Author/s:
Hirons, Ashley Sarah

Title:
Virology and pathogenesis of HTLV-1c infection in Central Australia

Date:
2023-04

Persistent Link:
https://hdl.handle.net/11343/337577

Virology and pathogenesis of HTLV-1c infection in Central Australia

Ashley Hirons

ORCID Identifier: 0000-0001-8845-8830

Submitted in total fulfilment of the requirements of the degree of

Doctor of Philosophy

April 2023

Department of Microbiology and Immunology, The University of Melbourne

The Doherty Institute

This thesis is dedicated to all the First Nations people of Mparntwe and surrounding regions in Central Australia who bravely and generously participated in this study, and who continue to battle with, or have lost the battle to, this insidious infection.

Abstract

Human T cell lymphotropic virus type 1 (HTLV-1) subtype-C, endemic in Central Australia, genetically differs from other subtypes and is associated with bronchiectasis, bloodstream infections, and other comorbidities that lower life expectancy. Understanding viral pathogenesis in this setting where health inequity prevails, complicates investigation of the direct effect that HTLV-1c has on the immune system. To address these shortfalls, this thesis uses primary human samples from Central Australia to characterise the host immune response to HTLV-1c infection, in context with the proviral landscape in HTLV-1c infected cells. We further assessed an HTLV-1c humanised (hu) mouse model to understand its potential applications for testing preventative and therapeutic treatments that are elusive for HTLV-1.

We detected HTLV-1c-specific T cells in both asymptomatic carriers and bronchiectasis patients, alongside a dysregulated Th1 response and extensive background activation in all T cell phenotypes. The implications of this chronic T cell activation were assessed from the plasma cytokine profile in HTLV-1c-infected donors. We found that HTLV-1c infection is associated with a combination of pro-inflammatory and immunosuppressive cytokines, however in HTLV-1c-assoiated bronchiectasis, the cytokine networks reveal a profound immune exhaustion.

We also investigated the viral instigators of pathogenesis by identifying HTLV-1c infected primary human cells and characterising the provirus structure within. Cells with surface markers showing potential lung homing phenotype were expanded and highly infected with HTLV-1c defective provirus. FOXP3 proxy phenotype cells, often considered Tregs, did not expand but were also infected with defective provirus. CCR4-CD4+ T cells were the only cells circulating in blood that contained exclusively full-length provirus, and may be a crucial source of persistent infection and transmission. The overall landscape of HTLV-1c proviruses showed a predomination of large internal deletions, likely formed through homologous recombinational DNA repair during cell proliferation at provirus breakpoints exhibiting LTR-like sequences. However, the retention of the *hbz* gene was common in defective proviruses and a prominent feature in Bex donors, supporting *hbz* contributing to pulmonary disease. We also identified numerous unique defective chimeric HTLV-1c-human genome (hg) provirus sequences which implicate some captured host cellular genes in pathogenesis.

We determined that the hu-mouse model of HTLV-1c infection recapitulates many aspects of human in vivo infection, despite lacking adaptive immunity. We found that phenotypic lung homing cells were analogously expanded and infected with HTLV-1c. Significant levels of background T cell activation were present in the hu-mouse model, akin to human infection. Moreover, we detected high levels of defective provirus in hu-mice, alongside HTLV-1c-hg chimeric proviruses. Taken together, the HTLV-1c hu-mouse model is an important tool for pre-clinical testing of potential therapeutic and preventative strategies.

Ultimately, this thesis shows that HTLV-1c infection in Central Australia directs pathogenesis through a complex network of chronic T cell activation, immune dysregulation leading to exhaustion, potential persistent expression of *hbz* in defective proviruses, with low levels of ongoing viral dissemination. These important findings will direct future studies and development of therapeutic strategies of HTLV-1c infection in Australia, where they are desperately needed.

Declaration

This is to certify that:

(i) This thesis comprises only my original work towards the degree of Doctor of Philosophy, except where indicated in the preface.

(ii) Due acknowledgement has been made in the text to all other materials used.

