. 2010. The immune response during acute HIV-1 infection: clues for vaccine development. *Nature Reviews. Immunology* 10(1): 11-23.
73. Stevenson, M. 2003. HIV-1 pathogenesis. *Nature Medicine* 9(7): 853-860.
74. Kuritzkes, D.R., Walker, B.D. 2007. HIV-1: Pathogenesis, Clinical Manifestations, and Treatment. Pages 2187-2214 in Knipe, D.M., Howley, P.M., Griffin, D.E., Lamb, R.A., Martin, M.A., Roizman, B., Straus, S.E., eds. *Fields Virology. 5th Edition*. Philadelphia: Lippincott Williams & Wilkins.
75. Haase, A.T. 2010. Targeting early infection to prevent HIV-1 mucosal transmission. *Nature* 464: 217-223.
76. Sagar, M. 2010. HIV-1 Transmission Biology: Selection and Characteristics of Infecting Viruses. *Journal of Infectious Diseases* 202(S2): S289-S296.
77. Derdeyn, C.A., Hunter, E. 2008. Viral characteristics of transmitted HIV. *Current Opinion in HIV and AIDS* 3(1): 16-21.
78. Collman, R.G., Perno, C., Crowe, S.M., Stevenson, M., Montaner, L.J. 2003. HIV and cells of macrophage/dendritic lineage and other non-T cell reservoirs: new answers yield new questions. *Journal of Leukocyte Biology* 74: 631-634.
79. Derdeyn, C.A., Silvestri, G. 2005. Viral and host factors in the pathogenesis of HIV infection. *Current Opinion in Immunology* 17(4): 366-373.
80. Lee, H.Y., Giorgi, E.E., Keele, B.F., Gaschen, B., Athreya, G.S., Salazar-Gonzalez, J.F., Pham, K.T., Goepfert, P.A., Kilby, J.M., Saag, M.S., Delwart, E.L., Busch, M.P., Hahn, B.H., Shaw, G.M., Korber, B.T., Bhattacharya, T., Perelson, A.S. 2009.

- Modeling sequence evolution in acute HIV-1 infection. *Journal of Theoretical Biology* 261: 341-360.
81. Fiebig, E.W., Wright, D.J., Rawal, B.D., Garrett, P.E., Schumacher, R.T., Peddada, L., Heldebrant, C., Smith, R., Conrad, A., Kleinman, S.H., Busch, M.P. 2003. Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. *AIDS* 17: 1871-1879.
 82. Streeck, H., Nixon, D.F. 2010. T Cell Immunity in Acute HIV-1 Infection. *Journal of Infectious Diseases* 202(S2): S302-S308.
 83. Lewin, S.R., Evans, V.A., Elliott, J.H., Spire, B., Chomont, N. 2011. Finding a cure for HIV: will it ever be achievable? *Journal of the International AIDS Society* 14:4.
 84. Groot, F., Welsch, S., Sattentau, Q.J. 2008. Efficient HIV-1 transmission from macrophages to T cells across transient virological synapses. *Blood* 111(9): 4660-4663.
 85. Appay, V., Sauce, D. 2008. Immune activation and inflammation in HIV-1 infection: causes and consequences. *The Journal of Pathology* 214: 231-241.
 86. Sodora, D.L., Allan, J.S., Apetrei, C., Brenchley, J.M., Douek, D.C., Else, J.G., Estes, J.D., Hahn, B.H., Hirsch, V.M., Kaur, A., Kirchhoff, F., Muller-Trutwin, M., Pandrea, I., Schmitz, J.E., Silvestri, G. 2009. Toward an AIDS vaccine: lessons from natural simian immunodeficiency virus infections of African nonhuman primate hosts. *Nature Medicine* 15(8): 861-865.
 87. Cadogan, M., Dalgleish, A.G. 2008. Pathogenesis of HIV: non-specific immune hyperactivity and its implications for vaccines. *Clinical Medicine* 8(3): 267-271.
 88. Gonzalez, V.D., Landay, A.L., Sandberg, J.K. 2010. Innate immunity and chronic immune activation in HCV/HIV-1 co-infection. *Clinical Immunology* 135: 12-25.
 89. Chang, J.J., Altfeld, M. 2010. Innate Immune Activation in Primary HIV-1 Infection. *Journal of Infectious Diseases* 202(S2): S297-S301.
 90. Chakrabarti, L.A., Simon, V. 2010. Immune mechanisms of HIV control. *Current Opinion in Immunology* 22: 488-496.
 91. Janeway, C.A., Travers, P., Walport, M., Shlomchik, M.J. 2005. *Immunobiology. The immune system in health and disease. 6th Edition*. New York: Garland Science.
 92. Cohen, G.B., Gandhi, R.T., Davis, D.M., Mandelboim, O., Chen, B.K., Strominger, J.L., Baltimore, D. 1999. The selective downregulation of class I major histocompatibility complex proteins by HIV-1 protects HIV-infected cells from NK cells. *Immunity* 10(6): 661-671.
 93. Alter, G., Moody, M.A. 2010. The Humoral Response to HIV-1: New Insights, Renewed Focus. *Journal of Infectious Diseases* 202(S2): S315-S322.
 94. Walker, B.D., Burton, D.R. 2008. Towards an AIDS Vaccine. *Science* 320(5877): 760-764.
 95. Koup, R.A., Safrit, J.T., Cao, Y., Andrews, C.A., McLeod, G., Borkowsky, W., Farthing, C., Ho, D.D. 1994. Temporal association of cellular immune responses with the initial control of viremia in primary human immunodeficiency virus type 1 syndrome. *Journal of Virology* 68(7): 4650-4655.
 96. Jin, X., Bauer, D.E., Tuttleton, S.E., Lewin, S., Gettie, A., Blanchard, J., Irwin, C.E., Safrit, J.T., Mittler, J., Weinberger, L., Kostrikis, L.G., Zhang, L., Perelson, A.S., Ho, D.D. 1999. Dramatic rise in plasma viremia after CD8(+) T cell depletion in simian immunodeficiency virus-infected macaques. *Journal of Experimental Medicine* 189(6): 991-998.
 97. Schmitz, J.E., Kuroda, M.J., Santra, S., Sasseville, V.G., Simon, M.A., Lifton, M.A., Racz, P., Tenner-Racz, K., Dalesandro, M., Scallon, B.J., Ghayeb, J., Forman, M.A., Montefiori, D.C., Rieber, E.P., Letvin, N.L., Reimann, K.A. 1999. Control of viremia

- in simian immunodeficiency virus infection by CD8+ lymphocytes. *Science* 283(5403): 857-860.
98. Schmitz, J.E., Kuroda, M.J., Santra, S., Simon, M.A., Lifton, M.A., Lin, W., Khunhkhun, R., Piatak, M., Lifson, J.D., Grosschupff, G., Gelman, R.S., Racz, P., Tenner-Racz, K., Mansfield, K.A., Letvin, N.L., Montefiori, D.C., Reimann, K.A. 2003. Effect of humoral immune responses on controlling viremia during primary infection of rhesus monkeys with simian immunodeficiency virus. *Journal of Virology* 77(3): 2165-2173.
 99. Klatt, N.R., Shudo, E., Ortiz, A.M., Engram, J.C., Paiardini, M., Lawson, B., Miller, M.D., Else, J., Pandrea, I., Estes, J.D., Apetrei, C., Schmitz, J.E., Ribeiro, R.M., Perelson, A.S., Silvestri, G. 2010. CD8+ Lymphocytes Control Viral Replication in SIVmac239-Infected Rhesus Macaques without Decreasing the Lifespan of Productively Infected Cells. *PLoS Pathogens* 6(1): e1000747.
 100. Wong, J.K., Strain, M.C., Porrata, R., Reay, E., Sankaran-Walters, S., Ignacio, C.C., Russell, T., Pillai, S.K., Looney, D.J., Dandekar, S. 2010. In vivo CD8+ T-cell suppression of siv viremia is not mediated by CTL clearance of productively infected cells. *PLoS Pathogens* 6(1): e1000748.
 101. European Bioinformatics Institute. IMGT/HLA Database. <http://www.ebi.ac.uk/imgt/hla> [Accessed 08/03/2011].
 102. Draenert, R., Le Gall, S., Pfafferott, K.J., Leslie, A.J., Chetty, P., Brander, C., Holmes, E.C., Chang, S.C., Feeney, M.E., Addo, M.M., Ruiz, L., Ramduth, D., Jeena, P., Altfeld, M., Thomas, S., Tang, Y., Verrill, C.L., Dixon, C., Prado, J.G., Kiepiela, P., *et al.* 2004. Immune selection for altered antigen processing leads to cytotoxic T lymphocyte escape in chronic HIV-1 infection. *Journal of Experimental Medicine* 199(7): 905-915.
 103. Jones, N.A., Wei, X., Flower, D.R., Wong, M., Michor, F., Saag, M.S., Hahn, B.H., Nowak, M.A., Shaw, G.M., Borrow, P. 2004. Determinants of human immunodeficiency virus type 1 escape from the primary CD8+ cytotoxic T lymphocyte response. *Journal of Experimental Medicine* 200(10): 1243-1256.
 104. Troyer, R.M., McNevin, J., Liu, Y., Zhang, S.C., Krizan, R.W., Abraha, A., Tebit, D.M., Zhao, H., Avila, S., Lobritz, M.A., McElrath, M.J., Le Gall, S., Mullins, J.I., Arts, E.J. 2009. Variable Fitness Impact of HIV-1 Escape Mutations to Cytotoxic T Lymphocyte (CTL) Response. *PLoS Pathogens* 5(4): 1-13.
 105. Allen, T.M., Altfeld, M., Yu, X.G., O'Sullivan, K.M., Lichterfeld, M., Le Gall, S., John, M., Mothe, B.R., Lee, P.K., Kalife, E.T., Cohen, D.E., Freedberg, K.A., Strick, D.A., Johnston, M.N., Sette, A., Rosenberg, E.S., Mallal, S.A., Goulder, P.J.R., Brander, C., Walker, B.D. 2004. Selection, Transmission, and Reversion of an Antigen-Processing Cytotoxic T-lymphocyte Escape Mutation in Human Immunodeficiency Virus Type 1 Infection. *Journal of Virology* 78(13): 7069-7078.
 106. Goonetilleke, N., Liu, M.K.P., Salazar-Gonzalez, J.F., Ferrari, G., Giorgi, E., Gantsov, V.V., Keele, B.F., Learn, G.H., Turnbull, E.L., Salazar, M.G., Weinhold, K.J., Moore, S., CHAVI Clinical Core B, Letvin, N., Haynes, B.F., Cohen, M.S., Hraber, P., Bhattacharya, T., Borrow, P., Perelson, A.S., *et al.* 2009. The first T cell response to transmitted/founder virus contributes to the control of acute viremia in HIV-1 infection. *The Journal of Experimental Medicine* 206(6): 1253-1272.
 107. Asquith, B., Edwards, C.T., Lipstich, M., Mclean, A.R. 2006. Inefficient cytotoxic T lymphocyte-mediated killing of HIV-1-infected cells *in vivo*. *PLoS Biology* 4(4): e90.
 108. Addo, M.M., Yu, X.G., Rathod, A., Cohen, D., Eldridge, R.L., Strick, D., Johnston, M.N., Corcoran, C., Wurcel, A.G., Fitzpatrick, C.A., Feeney, M.E., Rodriguez, W.R., Basgoz, N., Draenert, R., Stone, D.R., Brander, C., Goulder, P.J., Rosenberg, E.S., Altfeld, M., Walker, B.D. 2003. Comprehensive epitope analysis of human