(iii) This thesis is less than 100,000 works in length, exclusive of tables, maps bibliographies, appendices and footnotes.

Ashley Hirons

Preface

This work was carried out in the Retrovirus Molecular Virology Laboratory, in the Department of Microbiology and Immunology, The University of Melbourne, at the Doherty Institute in Victoria, Australia, under the supervision of Professor Damian Purcell, Dr Georges Khoury and Dr Lloyd Einsiedel. The study of primary human samples was approved by the Central Australian Research Ethics Committee (HREC-17-2930) and The University of Melbourne Human Research Ethics Committee (1442830.1). The study of humanised mice was approved by the Walter and Eliza Hall Institute Animal Ethics Committee (2017.016).

The collaborative contributions to this thesis by the following individuals are as follows:

Dr Natasha Jansz & Professor Geoffrey Faulkner (Mater Institute, University of Queensland, Australia). Dr Jansz performed all ONT sequencing and bioinformatics analysis for sequence alignments of HTLV-1c provirus, with insights and advice from Prof Faulkner.

Mr Nicholas Hirons (Chartr Co., London, UK). Mr Hirons performed all bioinformatic multivariate analysis of patient plasma cytokine levels.

Dr Jennifer Juno (University of Melbourne, Australia). Dr Juno assisted in the design of the AIM T cell assay flow cytometry panel and protocol, and provided advice for analyses.

Dr Paula Ellenberg (University of Melbourne, Australia). Dr Ellenberg performed Luminex experiment and assisted in the initial FACS sorting of patient samples in PC3.

Associate Professor John Zaunders (St. Vincent's Hospital, Sydney, Australia). A.Prof Zaunders assisted in the design of the CD4+ T cell phenotyping flow cytometry panel.

Dr Lloyd Einsiedel (Alice Springs Hospital, NT, Australia). Dr Einsiedel provided all primary patient blood samples from Central Australia through recruitment of donors.

Dr Radwan Talukder & Ms Allegra Vickas (Alice Springs Hospital, NT, Australia). Dr Talukder and Ms Holloway were involved in the donor recruitment and collection of blood.

Dr James Cooney & Professor Marc Pellegrini (Walter and Eliza Hall Institute, Melbourne, Australia). Dr Cooney performed all infections, monitoring, and organ harvest of humanised mice, with insights and advice from Prof Pellegrini.

Works arising from thesis

Some text, figures and data in this thesis are published in journal articles, and were presented at national and international conferences.

List of journal articles

Hirons A., Khoury G. and Purcell D. **Human T-cell lymphotropic virus type-1: a lifelong persistent infection, yet never truly silent.** Lancet ID. 2021. 21(1): E2-E10.

Cooney J., Allison C., Doerflinger M., Hickey P., Hirons A., Dagley L., Yousef J., Yurick D., Preston S., Einsiedel L., Purcell D. and Pellegrini M. **Combination antiretroviral therapy and MCL1 inhibition mitigate HTLV-1 infection in vivo.** Cell. 2023. (In revision).

Hirons A., Yurick D., Ellenberg P., Einsiedel L., Khoury G. and Purcell D. **Unique proviral genomic features of HTLV-1c in Central Australia may impact persistence and pathogenesis.** 2023. (Manuscript ready for submission to PLoS pathogens).

List of presentations

Hirons A., Khoury G., Jansz N., Ellenberg P., Juno J., Hirons N., Zaunders J., Talukder R., Vickas A., Yurick D., Grimley S., Faulkner G., Einsiedel L. and Purcell D. **Chronic immune activation and T-cell dysfunction mediated from the HTLV-1c provirus graveyard.** Australasian Virology Society Meeting. Oral Presentation. Dec 2022. Surfers Paradise, QLD, Australia.

Hirons A., Khoury G., Ellenberg P., Jansz N., Juno J., Zaunders J., Hirons N., Talukder R., Vickas A., Yurick D., Grimley S., Faulkner G., Einsiedel L. and Purcell D. **Immune activation and dysfunction are defining characteristics of every HTLV-1c infection.** Oral Presentation. International Conference on Human Retrovirology HTLV. May 2022. Melbourne, Vic, Australia.

Hirons A., Ellenberg P., Hirons N., Talukder R., Vickas A., Zaunders J., Einsiedel L., Khoury G. and Purcell D. **The complex and distinct host-virus interplay of HTLV-1c infection**. Australian Centre for HIV and Hepatitis Virology Research Meeting. Oral Presentation. June 2021. Bowral, NSW, Australia.

Hirons A., Ellenberg P., Khoury G., Cooney J., Pellegrini M., Einsiedel L. and Purcell D. **Novel *hbz* mRNA in HTLV-1c results in loss of activation domain.** Australasian Virology Society Meeting. Oral Presentation. Dec 2019. Queenstown, New Zealand.

Hirons A., Khoury G., Yurick D., Einsiedel L. and Purcell D. **HTLV-1c infection in Australia: Call to Action**. From Bush to Bench – Indigenous Health Public Forum. Oral Presentation. July 2019. Melbourne, Vic, Australia.

Hirons A., Khoury G., Ellenberg P., Cooney J., Pellegrini M., Einsiedel L. and Purcell D. **Novel *hbz* mRNA in HTLV-1c results in loss of activation domain**. Australian Centre for HIV and Hepatitis Virology Research Meeting. Oral Presentation. May 2019. Canungra, QLD, Australia.

Hirons A., Ellenberg P., Khoury G., Cooney J., Pellegrini M., Einsiedel L. and Purcell D. **Novel *hbz* mRNA in HTLV-1c results in loss of activation domain**. Doherty Symposium 5th Anniversary. Poster Presentation. Sept 2019. Melbourne, Vic, Australia.

Hirons A., Ellenberg P., Khoury G., Cooney J., Pellegrini M., Einsiedel L. and Purcell D. **Novel *hbz* mRNA in HTLV-1c results in loss of activation domain**. International Conference on Human Retrovirology HTLV. Poster Presentation. April 2019. Lima, Peru.

Hirons A., Khoury G., Cooney J., Pellegrini M., Einsiedel L. and Purcell D. **Novel *tax/rex* and *hbz* mRNA transcripts of HTLV-1c**. Victorian Infection and Immunity Network Young Investigators Symposium. Oral Presentation. Oct 2018. Melbourne, Vic, Australia.

Hirons A., Khoury G., Cooney J., Pellegrini M., Einsiedel L. and Purcell D. **Novel *tax/rex* and *hbz* mRNA transcripts of HTLV-1c**. Australian Centre for HIV and Hepatitis Virology Research Meeting. Oral Presentation. June 2018. Yarra Valley, Vic, Australia.

Acknowledgements

This time throughout my PhD has been filled with lots of triumphs and struggles, all the while supported by a wonderful community, without whom this wouldn't have been possible.

I am very grateful to my primary supervisor, Professor Damian Purcell, for taking a chance on a freshly retired basketball player with lots of determination but limited lab experience. He inspired my interest in molecular virology, especially for HTLV-1 infection in Australia. He is incredibly passionate and has a wealth of knowledge that he is always generous to share. I admire his advocacy for HTLV-1 to be more recognised in Australia, and I hope to follow in those footsteps.

I am deeply thankful to my co-supervisor, Dr Georges Khoury, who took me on as a fresh honours student and taught me many techniques and skills needed in the lab. His attention to detail is next to none, and I hope to keep emulating that as I move forwards. Despite moving overseas early in my PhD, Georges has remained in contact and been willing to provide advice for troubleshooting experiments, guidance for project directions, and giving feedback for presentations or manuscripts.

I would like to thank Dr Lloyd Einsiedel, for his continued collaboration and support with HTLV-1 pathogenesis research. As a clinician in Central Australia, Lloyd does incredible work fostering relationships with the community and loudly advocates for patient outcomes. I greatly admire his dedication and knowledge in this field, and I look forward to continuing to work with him over the coming years.