- immunodeficiency virus type 1 (HIV-1)-specific T-cell responses directed against the entire expressed HIV-1 genome demonstrate broadly directed responses, but no correlation to viral load. *Journal of Virology* 77: 2081-2092.
109. Sun, J.C., Williams, M.A., Bevan, M.J. 2004. CD4⁺ T cells are required for the maintenance, not programming, of memory CD8⁺ T cells after acute infection. *Nature Immunology* 5(9): 927-933.
 110. Douek, D.C., Brenchley, J.M., Betts, M.R., Ambrozak, D.R., Hill, B.J., Okamoto, Y., Casazza, J.P., Kuruppu, J., Kunstman, K., Wolinsky, S., Grossman, Z., Dybul, M., Oxenius, A., Price, D.A., Connors, M., Koup, R.A. 2002. HIV preferentially infects HIV-specific CD4⁺ T cells. *Nature* 417(6884): 95-98.
 111. Schnittman, S.M., Lane, H.C., Greenhouse, J., Justement, J.S., Baseler, M., Fauci, A.S. 1990. Preferential infection of CD4⁺ memory T cells by human immunodeficiency virus type 1: Evidence for a role in the selective T-cell functional defects observed in infected individuals. *Proceedings of the National Academy of Sciences of the United States of America* 87: 6058-6062.
 112. Esté, J.A., Cihlar, T. 2010. Current status and challenges of antiretroviral research and therapy. *Antiviral Research* 85: 25-33.
 113. Chen, L.F., Hoy, J., Lewin, S.R. 2007. New Drugs, Old Drugs. *The Medical Journal of Australia* 186(3): 146-151.
 114. Hutter, G., Nowak, D., Mossner, M., Ganepola, S., Mussig, A., Allers, K., Schneider, T., Hofmann, J., Kucherer, C., Blau, O., Blau, I.W., Hofmann, W.K., Thiel, E. 2009. Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *New England Journal of Medicine* 360: 692-698.
 115. Osterberg, L., Blaschke, T. 2005. Adherence to medication. *New England Journal of Medicine* 353: 487-497.
 116. Lohse, N., Hansen, A.B., Pedersen, G., Kronborg, G., Gerstoft, J., Sorensen, H.T., Vaeth, M., Obel, N. 2007. Survival of persons with and without HIV infection in Denmark, 1995-2005. *Annals of Internal Medicine* 146: 87-95.
 117. Wilkin, T.J., Shalev, N., Tieu, H., Hammer, S.M. 2010. Advances in Antiretroviral Therapy. *Topics in HIV medicine : a publication of the International AIDS Society, USA* 18(2): 66-92.
 118. Holt, N., Wang, J., Kim, K., Friedman, G., Wang, X., Taupin, V., Crooks, G.M., Kohn, D.B., Gregory, P.D., Holmes, M.C., Cannon, P.M. 2010. Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. *Nature Biotechnology* 28(8): 839-847.
 119. McElrath, M.J., Haynes, B.F. 2010. Induction of Immunity to Human Immunodeficiency Virus Type-1 by Vaccination. *Immunity* 33(4): 542-554.
 120. Abdool Karim, Q., Abdool Karim, S.S., Frohlich, J.A., Grobler, A.C., Baxter, C., Mansoor, L.E., Kharsany, A.B., Sibeko, S., Mlisana, K.P., Omar, Z., Gengiah, T.N., Maarschalk, S., Arulappan, N., Mlotshwa, M., Morris, L., Taylor, D., CAPRISA 004 Trial Group. 2010. Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women. *Science* 329(5996): 1168-1174.
 121. Allen, T.M., Altfeld, M. 2008. Crippling HIV one mutation at a time. *Journal of Experimental Medicine* 205(5): 1003-1007.
 122. Watkins, D.I., Burton, D.R., Kallas, E.G., Moore, J.P., Koff, W.C. 2008. Nonhuman primate models and the failure of the Merck HIV-1 vaccine in humans. *Nature Medicine* 14(6): 617-621.
 123. Johnston, M.I., Fauci, A.S. 2008. An HIV Vaccine — Challenges and Prospects. *New England Journal of Medicine* 359(9): 888-890.

124. Buchbinder, S.P., Mehrotra, D.V., Duerr, A., Fitzgerald, D.W., Mogg, R., Li, D., Gilbert, P.B., Lama, J.R., Marmor, M., Del Rio, C., McElrath, M.J., Casimiro, D.R., Gottesdiener, K.M., Chodakewitz, J.A., Corey, L., Robertson, M.N., Step Study Protocol Team. 2008. Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial. *Lancet* 372(9653): 1881-1893.
125. Gray, G., Buchbinder, S., Duerr, A. 2010. Overview of STEP and Phambili trial results: two phase IIb test-of-concept studies investigating the efficacy of MRK adenovirus type 5 gag/pol/nef subtype B HIV vaccine. *Current Opinion in HIV and AIDS* 5(5): 357-361.
126. Rerks-Ngarm, S., Pitisuttithum, P., Nitayaphan, S., Kaewkungwal, J., Chiu, J., Paris, R., Prem Sri, N., Namwat, C., de Souza, M., Adams, E., Benenson, M., Gurunathan, S., Tartaglia, J., McNeil, J.G., Francis, D.P., Stablein, D., Birx, D.L., Chunsuttiwat, S., Khamboonruang, C., Thongcharoen, P., *et al.* 2009. Vaccination with ALVAC and AIDSVAX to Prevent HIV-1 Infection in Thailand. *The New England Journal of Medicine* 361(23): 2209-2220.
127. McElrath, M.J. 2010. Immune Responses to HIV Vaccines and Potential Impact on Control of Acute HIV-1 Infection. *Journal of Infectious Diseases* 202(S2): S323-S326.
128. Deeks, S.G., Walker, B.D. 2007. Human immunodeficiency virus controllers: mechanisms of durable virus control in the absence of antiretroviral therapy. *Immunity* 27(3): 406-416.
129. Chatterjee, K. 2010. Host genetic factors in susceptibility to HIV-1 infection and progression to AIDS. *Journal of Genetics* 89(1): 109-116.
130. Mellors, J.W., Munoz, A., Giorgi, J.V., Margolick, J.B., Tassoni, C.J., Gupta, P., Kingsley, L.A., Todd, J.A., Saah, A.J., Detels, R., Phair, J.P., Rinaldo, C.R. 1997. Plasma Viral Load and CD4+ Lymphocytes as Prognostic Markers of HIV-1 Infection. *Annals of Internal Medicine* 126(12): 946-954.
131. Blankson, J. 2010. Control of HIV-1 Replication in Elite Suppressors. *Discovery Medicine* 9(46): 261-266.
132. Mothe, B., Ibarondo, J., Llano, A., Brander, C. 2009. Virological, immune and host genetics markers in the control of HIV infection. *Disease Markers* 27: 105-120.
133. Mellors, J.W., Kingsley, L.A., Rinaldo, C.R., Todd, J.A., Hoo, B.S., Kokka, R.P., Gupta, P. 1995. Quantitation of HIV-1 RNA in plasma predicts outcome after seroconversion. *Annals of Internal Medicine* 122(8): 573-579.
134. Giorgi, J.V., Lyles, R.H., Matud, J.L., Yamashita, T.E., Mellors, J.W., Hultin, L.E., Jamieson, B.D., Margolick, J.B., Rinaldo, C.R., Phair, J.P., Detels, R., Multicenter AIDS Cohort Study. 2002. Predictive value of immunologic and virologic markers after long or short duration of HIV-1 infection. *Journal of Acquired Immunodeficiency Syndrome* 29(4): 346-355.
135. Mellors, J.W., Margolick, J.B., Phair, J.P., Rinaldo, C.R., Detels, R., Jacobson, L.P., Muñoz, A. 2007. Prognostic Value of HIV-1 RNA, CD4 Cell Count, and CD4 Cell Count Slope for Progression to AIDS and Death in Untreated HIV-1 Infection. *JAMA : The Journal of the American Medical Association* 297(21): 2349-2350.
136. Mellors, J.W., Rinaldo, C.R., Gupta, P., White, R.M., Todd, J.A., Kingsley, L.A. 1996. Prognosis in HIV-1 Infection Predicted by the Quantity of Virus in Plasma. *Science* 272(5265): 1167-1170.
137. Langford, S.E., Ananworanich, J., Cooper, D.A. 2007. Predictors of disease progression in HIV infection: a review. *AIDS Research and Therapy* 4:11.