I wish to thank my PhD committee: Chair Professor Jason Mackenzie, Professor Katherine Kedzierska and Professor Marc Pellegrini, for their time and expertise over the years. They all gave careful thoughts and suggestions which helped progress my project greatly.

I would like to thank all the members of the Purcell lab throughout the years, whether we worked directly together or shared lab space and provided each other support. Thank you to the DPL team in the early years: Behrnaz, Michelle, Charlene, Natalia and Leigh. Thank you to those DPL lab members that joined the ranks later on: Sam, Sarah, Paula, Marvin, Yadana, Dani, Ashley and Julie. In particular, I am grateful for the assistance

and support from Paula with experiments, and for being my conference roomie. I look forward to continuing to work with the current DPL lab members in the next phase of my career.

I wish to thank everyone else who assisted me throughout my PhD, whether that was providing reagents, helping with machines or data processing. The Doherty Flow Cytometry team helped immensely in processing many samples. The SPASIM committee in 2019, and especially my VC Paulina, all did a wonderful job organising some amazing events together. Thank you to my collaborator and friend Natasha for being so generous with her time, and I look forward to more if this in the future.

I have made some awesome friends during my PhD. Early on I was lucky to find Charlene and Natalia, and together we shared many memories in the lab and out, and I know we will be friends for life. Belle and Dani have been amazing friends throughout my PhD, they always have my back and for that I am very grateful. Mish was a wonderful friend and housemate, always showing such kindness and generosity. Yianni has been my honours and PhD buddy, and I am glad we shared this experience and our struggles together. I am so grateful for the everlasting friendship with Gracie, she always cheered me on during my PhD. HX supported me through the many highs and lows over the last five years, and we managed to scatter in lots of fun adventures and activities that no doubt helped get me through this period of time. My sweet dog Maple has also been by my side through it all, with unconditional love to give.

And lastly, from the bottom of my heart, I wish to acknowledge and convey my deepest gratitude to my amazing family – Mum, Dad and Nick – for lifting me up through this time. Despite never living in the same city for long through adult life with my brother Nick, I always know he is in my corner. It has been awesome to collaborate with him on some work during this time. I wish to thank him for always thinking of me, supporting me, and loving me. I have watched how dedicated and hard-working my Dad has been throughout my life –always inspiring me. I am thankful that he pushed me, provided advice, and took care of me when I needed it. I am so thankful for the unwavering love and friendship my Mum has given me throughout this time, always willing to go the extra mile - I wouldn't have made it through without her. My family is truly the best.

Table of Contents

Abstract iii
Declaration v
Preface vi
Works arising from thesis viii
List of journal articles viii
List of presentations viii
Acknowledgements x
List of Figures xix
List of Tables xxii
Abbreviations xxiii
1. Introduction 1
1.1 Introduction to HTLV-1 2
1.2 HTLV-1 lifecycle 2
1.2.1 Viral particle 2
1.2.2 Viral replication cycle 4
1.3 HTLV-1 epidemiology and transmission 7
1.3.1 HTLV-1 subtypes 7
1.3.2 Human-to-human transmission 7
1.4 HTLV-1 associated diseases 9
1.4.1 Adult T cell Lymphoma/Leukemia (ATL) 9
1.4.2 HTLV-1 associated myelopathy (HAM) 10
1.4.3 HTLV-1 associated pulmonary disease (HAPD) 11
1.4.4 HTLV-1 Uveitis (HU) 11
1.4.5 Co-infections and other associated diseases 12
1.5 HTLV-1 gene expression 14
1.5.1 HTLV-1 genomic structure 14
1.5.2 The effect of Transcriptional activator of the *pX* region (Tax) on viral expression and host immune responses 16
1.5.3 The role of HTLV-1 basic leucine zipper factor (HBZ) in survival and proliferation of infected cells 17
1.5.4 The role of Rex in modulating viral mRNA expression and persistence 18