138. Donfack, J., Buchinsky, F.J., Post, J.C., Ehrlich, G.D. 2006. Human susceptibility to viral infection: the search for HIV-protective alleles among Africans by means of genome-wide studies. *AIDS Research and Human Retroviruses* 22(10): 925-930.
139. O'Brien, S.J., Nelson, G.W. 2004. Human genes that limit AIDS. *Nature Genetics* 36(6): 565-574.
140. Takeuchi, H., Matano, T. 2008. Host factors involved in resistance to retroviral infection. *Microbiology and Immunology* 52(6): 318-325.
141. Liu, R., Paxton, W.A., Choe, S., Ceradini, D., Martin, S.R., Horuk, R. 1996. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* 86: 367-377.
142. An, P., Wang, L.H., Hutcheson-Dilks, H., Nelson, G., Donfield, S., Goedert, J.J., Rinaldo, C.R., Buchbinder, S., Kirk, G.D., O'Brien, S.J., Winkler, C.A. 2007. Regulatory Polymorphisms in the Cyclophilin A Gene, PPIA, Accelerate Progression to AIDS. *PLoS Pathogens* 3(6): 849-857.
143. Diop, G., Hirtzig, T., Do, H., Coulonges, C., Vasilescu, A., Labib, T., Spadoni, J.L., Therwath, A., Lathrop, M., Matsuda, F., Zagury, J.F. 2006. Exhaustive genotyping of the interferon alpha receptor 1 (IFNAR1) gene and association of an IFNAR1 protein variant with AIDS progression or susceptibility to HIV-1 infection in a French AIDS cohort. *Biomedicine & Pharmacotherapy* 60(9): 569-577.
144. Ross, S.R. 2009. Are viruses inhibited by APOBEC3 molecules from their host species? *PLoS Pathogens* 5(4): e1000347.
145. van Manen, D., Rits, M.A., Beugeling, C., van Dort, K., Schuitemaker, H., Kootstra, N.A. 2008. The effect of Trim5 polymorphisms on the clinical course of HIV-1 infection. *PLoS Pathogens* 4(2): e18.
146. Shearer, G., Clerici, M. 2010. Historical Perspective on HIV-Exposed Seronegative Individuals: Has Nature Done the Experiment for Us? *The Journal of Infectious Diseases* 202(S3): S329:S332.
147. Fellay, J., Shianna, K.V., Ge, D., Colombo, S., Ledergerber, B., Weale, M., Zhang, K., Gumbs, C., Castagna, A., Cossarizza, A., Cozzi-Lepri, A., De Luca, A., Easterbrook, P., Francioli, P., Mallal, S., Martinez-Picado, J., Miro, J.M., Obel, N., Smith, J.P., Wyniger, J., *et al.* 2007. A whole-genome association study of major determinants for host control of HIV-1. *Science* 317: 944-947.
148. Han, Y., Lai, J., Barditch-Crovo, P., Gallant, J.E., Williams, T.M., Siliciano, R.F., Blankson, J.N. 2008. The role of protective HCP5 and HLA-C associated polymorphisms in the control of HIV-1 replication in a subset of elite suppressors. *AIDS* 22: 541-544.
149. Martin, M.P., Qi, Y., Gao, X., Yamada, E., Martin, J.N., Pereyra, F., Colombo, S., Brown, E.E., Shupert, W.L., Phair, J., Goedert, J.J., Buchbinder, S., Kirk, G.D., Telenti, A., Connors, M., O'Brien, S.J., Walker, B.D., Parham, P., Deeks, S.G., McVicar, D.W., *et al.* 2007. Innate partnership of HLA-B and KIR3DL1 subtypes against HIV-1. *Nature Genetics* 39: 733-740.
150. Deacon, N.J., Tsykin, A., Solomon, A., Smith, K., Ludford-Menting, M., Hooker, D.J., McPhee, D.A., Greenway, A.L., Ellett, A., Chatfield, C., Lawson, V.A., Crowe, S., Maerz, A., Sonza, S., Learmont, J., Sullivan, J.S., Cunningham, A., Dwyer, D., Dowton, D., Mills, J. 1995. Genomic structure of an attenuated quasi species of HIV-1 from a blood transfusion donor and recipients. *Science* 270(5238): 988-991.
151. Kirchhoff, F., Greenough, T.C., Brettler, D.B., Sullivan, J.L., Desrosiers, R.C. 1995. Brief report: absence of intact *nef* sequences in a long-term survivor with nonprogressive HIV-1 infection. *New England Journal of Medicine* 332(4): 228-232.
152. Miura, T., Brockman, M.A., Brumme, C.J., Brumme, Z.L., Carlson, J.M., Pereyra, F., Trocha, A., Addo, M.M., Block, B.L., Rothchild, A.C., Baker, B.M., Flynn, T.,

- Schneidewind, A., Li, B., Wang, Y.E., Heckerman, D., Allen, T.M., Walker, B.D. 2008. Genetic Characterization of Human Immunodeficiency Virus Type 1 in Elite Controllers: Lack of Gross Genetic Defects or Common Amino Acid Changes. *Journal of Virology* 82(17): 8422-8430.
153. Hessel, A.J., Hangartner, L., Hunter, M., Havenith, C.E., Beurskens, F.J., Bakker, J.M., Lanigan, C.M., Landucci, G., Forthal, D.N., Parren, P.W., Marx, P.A., Burton, D.R. 2007. Fc receptor but not complement binding is important in antibody protection against HIV. *Nature* 449(7158): 101-104.
154. Doria-Rose, N.A., Klein, R.M., Manion, M.M., O'Dell, S., Phogat, A., Chakrabarti, B., Hallahan, C.W., Migueles, S.A., Wrammert, J., Ahmed, R., Nason, M., Wyatt, R.T., Mascola, J.R., Connors, M. 2009. Frequency and phenotype of human immunodeficiency virus envelope-specific B cells from patients with broadly cross-neutralizing antibodies. *Journal of Virology* 83: 188-199.
155. Bailey, J.R., Lassen, K.G., Yang, H.C., Quinn, T.C., Ray, S.C., Blankson, J.N., Siliciano, R.F. 2006. Neutralizing antibodies do not mediate suppression of human immunodeficiency virus type 1 in elite suppressors or selection of plasma virus variants in patients on highly active antiretroviral therapy. *Journal of Virology* 80: 4758-4770.
156. Pereyra, F., Addo, M.M., Kaufmann, D.E., Liu, Y., Miura, T., Rathod, A., Baker, B., Trocha, A., Rosenberg, R., Mackey, E., Ueda, P., Lu, Z., Cohen, D., Wrin, T., Petropoulos, C.J., Rosenberg, E.S., Walker, B.D. 2008. Genetic and immunologic heterogeneity among persons who control HIV infection in the absence of therapy. *Journal of Infectious Diseases* 197(4): 563-571.
157. Pantaleo, G., Koup, R.A. 2004. Correlates of immune protection in HIV-1 infection: what we know, what we don't know, what we should know. *Nature Medicine* 10(8): 806-810.
158. Potter, S.J., Lacabaratz, C., Lambotte, O., Perez-Patrigeon, S., Vingert, B., Sinet, M., Colle, J.H., Urrutia, A., Scott-Algara, D., Boufassa, F., Delfraissy, J.F., Thèze, J., Venet, A., Chakrabarti, L.A. 2007. Preserved central memory and activated effector memory CD4+ T-cell subsets in human immunodeficiency virus controllers: an ANRS EP36 study. *Journal of Virology* 81: 13904-13915.
159. Chevalier, M.F., Jülg, B., Pyo, A., Flanders, M., Ranasinghe, S., Soghoian, D.Z., Kwon, D.S., Rychert, J., Lian, J., Muller, M.I., Cutler, S., McAndrew, E., Jessen, H., Pereyra, F., Rosenberg, E.S., Altfeld, M., Walker, B.D., Streeck, H. 2011. HIV-1-specific interleukin-21+ CD4+ T cell responses contribute to durable viral control through the modulation of HIV-specific CD8+ T cell function. *Journal of Virology* 85(2): 733-741.
160. Migueles, S.A., Weeks, K.A., Nou, E., Berkley, A.M., Rood, J.E., Osborne, C.M., Hallahan, C.W., Cogliano-Shutta, N.A., Metcalf, J.A., McLaughlin, M., Kwan, R., Mican, J.M., Davey, R.T., Connors, M. 2009. Defective human immunodeficiency virus-specific CD8+ T-cell polyfunctionality, proliferation, and cytotoxicity are not restored by antiretroviral therapy. *Journal of Virology* 83(22): 11876-11889.
161. Betts, M.R., Nason, M.C., West, S.M., De Rosa, S.C., Migueles, S.A., Abraham, J., Lederman, M.M., Benito, J.M., Goepfert, P.A., Connors, M., Roederer, M., Koup, R.A. 2006. HIV nonprogressors preferentially maintain highly functional HIV-specific CD8+ T cells. *Blood* 107: 4781-4789.
162. Migueles, S.A., Laborico, A.C., Shupert, W.L., Sabbaghian, M.S., Rabin, R., Hallahan, C.W., Van Baarle, D., Kostense, S., Miedema, F., McLaughlin, M., Ehler, L., Metcalf, J., Liu, S., Connors, M. 2002. HIV-specific CD8+ T cell proliferation is coupled to perforin expression and is maintained in nonprogressors. *Nature Immunology* 3(11): 1061-1068.

163. Migueles, S.A., Osborne, C.M., Royce, C., Compton, A.A., Joshi, R.P., Weeks, K.A., Rood, J.E., Berkley, A.M., Sacha, J.B., Cogliano-Shutta, N.A., Lloyd, M., Roby, G., Kwan, R., McLaughlin, M., Stallings, S., Rehm, C., O'Shea, M.A., Mican, J., Packard, B.Z., Komoriya, A., *et al.* 2008. Lytic granule loading of CD8⁺ T cells is required for HIV-infected cell elimination associated with immune control. *Immunity* 29(6): 1009-1021.
164. Rehr, M., Cahenzli, J., Haas, A., Price, D.A., Gostick, E., Huber, M., Karrer, U., Oxenius, A. 2008. Emergence of polyfunctional CD8⁺ T cells after prolonged suppression of human immunodeficiency virus replication by antiretroviral therapy. *Journal of Virology* 82: 3391-3404.
165. Streeck, H., Brumme, Z.L., Anastario, M., Cohen, K.W., Jolin, J.S., Meier, A., Brumme, C.J., Rosenberg, E.S., Alter, G., Allen, T.M., Walker, B.D., Altfeld, M. 2008. Antigen Load and Viral Sequence Diversification Determine the Functional Profile of HIV-1-Specific CD8⁺ T Cells. *PLoS Medicine* 5(5): e100.
166. Rowland-Jones, S.L., Dong, T., Fowke, K.R., Kimani, J., Krausa, P., Newell, H., Blanchard, T., Ariyoshi, K., Oyugi, J., Ngugi, E., Bwayo, J., MacDonald, K.S., McMichael, A.J., Plummer, F.A. 1998. Cytotoxic T cell responses to multiple conserved HIV epitopes in HIV-resistant prostitutes in Nairobi. *The Journal of Clinical Investigation* 102(9): 1758-1765.
167. Carrington, M., O'Brien, S.J. 2003. The Influence of HLA Genotype on AIDS. *Annual Review of Medicine* 54: 535-551.
168. Gao, X., Bashirova, A., Iversen, A.K.N., Phair, J., Goedert, J.J., Buchbinder, S., Hoots, K., Vlahov, D., Altfeld, M., O'Brien, S.J., Carrington, M. 2005. AIDS restriction HLA allotypes target distinct intervals of HIV-1 pathogenesis. *Nature Medicine* 11(2): 1290-1292.
169. Pereyra, F., Jia, X., McLaren, P.J., Telenti, A., de Bakker, P.I., Walker, B.D., Ripke, S., Brumme, C.J., Pulit, S.L., Carrington, M., Kadie, C.M., Carlson, J.M., Heckerman, D., Graham, R.R., Plenge, R.M., Deeks, S.G., Gianniny, L., Crawford, G., Sullivan, J., Gonzalez, E., *et al.*, for the International HIV Controllers Study. 2010. The Major Genetic Determinants of HIV-1 Control Affect HLA Class I Peptide Presentation. *Science* 330(6010): 1551-1557.
170. Kiepiela, P., Leslie, A.J., Honeyborne, I., Ramduth, D., Thobakgale, C., Chetty, S., Rathnavalu, P., Moore, C., Pfafferott, K.J., Hilton, L., Zimbwa, P., Moore, S., Allen, T., Brander, C., Addo, M.M., Altfeld, M., James, I., Mallal, S., Bunce, M., Barber, L.D., *et al.* 2004. Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. *Nature* 432: 769-774.
171. Tang, J., Tang, S., Lobashevsky, E., Myracle, A.D., Fideli, U., Aldrovandi, G., Allen, S., Musonda, R., Kaslow, R.A., and the Zambian-UAB HIV Research Project. 2002. Favorable and Unfavorable HLA Class I Alleles and Haplotypes in Zambians Predominantly Infected with Clade C Human Immunodeficiency Virus Type 1. *Journal of Virology* 76(16): 8276-8284.
172. Kaslow, R.A., Carrington, M., Apple, R., Park, L., Muñoz, A., Saah, A.J., Goedert, J.J., Winkler, C., O'Brien, S.J., Rinaldo, C., Detels, R., Blattner, W., Phair, J., Erlich, H., Mann, D.L. 1996. Influence of combinations of human major histocompatibility complex genes on the course of HIV-1 infection. *Nature Medicine* 2: 405-411.
173. Julg, B., Moodley, E.S., Qi, Y., Ramduth, D., Reddy, S., Mncube, Z., Gao, X., Goulder, P.J., Detels, R., Ndung'u, T., Walker, B.D., Carrington, M. 2011. Possession of HLA Class II DRB1*1303 Associates with Reduced Viral Loads in Chronic HIV-1 Clade C and B Infection. *Journal of Infectious Diseases* 203(6): 803-809.

174. Frater, A.J., Brown, H., Oxenius, A., Günthard, H.F., Hirschel, B., Robinson, N., Leslie, A.J., Payne, R., Crawford, H., Prendergast, A., Brander, C., Kiepiela, P., Walker, B.D., Goulder, P.J.R., McLean, A., Phillips, R.E., and the Swiss HIV-Cohort Study. 2007. Effective T-cell Responses Select Human Immunodeficiency Virus Mutants and Slow Disease Progression. *Journal of Virology* 81(12): 6742-6751.
175. Leslie, A., Matthews, P.C., Listgarten, J., Carslon, J., Kadie, C., Ndung'u, T., Coovadia, H., Walker, B.D., Heckerman, D., Goulder, P.J.R. 2010. Additive contribution of HLA class I alleles in the immune control of HIV-1 infection. *Journal of Virology* 84: 9879-9888.
176. Ngumbela, K.C., Day, C.L., Mncube, Z., Nair, K., Ramduth, D., Thobakgale, C., Moodley, E., Reddy, S., De Pierres, C., Mkhwanazi, N., Bishop, K., van der Stok, M., Ismail, N., Honeyborne, I., Crawford, H., Kavanagh, D.G., Rousseau, C., Nickle, D., Mullins, J., Heckerman, D., *et al.* 2008. Targeting of a CD8 T Cell Env Epitope Presented by HLA-B*5802 Is Associated with Markers of HIV Disease Progression and Lack of Selection Pressure. *AIDS Research and Human Retroviruses* 24(1): 72-82.
177. Boulet, S., Kleyman, M., Kim, J.Y., Kanya, P., Sharafi, S., Simic, N., Bruneau, J., Routy, J.P., Tsoukas, C.M., Bernard, N.F. 2008. A combined genotype of KIR3DL1 high expressing alleles and HLA-B*57 is associated with a reduced risk of HIV infection. *AIDS* 22(12): 1487-1491.
178. Hamlyn, E., Hickling, S., Frater, J., Phillips, R., Babiker, A., McClure, M., Fidler, S. 2010. Protective HLA class I alleles are associated with reduced immune activation and fibrinolysis in individuals with primary HIV infection. *HIV Medicine* 11: 3-3.
179. Loomis-Price, L.D., Cox, J.H., Mascola, J.R., VanCott, T.C., Michael, N.L., Fouts, T.R., Redfield, R.R., Robb, M.L., Wahren, B., Sheppard, H.W., Birx, D.L. 1998. Correlation between humoral responses to human immunodeficiency virus type 1 envelope and disease progression in early-stage infection. *Journal of Infectious Diseases* 178(5): 1306-1316.
180. Hazenberg, M.D., Otto, S.A., van Benthem, B.H., Roos, M.T., Coutinho, R.A., Lange, J.M., Hamann, D., Prins, M., Miedema, F. 2003. Persistent immune activation in HIV-1 infection is associated with progression to AIDS. *AIDS* 17(13): 1881-1888.
181. Limou, S., Coulonges, C., Herbeck, J.T., Van Manen, D., An, P., Le Clerc, S., Delaneau, O., Diop, G., Taing, L., Montes, M., van't Wout, A.B., Gottlieb, G.S., Therwath, A., Rouzioux, C., Delfraissy, J.F., Lelièvre, J.D., Lévy, Y., Herberg, S., Dina, C., Phair, J., *et al.* 2010. Multiple-Cohort Genetic Association Study Reveals CXCR6 as a New Chemokine Receptor Involved in Long-Term Nonprogression to AIDS. *Journal of Infectious Diseases* 202(6): 908-915.
182. Owen, R.E., Heitman, J.W., Hirschhorn, D.F., Lanterna, M.C., Biswas, H.H., Martin, J.N., Krone, M.R., Deeks, S.G., Norris, P.J., NIAID Center for HIV/AIDS Vaccine Immunology. 2010. HIV+ elite controllers have low HIV-specific T-cell activation yet maintain strong, polyfunctional T-cell responses. *AIDS* 24: 1095-1105.
183. Altfeld, M., Kalife, E.T., Qi, Y., Streeck, H., Lichterfeld, M., Johnston, M.N., Burgett, N., Swartz, M.E., Yang, A., Alter, G., Yu, X.G., Meier, A., Rockstroh, J.K., Allen, T.M., Jessen, H., Rosenberg, E.S., Carrington, M., Walker, B.D. 2006. HLA Alleles Associated with Delayed Progression to AIDS Contribute Strongly to the Initial CD8+ T Cell Response against HIV-1. *PLoS Medicine* 3(10): 1851-1864.
184. Streeck, H., Jolin, J.S., Qi, Y., Yassine-Diab, B., Johnson, R.C., Kwon, D.S., Addo, M.M., Brumme, C., Routy, J.P., Little, S., Jessen, H.K., Kelleher, A.D., Hecht, F.M., Sekaly, R.P., Rosenberg, E.S., Walker, B.D., Carrington, M., Altfeld, M. 2009. Human immunodeficiency virus type 1-specific CD8+ T-cell responses during

- primary infection are major determinants of the viral set point and loss of CD4+ T cells. *Journal of Virology* 83(15): 7641-7648.
185. Harari, A., Cellerai, C., Enders, F.B., Köstler, J., Codarri, L., Tapia, G., Boyman, O., Castro, E., Gaudieri, S., James, I., John, M., Wagner, R., Mallal, S., Pantaleo, G. 2007. Skewed association of polyfunctional antigen-specific CD8 T cell populations with HLA-B genotype. *Proceedings of the National Academy of Sciences of the United States of America* 104(41): 16233-16238.
 186. Košmrlj, A., Read, E.L., Qi, Y., Allen, T.M., Altfeld, M., Deeks, S.G., Pereyra, F., Carrington, M., Walker, B.D., Chakraborty, A.K. 2010. Effects of thymic selection of the T-cell repertoire on HLA class I-associated control of HIV infection. *Nature* 465: 350-356.
 187. Crawford, H., Lumm, W., Leslie, A., Schaefer, M., Boeras, D., Prado, J.G., Tang, J., Farmer, P., Ndung'u, T., Lakhi, S., Gilmour, J., Goepfert, P., Walker, B.D., Kaslow, R., Mulenga, J., Allen, S., Goulder, P.J.R., Hunter, E. 2009. Evolution of HLA-B*5703 HIV-1 escape mutations in HLA-B*5703-positive individuals and their transmission recipients. *Journal of Experimental Medicine* 206: 909-919.
 188. Honeyborne, I., Prendergast, A., Pereyra, F., Leslie, A., Crawford, H., Payne, R., Reddy, S., Bishop, K., Moodley, E., Nair, K., van der Stok, M., McCarthy, N., Rousseau, C.M., Addo, M., Mullins, J.I., Brander, C., Kiepiela, P., Walker, B.D., Goulder, P.J.R. 2007. Control of Human Immunodeficiency Virus Type 1 is Associated with HLA-B*13 and Targeting of Multiple Gag-Specific CD8+ T-Cell Epitopes. *Journal of Virology* 81(7): 3667-3672.
 189. Wang, Y.E., Li, B., Carlson, J.M., Streeck, H., Gladden, A.D., Goodman, R., Schneidewind, A., Power, K.A., Toth, I., Frahm, N., Alter, G., Brander, C., Carrington, M., Walker, B.D., Altfeld, M., Heckerman, D., Allen, T.M. 2009. Protective HLA Class I Alleles That Restrict Acute-Phase CD8 T-Cell Responses Are Associated with Viral Escape Mutations Located in Highly Conserved Regions of Human Immunodeficiency Virus Type 1. *Journal of Virology* 83(4): 1845-1855.
 190. Altfeld, M., Kalife, E.T., Qi, Y., Streeck, H., Lichterfeld, M., Johnston, M.N., Burgett, N., Swartz, M.E., Yang, A., Alter, G., Yu, X.G., Meier, A., Rockstroh, J.K., Allen, T.M., Jessen, H., Rosenberg, E.S., Carrington, M., Walker, B.D. 2006. HLA Alleles Associated with Delayed Progression to AIDS Contribute Strongly to the Initial CD8+ T Cell Response against HIV-1. *PLoS Medicine* 3(10): 1851-1864.
 191. Zuniga, R., Mothe, B., Llano, A., Ibarondo, J., Daniels, M., Farfan, M., Suarez, J., Szinger, J.J., Hildebrand, W., Lucchetti, A., Heckerman, D., Goulder, P., Walker, B.D., Sanchez, J., Korber, B.T., Brander, C. 2009. *HIV specific T cell responses and response patterns associated with viral control independent of classical non-progressor HLA class I alleles, S01-06 OA*. Presented at AIDS Vaccine 2009 Conference, Paris, France.
 192. Llano, A., Ibarondo, J., Zamarreño, J., Miranda, C., Puertas, M.C., Daniels, M., Szinger, J., Rolland, M., Martinez-Picado, J., Yang, O., Mullins, J.I., Korber, B., Heckerman, D., Sanchez, J., Brander, C., Mothe, B. 2010. *Identification of HIV-1 peptide sequences associated with the induction of protective T cell immunity across different two clade B infected host populations, MOPE0048*. Presented at XVIII International AIDS Conference, Vienna, Austria.
 193. Kiepiela, P., Ngumbela, K., Thobakgale, C., Ramduth, D., Honeyborne, I., Moodley, E., Reddy, S., de Pierres, C., Mncube, Z., Mkhwanazi, N., Bishop, K., van der Stok, M., Nair, K., Khan, N., Crawford, H., Payne, R., Leslie, A., Prado, J., Prendergast, A., Frater, J., *et al.* 2007. CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. *Nature Medicine* 13(1): 46-53.