1.5.5 The role of p12/p8 in promoting T cell activation and proliferation and in avoiding immune recognition 19
1.6 The persistence and proliferation of HTLV-1 and associated pathogenesis 20
1.6.1 HTLV-1 cell-to-cell transmission modes avoid immune system surveillance . 20
1.6.2 The survival of HTLV-1 infected cells and non-infectious proviruses 20
1.6.3 The interplay between the host genomic environment and the HTLV-1 provirus 21
1.6.4 The proliferation of HTLV-1 infected cells 22
1.7 The immune response to HTLV-1 infection 26
1.7.1 HTLV-1 infection impairs host T cell responses 26
1.7.2 Cytokine network is disrupted in HTLV-1 infection 26
1.7.3 The impact of host genetics on the progression of HTLV-1 infection 27
1.8 HTLV-1c in Australia 28
1.8.1 Central Australia is endemic for HTLV-1c infection 28
1.8.2 HTLV-1c is genetically divergent from HTLV-1a 28
1.9 Project overview 29
1.9.1 Project aims 29
2. Materials and Methods 30
2.1 Materials 31
2.1.1 Chemicals 31
2.1.2 Oligonucleotides 31
2.1.3 Probes 31
2.1.4 Antibodies 31
2.1.5 Antibody-Immobilized Magnetic beads 31
2.1.6 Peptide libraries 31
2.2 Primary patient blood samples 35
2.2.1 Patient sample collection 35
2.2.2 PBMC and plasma isolation 35
2.2.3 Diagnosis of HTLV-1 serostatus 36
2.3 Genomic DNA (gDNA) 36
2.3.1 gDNA extraction 36
2.3.2 gDNA ethanol precipitation 37
2.3.3 gDNA quantification 37
2.4 Droplet digital (dd) PCR 37
2.4.1 ddPCR primer and probe design 37

2.4.2 ddPCR droplet generation and detection 38
2.4.3 ddPCR raw data analysis 38
2.5 Activation Induced Marker (AIM) Assay 39
2.5.1 AIM assay design 39
2.5.2 HTLV-1c peptide stimulation 39
2.5.3 AIM staining and flow cytometry analysis 40
2.6 Multiplex bead-based cytokine assay 40
2.7 Statistical Analysis and Modelling 41
2.7.1 Uni- and bivariate analyses 41
2.7.2 Overview of multivariate analyses 41
2.7.3 Correlation heatmap 41
2.7.4 Principal Component Analysis (PCA) 42
2.7.5 Elastic net logistic regression model 42
2.7.6 Partial Least Squares Discriminant Analysis (PLS-DA) 42
2.7.7 Logistic regression analysis 43
2.8 CD4+ T cell phenotyping 43
2.9 Single provirus amplification (SPA) PCR assay 44
2.9.1 Design and optimisation of SPA assay 44
2.9.2 SPA PCR protocol 44
2.9.3 Visualisation and purification of single provirus amplifications 45
2.10 Long-read HTLV-1c provirus sequencing 45
2.10.1 Oxford Nanopore Technologies (ONT) sequencing 45
2.10.2 Sequence alignments and analysis 46
2.11 HTLV-1c-infected humanised mouse model 47
2.11.1 Maintenance of hu-NSG mouse strain 47
2.11.2 Generation of humanised mice 47
2.11.3 Infection of hu-NSG mice with HTLV-1c 47
3. Characterising the host immune response to HTLV-1c infection . 49
3.1 Chapter three background 50
3.1.1 CD4+ T cell responses 50
3.1.2 CD8+ T cell responses 51
3.1.3 Double negative (DN) T cell responses 52
3.1.4 Cytokines and chemokines 52
3.2 Chapter three aims 53
3.3 Characterising the HTLV-1c-specific T cell responses 53