194. Edwards, B.H., Bansal, A., Sabbaj, S., Bakari, J., Mulligan, M.J., Goepfert, P.A. 2002. Magnitude of Functional CD8+ T-Cell Responses to the Gag Protein of Human Immunodeficiency Virus Type 1 Correlates Inversely with Viral Load in Plasma. *Journal of Virology* 76(5): 2298-2305.
195. Rolland, M., Heckerman, D., Deng, W., Rousseau, C., Coovadia, H., Bishop, K., Goulder, P.J.R., Walker, B.D., Brander, C., Mullins, J.I. 2008. Broad and Gag-Biased HIV-1 Epitope Repertoires Are Associated with Lower Viral Loads. *PLoS One* 3(1): e1424.
196. Zuniga, R., Lucchetti, A., Galvan, P., Sanchez, S., Sanchez, C., Hernandez, A., Sanchez, H., Frahm, N., Linde, C.H., Hewitt, H.S., Hildebrand, W., Altfeld, M., Allen, T.M., Walker, B.D., Korber, B.T., Leitner, T., Sanchez, J., Brander, C. 2006. Relative Dominance of Gag p24-Specific Cytotoxic T Lymphocytes Is Associated with Human Immunodeficiency Virus Control. *Journal of Virology* 80(6): 3122-3125.
197. Saez-Cirion, A., Sinet, M., Shin, S.Y., Urrutia, A., Versmisse, P., Lacabaratz, C., Boufassa, F., Avettand-Fenoel, V., Rouzioux, C., Delfraissy, J.F., Barre-Sinoussi, F., Lambotte, O., Venet, A., Pancino, G., ANRS EP36 HIV Controllers Study Group. 2009. Heterogeneity in HIV suppression by CD8 T cells from HIV controllers: association with Gag-specific CD8 T cell responses. *Journal of Immunology* 182: 7828-7837.
198. Julg, B., Williams, K.L., Reddy, S., Bishop, K., Qi, Y., Carrington, M., Goulder, P.J., Ndung'u, T., Walker, B.D. 2010. Enhanced anti-HIV functional activity associated with Gag-specific CD8 T-cell responses. *Journal of Virology* 84: 5540-5549.
199. Patke, D.S., Langan, S.J., Carruth, L.M., Keating, S.M., Sabundayo, B.P., Margolick, J.B., Quinn, T.C., Bollinger, R.C. 2002. Association of Gag-Specific T Lymphocyte Responses during the Early Phase of Human Immunodeficiency Virus Type 1 Infection and Lower Virus Load Set Point. *Journal of Infectious Diseases* 186(8): 1177-1180.
200. Gray, C.M., Mlotshwa, M., Riou, C., Mathebula, T., de Assis Rosa, D., Mashishi, T., Seoighe, C., Ngandu, N., van Loggerenberg, F., Morris, L., Mlisana, K., Williamson, C., Karim, S.A. 2009. Human Immunodeficiency Virus-Specific Gamma Interferon Enzyme-Linked Immunospot Assay Responses Targeting Specific Regions of the Proteome during Primary Subtype C Infection are Poor Predictors of the Course of Viremia and Set Point. *Journal of Virology* 83(1): 470-478.
201. Schellens, I.M.M., Borghans, J.A.M., Jansen, C.A., De Cuyper, I.M., Geskus, R.B., van Baarle, D., Miedema, F. 2008. Abundance of Early Functional HIV-Specific CD8+ T Cells Does Not Predict AIDS-Free Survival Time. *PLoS One* 3(7): e2745,1-8.
202. Vingert, B., Perez-Patrigéon, S., Jeannin, P., Lambotte, O., Boufassa, F., Lemaitre, F., Kwok, W.W., Theodorou, I., Delfraissy, J.F., Thèze, J., Chakrabarti, L.A., ANRS EP36 HIV Controllers Study Group. 2010. HIV controller CD4+ T cells respond to minimal amounts of gag antigen due to high TCR avidity. *PLoS Pathogens* 6: e1000780.
203. Jansen, C.A., De Cuyper, I.M., Hooibrink, B., van der Bij, A.K., van Baarle, D., Miedema, F. 2006. Prognostic value of HIV-1 Gag-specific CD4+ T-cell responses for progression to AIDS analyzed in a prospective cohort study. *Blood* 107: 1427-1433.
204. Ndongala, M.L., Peretz, Y., Boulet, S., Doroudchi, M., Yassine-Diab, B., Boulassel, M.R., Rouleau, D., Tremblay, C., LeBlanc, R., Routy, J.P., Sékaly, R.P., Bernard, N.F. 2009. HIV Gag p24 specific responses secreting IFN-gamma and/or IL-2 in

- treatment-naïve individuals in acute infection early disease (AIED) are associated with low viral load. *Journal of Clinical Immunology* 131(2): 277-287.
205. Chevalier, M., Pyo, A., Jolin, J.S., Addo, M., Kwon, D.S., Toth, I., Walker, B., Streeck, H. 2009. *Skewed HIV-1-specific CD4+ Th2 helper cell contribution in progressive HIV-1 infection, P16-31*. Presented at AIDS Vaccine 2009 Conference, Paris, France.
 206. Rousseau, C.M., Daniels, M.G., Carlson, J.M., Kadie, C., Crawford, H., Prendergast, A., Matthews, P., Payne, R., Rolland, M., Raugi, D.N., Maust, B.S., Learn, G.H., Nickle, D.C., Coovadia, H., Ndung'u, T., Frahm, N., Brander, C., Walker, B.D., Goulder, P.J.R., Bhattacharya, T., *et al.* 2008. HLA Class I-Driven Evolution of Human Immunodeficiency Virus Type 1 Subtype C Proteome: Immune Escape and Viral Load. *Journal of Virology* 82(13): 6434-6446.
 207. Goepfert, P.A., Lumm, W., Farmer, P., Matthews, P., Prendergast, A., Carlson, J.M., Derdeyn, C.A., Tang, J., Kaslow, R.A., Bansal, A., Yusim, K., Heckerman, D., Mulenga, J., Allen, S., Goulder, P.J.R., Hunter, E. 2008. Transmission of HIV-1 Gag immune escape mutations is associated with reduced viral load in linked recipients. *Journal of Experimental Medicine* 205: 1009-1017.
 208. Chopera, D.R., Woodman, Z., Mlisana, K., Mlotshwa, M., Martin, D.P., Seoighe, C., Treurnicht, F., Assis de Rosa, D., Hide, W., Karim, S.A., Gray, C.M., Williamson, C., CAPRISA 002 Study Team. 2008. Transmission of HIV-1 CTL Escape Variants Provides HLA-Mismatched Recipients with a Survival Advantage. *PLoS Pathogens* 4(3) 1-12.
 209. Matthews, P.C., Prendergast, A., Leslie, A., Crawford, H., Payne, R., Rousseau, C., Rolland, M., Honeyborne, I., Carlson, J., Kadie, C., Brander, C., Bishop, K., Mlotshwa, N., Mullins, J.I., Coovadia, H., Ndung'u, T., Walker, B.D., Heckerman, D., Goulder, P.J.R. 2008. Central Role of Reverting Mutations in HLA Associations with Human Immunodeficiency Virus Set Point. *Journal of Virology* 82 (17): 8548–8559.
 210. Leslie, A.J., Pfafferott, K.J., Chetty, P., Draenert, R., Addo, M.M., Feeney, M., Tang, Y., Holmes, E.C., Allen, T., Prado, J.G., Altfeld, M., Brander, C., Dixon, C., Ramduth, D., Jeena, P., Thomas, S.A., St John, A., Roach, T.A., Kupfer, B., Luzzi, G., *et al.* 2004. HIV evolution: CTL escape mutation and reversion after transmission. *Nature Medicine* 10(3): 282-288.
 211. Martinez-Picado, J., Prado, J.G., Fry, E.E., Pfafferott, K., Leslie, A., Chetty, S., Thobakgale, C., Honeyborne, I., Crawford, H., Matthews, P., Pillay, T., Rousseau, C., Mullins, J.I., Brander, C., Walker, B.D., Stuart, D.I., Kiepiela, P., Goulder, P. 2006. Fitness Cost of Escape Mutations in p24 Gag in Association with Control of Human Immunodeficiency Virus Type 1. *Journal of Virology* 80(7): 3617-3623.
 212. Crawford, H., Prado, J.G., Leslie, A., Hué, S., Honeyborne, I., Reddy, S., van der Stok, M., Mncube, Z., Brander, C., Rousseau, C., Mullins, J.I., Kaslow, R., Goepfert, P., Allen, S., Hunter, E., Mulenga, J., Kiepiela, P., Walker, B.D., Goulder, P.J.R. 2007. Compensatory Mutation Partially Restores Fitness and Delays Reversion of Escape Mutation within the Immunodominant HLA-B*5703-Restricted Gag Epitope in Chronic Human Immunodeficiency Virus Type 1 Infection. *Journal of Virology* 81(15): 8346-8351.
 213. Schneidewind, A., Brockman, M.A., Yang, R., Adam, I., Li, B., Le Gall, S., Rinaldo, C.R., Craggs, S.L., Allgaier, R.L., Power, K.A., Kuntzen, T., Tung, C.S., LaBute, M.X., Mueller, S.M., Harrer, T., McMichael, A.J., Goulder, P.J., Aiken, C., Brander, C., Kelleher, A.D., *et al.* 2007. Escape from the dominant HLA-B27-restricted cytotoxic T-lymphocyte response in Gag is associated with a dramatic reduction in

- human immunodeficiency virus type 1 replication. *Journal of Virology* 81: 12382-12393.
214. Schneidewind, A., Brockman, M.A., Sidney, J., Wang, Y.E., Chen, H., Suscovich, T.J., Li, B., Adam, R.I., Allgaier, R.L., Mothé, B.R., Kuntzen, T., Oniangue-Ndza, C., Trocha, A., Yu, X.G., Brander, C., Sette, A., Walker, B.D., Allen, T.M. 2008. Structural and functional constraints limit options for cytotoxic T-lymphocyte escape in the immunodominant HLA-B27-restricted epitope in human immunodeficiency virus type 1 capsid. *Journal of Virology* 82: 5594-5605.
 215. Goulder, P.J.R., Phillips, R.E., Colbert, R.A., McAdam, S., Ogg, G., Nowak, M.A., Gianfranco, P., Luzzi, G., Morgan, B., Edwards, A., McMichael, A.J., Rowland-Jones, S. 1997. Late escape from an immunodominant cytotoxic T-lymphocyte response associated with progression to AIDS. *Nature Medicine* 3: 212-217.
 216. Sanchez-Merino, V., Farrow, M.A., Brewster, F., Somasundaran, M., Luzuriaga, K. 2008. Identification and Characterization of HIV-1 CD8+ T Cell Escape Variants with Impaired Fitness. *Journal of Infectious Diseases* 197: 300-308.
 217. Blaak, H., Brouwer, M., Ran, L.J., de Wolf, F., Schuitemaker, H. 1998. In vitro replication kinetics of human immunodeficiency virus type 1 (HIV-1) variants in relation to virus load in long-term survivors of HIV-1 infection. *Journal of Infectious Diseases* 177: 600-610.
 218. Quiñones-Mateu, M.E., Ball, S.C., Marozsan, A.J., Torre, V.S., Albright, J.L., Vanham, G., van der Groen, G., Colebunders, R.L., Arts, E.J. 2005. A Dual Infection/Competition Assay Shows a Correlation between Ex Vivo Human Immunodeficiency Virus Type 1 Fitness and Disease Progression. *Journal of Virology* 74(19): 9222-9233.
 219. Prado, J.G., Prendergast, A., Thobakgale, C., Molina, C., Tudor-Williams, G., Ndung'u, T., Walker, B.D., Goulder, P. 2010. Replicative Capacity of Human Immunodeficiency Virus Type 1 Transmitted from Mother to Child Is Associated with Pediatric HIV-1 Disease Progression Rate. *Journal of Virology* 84(1): 492-502.
 220. Trkola, A., Kuster, H., Leemann, C., Ruprecht, C., Joos, B., Telenti, A., Hirschel, B., Weber, R., Bonhoeffer, S., Günthard, H.F., and the Swiss HIV Cohort Study. 2003. Human Immunodeficiency Virus Type 1 Fitness Is a Determining Factor in Viral Rebound and Set Point in Chronic Infection. *Journal of Virology* 77(24): 13146-13155.
 221. Abraha, A., Nankya, I.L., Gibson, R., Demers, K., Tebit, D.M., Johnston, E., Katzenstein, D., Siddiqui, A., Herrera, C., Fischetti, L., Shattock, R.J., Arts, E.J. 2009. CCR5- and CXCR4-tropic subtype C HIV-1 isolates have lower pathogenic fitness as compared to the other dominant group M subtypes: Implications for the epidemic. *Journal of Virology* 83(11): 5592-5605.
 222. Ball, S.C., Abraha, A., Collins, K.R., Marozsan, A.J., Baird, H., Quiñones-Mateu, M.E., Penn-Nicholson, A., Murray, M., Richard, N., Lobritz, M., Zimmerman, P.A., Kawamura, T., Blauvelt, A., Arts, E.J. 2003. Comparing the Ex Vivo Fitness of CCR5-Tropic Human Immunodeficiency Virus Type 1 Isolates of Subtypes B and C. *Journal of Virology* 77(2): 1021-1038.
 223. Rangel, H.R., Weber, J., Chakraborty, B., Gutierrez, A., Marotta, M.L., Mirza, M., Kiser, P., Martinez, M.A., Este, J.A., Quiñones-Mateu, M.E. 2003. Role of the Human Immunodeficiency Virus Type 1 Envelope Gene in Viral Fitness. *Journal of Virology* 77(16): 9069-9073.
 224. Campbell, T.B., Schneider, K., Wrin, T., Petropoulos, C.J., Connick, E. 2003. Relationship between in vitro human immunodeficiency virus type 1 replication rate and virus load in plasma. *Journal of Virology* 77(22): 12105-12112.

225. Miura, T., Brockman, M.A., Brumme, Z.L., Brumme, C.J., Pereyra, F., Trocha, A., Block, B.L., Schneidewind, A., Allen, T.M., Heckerman, D., Walker, B.D. 2009. HLA-Associated Alterations in Replication Capacity of Chimeric NL4-3 Viruses Carrying *gag-protease* from Elite Controllers of Human Immunodeficiency Virus Type 1. *Journal of Virology* 83(1): 140-149.
226. Miura, T., Brumme, Z.L., Brockman, M.A., Rosato, P., Sela, J., Brumme, C.J., Pereyra, F., Kaufmann, D.E., Trocha, A., Block, B.L., Daar, E.S., Connick, E., Jessen, H., Kelleher, A.D., Rosenberg, E., Markowitz, M., Schafer, K., Vaida, F., Iwamoto, A., Little, S., *et al.* 2010. Impaired Replication Capacity of Acute/Early Viruses in Persons Who Become HIV Controllers. *Journal of Virology* 84(15): 7581-7591.
227. Schneidewind, A., Tang, Y., Brockman, M.A., Ryland, E.G., Dunkley-Thompson, J., Steel-Duncan, J.C., St John, M.A., Conrad, J.A., Kalams, S.A., Noel, F., Allen, T.M., Christie, C.D., Feeney, M.E. 2009. Maternal transmission of human immunodeficiency virus escape mutations subverts HLA-B57 immunodominance but facilitates viral control in the haploidentical infant. *Journal of Virology* 83(17): 8616-8627.
228. Kawada, M., Igarashi, H., Takeda, A., Tsukamoto, T., Yamamoto, H., Dohki, S., Takiguchi, M., Matano, T. 2006. Involvement of Multiple Epitope-Specific Cytotoxic T-Lymphocyte Responses in Vaccine-Based Control of Simian Immunodeficiency Virus Replication in Rhesus Macaques. *Journal of Virology* 80(4): 1949-1958.
229. Miura, T., Brockman, M.A., Schneidewind, A., Lobritz, M., Pereyra, F., Rathod, A., Block, B.L., Brumme, Z.L., Brumme, C.J., Baker, B., Rothchild, A.C., Li, B., Trocha, A., Cutrell, E., Frahm, N., Brander, C., Toth, I., Arts, E.J., Allen, T.M., Walker, B.D. 2009. HLA-B57/B*5801 Human Immunodeficiency Virus Type 1 Elite Controllers Select for Rare Gag Variants Associated with Reduced Viral Replication Capacity and Strong Cytotoxic T-Lymphocyte Recognition. *Journal of Virology* 83(6): 2743-2755.
230. Feeney, M.E., Tang, Y., Pfafferott, K., Roosevelt, K.A., Draenert, R., Trocha, A., Yu, X.G., Verrill, C., Allen, T., Moore, C., Mallal, S., Burchett, S., McIntosh, K., Pelton, S.I., St John, M.A., Hazra, R., Klenerman, P., Altfeld, M., Walker, B.D., Goulder, P.J. 2005. HIV-1 viral escape in infancy followed by emergence of a variant-specific CTL response. *Journal of Immunology* 174(12): 7524-7530.
231. Goulder, P.J., Brander, C., Tang, Y., Tremblay, C., Colbert, R.A., Addo, M.M., Rosenberg, E.S., Nguyen, T., Allen, R., Trocha, A., Altfeld, M., He, S., Bunce, M., Funkhouser, R., Pelton, S.I., Burchett, S.K., McIntosh, K., Korber, B.T., Walker, B.D. 2001. Evolution and transmission of stable CTL escape mutations in HIV infection. *Nature* 412: 334-338.
232. Dykes, C., Demeter, L.M. 2007. Clinical Significance of Human Immunodeficiency Virus Type 1 Replication Fitness. *Clinical Microbiology Reviews* 20(4): 550-578.
233. Singh, A., Page, T., Moore, P.L., Allgaier, R.L., Hiramani, K., Coovadia, H.M., Walker, B.D., Morris, L., Ndung'u, T. 2009. Functional and genetic analysis of coreceptor usage by dualtropic HIV-1 subtype C isolates. *Virology* 393: 56-67.
234. Weber, J., Weberova, J., Carobene, M., Mirza, M., Martinez-Picado, J., Kazanjian, P., Quiñones-Mateu, M.E. 2006. Use of a novel assay based on intact recombinant viruses expressing green (EGFP) or red (DsRed2) fluorescent proteins to examine the contribution of *pol* and *env* genes to overall HIV-1 replicative fitness. *Journal of Virological Methods* 136: 102-117.
235. Gervaix, A., West, D., Leoni, L.M., Richman, D.D., Wong-Staal, F., Corbeil, J. 1997. A new reporter cell line to monitor HIV infection and drug susceptibility in

- vitro. Proceedings of the National Academy of Sciences of the United States of America* 94: 4653-4658.
236. van Opijnen, T., Berkhout, B. 2005. The host environment drives HIV-1 fitness. *Reviews in Medical Virology* 15: 219-233.
 237. Fikkert, V., Cherepanov, P., Van Laethem, K., Hantson, A., Van Remoortel, B., Pannecouque, C., De Clercq, E., Debyser, Z., Vandamme, A.M., Witvrouw, M. 2002. env Chimeric Virus Technology for Evaluating Human Immunodeficiency Virus Susceptibility to Entry Inhibitors. *Antimicrobial Agents and Chemotherapy* 46(12): 3954-3962.
 238. Dudley, D.M., Gao, Y., Nelson, K.N., Henry, K.R., Nankya, I., Gibson, R.M., Arts, E.J. 2009. A novel yeast-based recombination method to clone and propagate diverse HIV-1 isolates. *BioTechniques* 46: 458-467.
 239. Marozsan, A.J., Arts, E.J. 2003. Development of a yeast-based recombination cloning/system for the analysis of gene products from diverse human immunodeficiency virus type 1 isolates. *Journal of Virological Methods* 111: 111-120.
 240. Marozsan, A.J., Moore, D.M., Lobritz, M.A., Fraundorf, E., Abraha, A., Reeves, J.D., Arts, E.J. 2005. Differences in the Fitness of Two Diverse Wild-Type Human Immunodeficiency Virus Type 1 Isolates Are Related to the Efficiency of Cell Binding and Entry. *Journal of Virology* 79(11): 7121-7134.
 241. Lobritz, M.A., Marozsan, A.J., Troyer, R.M., Arts, E.J. 2007. Natural Variation in the V3 Crown of Human Immunodeficiency Virus Type 1 Affects Replicative Fitness and Entry Inhibitor Sensitivity. *Journal of Virology* 81(15): 8258-8269.
 242. Costa, C.I., Morgado, M.G., Santos, V.G.V., HEC/FIOCRUZ AIDS Clinical Research Group, Bongertz, V. 1996. HIV-1 Isolation from Plasma Specimens. *Memorias do Instituto Oswaldo Cruz* 91(6): 745-746.
 243. Paul, M.O., Tetali, S., Pahwa, S. 1994. Effective Use of Frozen Donor Peripheral Blood Mononuclear Cells for Human Immunodeficiency Virus Type 1 Isolation from Vertically Infected Pediatric Patients. *Journal of Clinical Microbiology* 32(5): 1379-1382.
 244. Mild, M., Björndal, A., Medstrand, P., Fenyö, E.M. 2006. Isolation of human immunodeficiency virus-type 1 (HIV-1) clones with biological and molecular properties of the primary isolate. *Virology* 350(1): 58-66.
 245. Back, N.K., Nijhuis, M., Keulen, W., Boucher, C.A., Essink, B.O.O., van Kuilenburg, A.B., van Gennip, A.H., Berkhout, B. 1996. Reduced replication of 3TC-resistant HIV-1 variants in primary cells due to a processivity defect of the reverse transcriptase enzyme. *The EMBO Journal* 15: 4040-4049.
 246. Huang, K.H., Goedhals, D., Carlson, J.M., Brockman, M.A., Mishra, S., Brumme, Z.L., Hickling, S., Tang, C.S., Miura, T., Seebregts, C., Heckerman, D., Ndung'u, T., Walker, B., Klenerman, P., Steyn, D., Goulder, P., Phillips, R., Bloemfontein-Oxford Collaborative Group, van Vuuren, C., Frater, J. 2011. Progression to AIDS in South Africa Is Associated with both Reverting and Compensatory Viral Mutations. *PLoS One* 6(4): e19018.
 247. Troyer, R.M., Collins, K.R., Abraha, A., Fraundorf, E., Moore, D.M., Krizan, R.W., Toossi, Z., Colebunders, R.L., Jensen, M.A., Mullins, J.I., Vanham, G., Arts, E.J. 2005. Changes in Human Immunodeficiency Virus Type 1 Fitness and Genetic Diversity during Disease Progression. *Journal of Virology* 79(14): 9006-9018.
 248. Dykes, C., Wang, J., Jin, X., Planelles, V., An, D.S., Tallo, A., Huang, Y., Wu, H., Demeter, L.M. 2006. Evaluation of a Multiple-Cycle, Recombinant Virus, Growth Competition Assay That Uses Flow Cytometry To Measure Replication Efficiency of

- Human Immunodeficiency Virus Type 1 in Cell Culture. *Journal of Clinical Microbiology* 44(6): 1930-1943.
249. de Oliveira, T., Engelbrecht, S., Janse van Rensburg, E., Gordon, M., Bishop, K., zur Megede, J., Barnett, S.W., Cassol, S. 2003. Variability at Human Immunodeficiency Virus Type 1 Subtype C Protease Cleavage Sites: an Indication of Viral Fitness? *Journal of Virology* 77(17): 9422-9430.
250. Banke, S., Lillemark, M.R., Gerstoft, J., Obel, N., Jørgensen, L.B. 2009. Positive Selection Pressure Introduces Secondary Mutations at Gag Cleavage Sites in Human Immunodeficiency Virus Type 1 Harboring Major Protease Resistance Mutations. *Journal of Virology* 83(7): 8916-8924.
251. Brockman, M., Brumme, Z., Brumme, C., Miura, T., Sela, J., Rosato, P., Kadie, C., Carlson, J., Markle, T., Streeck, H., Kelleher, A., Markowitz, M., Jessen, H., Rosenberg, E., Altfeld, M., Harrigan, P., Heckerman, D., Walker, B., Allen, T. 2010. Early selection in Gag by protective HLA alleles contributes to reduced HIV-1 replication capacity that may be largely compensated for in chronic infection. *Journal of Virology* 84(22): 11937-11949.
252. Radebe, M., Nair, K., Chonco, F., Bishop, K., Wright, J.K., van der Stok, M., Bassett, I.V., Mncube, Z., Altfeld, M., Walker, B.D., Ndung'u, T. 2011. Limited immunogenicity of HIV CD8+ T-cell epitopes in acute clade C virus infection. *Journal of Infectious Diseases*, 204(5): 768-76.
253. van Loggerenberg, F., Mlisana, K., Williamson, C., Auld, S.C., Morris, L., Gray, C.M., Abdool Karim, Q., Grobler, A., Barnabas, N., Iriogbe, I., Abdool Karim, S.S., the CAPRISA 002 Acute Infection Study Team. 2008. Establishing a Cohort at High Risk of HIV Infection in South Africa: Challenges and Experiences of the CAPRISA 002 Acute Infection Study. *PLoS One* 3(4): e1954.
254. Novitsky, V., Wang, R., Kebaabetswe, L., Greenwald, J., Rossenkhan, R., Moyo, S., Musonda, R., Woldegabriel, E., Lagakos, S., Essex, M. 2009. Better control of early viral replication is associated with slower rate of elicited antiviral antibodies in the detuned enzyme immunoassay during primary HIV-1C infection. *Journal of Acquired Immunodeficiency Syndrome* 52(2): 265-272.
255. Novitsky, V., Woldegabriel, E., Kebaabetswe, L., Rossenkhan, R., Mlotshwa, B., Bonney, C., Finucane, M., Musonda, R., Moyo, S., Wester, C., van Widenfelt, E., Makhema, J., Lagakos, S., Essex, M. 2009. Viral Load and CD4+ T Cell Dynamics in Primary HIV-1 Subtype C Infection. *Journal of Acquired Immunodeficiency Syndrome* 50: 65-76.
256. Huang, X., Zhang, J. 1996. Methods for comparing a DNA sequence with a protein sequence. *Computer Applications in the Biosciences* 12: 497-506.
257. de Oliveira, T., Deforche, K., Cassol, S., Rambaut, A., Vandamme, A.M. Stanford University HIV Drug Resistance Database: Rega HIV-1 Subtyping Tool - Version 2.0. . <http://dbpartners.stanford.edu/RegaSubtyping/> [Accessed 01/08/2009].
258. Ndung'u, T., Renjifo, B., Essex, M. 2001. Construction and Analysis of an Infectious Human Immunodeficiency Virus Type 1 Subtype C Molecular Clone. *Journal of Virology* 75(11): 4964-4972.
259. Freshney, R.I. 1987. *Culture of animal cells. A Manual of Basic Technique*. New York: Alan R. Liss Inc.
260. Meryman, H.T. 1974. Freezing injury and its prevention in living cells. *Annual Review of Biophysics and Bioengineering* 3(0): 341-363.
261. Naidoo, S. 2005. *The role of the kallikrein-kinin cascade and metabolites in angiogenesis and tumourigenesis*. Doctoral thesis. University of KwaZulu-Natal, Nelson R. Mandela School of Medicine: Durban, South Africa.

262. Spencer, S.C. 1991. Electroporation technique of DNA transfection. Pages 45-52 in Murray, E.J., ed. *Methods in Molecular Biology, Vol. 7: Gene Transfer and Expression Protocols*. New Jersey: The Humana Press Inc.
263. Shapiro, H.M. 2003. *Practical Flow Cytometry. 4th Edition*. New Jersey: John Wiley & Sons, Inc.
264. FlowJo. FlowJo in Africa. <http://www.flowjo.com/home/africa/> [Accessed 01/11/2008].
265. Los Alamos National Laboratory. Highlighter. <http://www.hiv.lanl.gov/content/sequence/HIGHLIGHT/highlighter.html> [Accessed 15/09/2009].
266. Rambaut, A. Figtree. <http://tree.bio.ed.ac.uk/software/figtree/> [Accessed 01/10/2009].
267. Los Alamos National Laboratory. HIV Sequence Database. <http://www.hiv.lanl.gov/content/sequence/HIV/> [Accessed 01/09/2009].
268. Hall, T. BioEdit. <http://www.mbio.ncsu.edu/BioEdit/bioedit.html> [Accessed 03/04/2009].
269. Carlson, J.M., Brumme, Z.L., Rousseau, C.M., Brumme, C.J., Matthews, P., Kadie, C., Mullins, J.I., Walker, B.D., Harrigan, P.R., Goulder, P.J.R., Heckerman, D. 2008. Phylogenetic Dependency Networks: Inferring Patterns of CTL Escape and Codon Covariation in HIV-1 Gag. *PLoS Computational Biology* 4(11): 1-23.
270. Storey, J.D., Tibshirani, R. 2003. Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences of the United States of America* 100: 9440-9445.
271. Strachan, T., Read, A.P. 1999. *Human Molecular Genetics. 2nd Edition*. New York: Wiley-Liss.
272. Stratagene. QuikChange+ II XL Site-Directed Mutagenesis Kit, Instruction Manual. stanxterm.aecom.yu.edu/wiki/data/Product.../quikchange2xl.pdf [Accessed 12/01/2010].
273. Thobakgale, C.F., Prendergast, A., Crawford, H., Mkhwanazi, N., Ramduth, D., Reddy, S., Molina, C., Mncube, Z., Leslie, A., Prado, J., Chonco, F., Mphatshwe, W., Tudor-Williams, G., Jeena, P., Blanckenberg, N., Dong, K., Kiepiela, P., Coovadia, H., Ndung'u, T., Walker, B.D., *et al.* 2009. Impact of HLA in Mother and Child on Disease Progression of Pediatric Human Immunodeficiency Virus Type 1 Infection. *Journal of Virology* 83(19): 1-11.
274. Brumme, Z., Wang, B., Nair, K., Brumme, C., de Pierres, C., Reddy, S., Julg, B., Moodley, E., Thobakgale, C., Lu, Z., van der Stok, M., Bishop, K., Mncube, Z., Chonco, F., Yuki, Y., Frahm, N., Brander, C., Carrington, M., Freedberg, M., Kiepiela, P., *et al.* 2009. Impact of select immunologic and virologic biomarkers on CD4 cell count decrease in chronic HIV-1 subtype C infection: Results from Sinikithemba Cohort, Durban, South Africa. *Clinical Infectious Diseases* 49(6): 956-964.
275. Los Alamos National Laboratory. HIV Molecular Immunology Database. <http://www.hiv.lanl.gov/content/immunology/> [Accessed 01/09/2009].
276. Tang, C., Ndassa, Y., Summers, M.F. 2002. Structure of the N-terminal 283-residue fragment of the immature HIV-1 Gag polyprotein. *Nature Structural Biology* 9(7): 537-543.
277. NCBI. Structure, Cn3D homepage. <http://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3d.shtml> [Accessed 30/03/2009].
278. Novitsky, V., Wang, R., Margolin, L., Baca, J., Kebaabetswe, L., Rossenkhan, R., Bonney, C., Herzig, M., Nkwe, D., Moyo, S., Musonda, R., Woldegabriel, E., van Widenfelt, E., Makhema, J., Lagakos, S., Essex, M. 2009. Timing Constraints of In

- Vivo Gag Mutations during Primary HIV-1 Subtype C Infection. *PLoS One* 4(11): e7727.
279. Yang, R., Aiken, C. 2007. A Mutation in Alpha Helix 3 of CA Renders Human Immunodeficiency Virus Type 1 Cyclosporin A Resistant and Dependent: Rescue by a Second-Site Substitution in a Distal Region of CA. *Journal of Virology* 81: 3749-3756.
 280. Price, D.A., Asher, T.E., Wilson, N.A., Nason, M.C., Brenchley, J.M., Metzler, I.S., Venturi, V., Gostick, E., Chattopadhyay, P.K., Roederer, M., Davenport, M.P., Watkins, D.I., Douek, D.C. 2009. Public clonotype usage identifies protective Gag-specific CD8+ T cell responses in SIV infection. *Journal of Experimental Medicine* 206(4): 923-936.
 281. Peyerl, F.W., Barouch, D.H., Yeh, W.W., Bazick, H.S., Kunstman, J., Kunstman, K.J., Wolinsky, S.M., Letvin, N.L. 2003. Simian-Human Immunodeficiency Virus Escape from Cytotoxic T-Lymphocyte Recognition at a Structurally Constrained Epitope. *Journal of Virology* 77(23): 12572–12578.
 282. Mellors, J.W., Margolick, J.B., Phair, J.P., Rinaldo, C.R., Detels, R., Jacobson, L.P., Muñoz, A. 2007. Prognostic Value of HIV-1 RNA, CD4 Cell Count, and CD4 Cell Count Slope for Progression to AIDS and Death in Untreated HIV-1 Infection. *JAMA : The Journal of the American Medical Association* 297(21): 2349-2350.
 283. Mellors, J.W., Kingsley, L.A., Rinaldo, C.R., Todd, J.A., Hoo, B.S., Kokka, R.P., Gupta, P. 1995. Quantitation of HIV-1 RNA in plasma predicts outcome after seroconversion. *Annals of Internal Medicine* 122(8): 573-579.
 284. Brumme, Z.L., Brumme, C.J., Carlson, J., Streeck, H., John, M., Eichbaum, Q., Block, B.L., Baker, B., Kadie, C., Markowitz, M., Jessen, H., Kelleher, A.D., Rosenberg, E., Kaldor, J., Yuki, Y., Carrington, M., Allen, T.M., Mallal, S., Altfeld, M., Heckerman, D., *et al.* 2008. Marked Epitope- and Allele-Specific Differences in Rates of Mutation in Human Immunodeficiency Type 1 (HIV-1) Gag, Pol, and Nef Cytotoxic T-Lymphocyte Epitopes in Acute/Early HIV-1 Infection. *Journal of Virology* 82 (18): 9216–9227.
 285. Duda, A., Lee-Turner, L., Fox, J., Robinson, N., Dustan, S., Kaye, S., Fryer, H., Carrington, M., McClure, M., Mclean, A.R., Fidler, S., Weber, J., Phillips, R.E., Frater, A.J., SPARTAC Trial Investigators. 2009. HLA-Associated Clinical Progression Correlates with Epitope Reversion Rates in Early Human Immunodeficiency Virus Infection. *Journal of Virology* 83(3): 1228-1239.
 286. Schaefer, M., Mulenga, J., Tang, J., Goepfert, P., Farmer, P., Kaslow, R., Allen, S., Hunter, E. 2009. *Dynamics of CTL epitope escape and reversion in an African subtype C cohort, OA06-03*. Presented at AIDS Vaccine 2009 Conference, Paris, France.
 287. Hunter, E., Boeras, D., Schaefer, M., Alexander, M., Yue, L., Mulenga, J., Karita, E., Derdeyn, C., Allen, S. 2009. *Characterization of viral variants that initiate systemic infection after mucosal transmission, S05-01*. Presented at AIDS Vaccine 2009 Conference, Paris, France.
 288. Kawashima, Y., Pfafferott, K., Frater, J., Matthews, P., Payne, R., Addo, M., Gatanaga, H., Fujiwara, M., Hachiya, A., Koizumi, H., Kuse, N., Oka, S., Duda, A., Prendergast, A., Crawford, H., Leslie, A., Brumme, Z., Brumme, C., Allen, T., Brander, C., *et al.* 2009. Adaptation of HIV-1 to human leukocyte antigen class I. *Nature* 458(7238): 641-645.
 289. Abdool Karim, Q., Abdool Karim, S.S. 2002. The evolving HIV epidemic in South Africa. *International Journal of Epidemiology* 31: 37-40.
 290. Lassen, K.G., Lobritz, M.A., Bailey, J.R., Johnston, S., Nguyen, S., Lee, B., Chou, T., Siliciano, R.F., Markowitz, M., Arts, E.J. 2009. Elite Suppressor-Derived HIV-1

- Envelope Glycoproteins Exhibit Reduced Entry Efficiency and Kinetics. *PLoS Pathogens* 5(4): e1000377.
291. Hatziioannou, T., Cowan, S., von Schwedler, U.K., Sundquist, W.I., Bieniasz, P.D. 2004. Species-Specific Tropism Determinants in the Human Immunodeficiency Virus Type 1 Capsid. *Journal of Virology* 78(11): 6005-6012.
 292. Crawford, H., Matthews, P.C., Schaefer, M., Carlson, J.M., Leslie, A., Kilembe, W., Allen, S., Ndung'u, T., Heckerman, D., Hunter, E., Goulder, P.J.R. 2010. The hypervariable HIV-1 capsid protein 1 residues comprise HLA-driven CD8+ T-cell escape mutations and covarying HLA-independent polymorphisms. *Journal of Virology* 85(3): 1384-1390.
 293. Honeyborne, I., Codoñer, F.M., Leslie, A., Tudor-Williams, G., Luzzi, G., Ndung'u, T., Walker, B.D., Goulder, P.J., Prado, J.G. 2010. HLA-Cw*03-restricted CD8+ T-cell responses targeting the HIV-1 gag major homology region drive virus immune escape and fitness constraints compensated for by intracodon variation. *Journal of Virology* 84(21): 11279-11288.
 294. Iordanskiy, S., Waltke, M., Feng, Y., Wood, C. 2010. Subtype-associated differences in HIV-1 reverse transcription affect viral replication. *Retrovirology* 7: 85.
 295. Velazquez-Campoy, A., Todd, M.J., Vega, S., Freire, E. 2001. Catalytic efficiency and vitality of HIV-1 proteases from African viral subtypes. *Proceedings of the National Academy of Sciences of the United States of America* 98(11): 6062-6067.
 296. Wyma, D.J., Kotov, A., Aiken, C. 2000. Evidence for a Stable Interaction of gp41 with Pr55Gag in Immature Human Immunodeficiency Virus Type 1 Particles. *Journal of Virology* 74(20): 9381-9387.
 297. Alexander, L., Weiskopf, E., Greenough, T.C., Gaddis, N.C., Auerbach, M.R., Malim, M.H., O'Brien, S.J., Walker, B.D., Sullivan, J.L., Desrosiers, R.C. 2000. Unusual polymorphisms in human immunodeficiency virus type 1 associated with nonprogressive infection. *Journal of Virology* 74(9): 4361-4376.
 298. Gonzalez-Galarza, F.F., Christmas, S., Middleton, D., Jones, A.R. 2011. Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations. *Nucleic Acid Research* 39: D913-D919.
 299. Rolland, M., Nickle, D.C., Mullins, J.I. 2007. HIV-1 Group M Conserved Elements Vaccine. *PLoS Pathogens* 3(11): e157.
 300. Chopera, D.R., Wright, J.K., Brockman, M.A., Brumme, Z.L. 2011. Immune-mediated attenuation of HIV-1. *Future Virology* 6(8): 917-928.
 301. Dahirel, V., Shekhar, K., Pereyra, F., Miura, T., Artyomov, N., Talsania, S., Allen, T.M., Altfeld, M., Carrington, M., Irvine, D.J., Walker, B.D., Chakraborty, A.K. 2011. Coordinate linkage of HIV evolution reveals regions of immunological vulnerability. *Proceedings of the National Academy of Sciences of the United States of America* 108(28): 11530-11535.
 302. Wang, Y., Oniangue-Ndza, C., Schneidewind, A., Kemper, M., Mellors, E., Qi, Y., Gladden, A., Power, K., Pereyra, F., Walker, B., Carrington, M., Allen, T. 2011. *Control of HIV Viral Replication Mediated by the Synergistic Effects of Unique Combinations of HLA Class I Alleles*, Abstract 428. Presented at Keystone Symposium on Protection from HIV: Targeted Intervention Strategies (X8) 2011, Whistler, British Columbia, Canada.
 303. Ueno, T., Motozono, C., Dohki, S., Mwimanzi, P., Rauch, S., Fackler, O.T., Oka, S., Takiguchi, M. 2008. CTL-Mediated Selective Pressure Influences Dynamic Evolution and Pathogenic Functions of HIV-1 Nef. *Journal of Immunology* 180: 1107-1116.