3.3.1 HTLV-1c infection causes expansion of CD4+ Tmem and cTFH phenotypes and dysregulation of Th1 frequency 57
3.3.2 HTLV-1c infection causes significant background CD4+ T cell activation 58
3.3.3 HTLV-1c infection induces Tax- and Env-specific CD4+ T cells 62
3.3.4 HTLV-1c infection elicits a robust CD8+ T cell response 67
3.3.5 DN T cells are activated and virus-specific in HTLV-1c infection 70
3.4 Determining the cytokine response to HTLV-1c infection 73
3.4.1 Univariate analysis shows elevated plasma cytokines in HTLV-1c infection . 75
3.4.2 Multivariate analyses expose dysregulated cytokine networks in HTLV-1c infection 76
3.4.3 Univariate analyses of plasma cytokines does not reveal any trends for HTLV-1c-associated bronchiectasis 79
3.4.4 Multivariate analyses reveal unique cytokine profile of HTLV-1c-associated bronchiectasis distinct from other HTLV-1a associated diseases 82
3.5 Chapter three discussion 85
3.5.1 HTLV-1c causes high levels of immune activation and dysfunctional CD4+ T cell responses 86
3.5.2 HTLV-1c elicits very strong CD8+ T cell responses 87
3.5.3 DN T cells may be an untapped reservoir of HTLV-1c infection and inflammation 88
3.5.4 Future directions of the HTLV-1c-specific T cell response studies..... 89
3.5.5 Distinct and dysfunctional host cytokine response in HTLV-1c infection..... 90
3.5.6 Unique cytokine signature of HTLV-1c-associated bronchiectasis 92
3.5.7 Dysfunctional Th1 response in HTLV-1c infection 95
3.5.8 Chapter three conclusion 95
4. Characterising the phenotype of HTLV-1c infected cells and the proviral structure within these cells 97
4.1 Chapter four background 98
4.1.1 HTLV-1 exhibits CD4+ T cell tropism 98
4.1.2 HTLV-1-infected cells are found in tissue reservoirs..... 99
4.1.3 Defective provirus is present in HTLV-1a infection.....100
4.2 Chapter four aims.....101
4.3 Characterising the phenotype of HTLV-1c infected cells 102
4.3.1 Strategy for infected cell phenotype characterisation was structured around CCR4 marker 102
4.3.2 Lung homing proxy CD4+ T cell phenotype is expanded in HTLV-1c infection 106

4.3.3 HTLV-1c provirus is enriched in both CCR4+ and CCR4- CD4+ T cell reservoirs 113
4.4 Characterising the proviral landscape in HTLV-1c infection in vivo 115
4.4.1 *tax* and *hbz* gene regions are detected at higher levels than *gag* or *env* in HTLV-1c proviral reservoirs by ddPCR 115
4.4.2 High levels of defective provirus are present in all HTLV-1c infected donors 118
4.4.3 HAPD+ donors show greater retention of HTLV-1c *pX* region in defective provirus than HAPD- 122
4.4.4 Exclusively full-length HTLV-1c provirus detected in CCR4- reservoir 125
4.4.5 HTLV-1c defective provirus breakpoints share LTR-like motifs 128
4.4.6 Chimeric HTLV-1c-human genome provirus sequences occur most commonly in centromeric regions 130
4.5 Chapter four discussion 134
4.5.1 Novel characterisation of the phenotype of HTLV-1c infected cells 134
4.5.2 Detection of multiple HTLV-1c gene coding regions with ddPCR prompted redesign of HTLV-1 PVL assay in Australia 136
4.5.3 SPA and ONT sequencing methodology provides high resolution of HTLV-1c proviruses 137
4.5.4 HTLV-1c provirus graveyard with large internal deletions is capable of driving pathogenesis, while CCR4- cells maintain infection 139
4.5.5 Defective HTLV-1c proviruses may have arisen from homologous recombinational deletion 141
4.5.6 Novel HTLV-1c-human genome chimeric proviruses may impact gene expression and pathogenesis 142
4.5.7 Chapter four conclusion 144
5. Assessing the ability of a novel HTLV-1c humanised mouse model to recapitulate human in vivo infection 146
5.1 Chapter five background 147
5.1.1 Mouse models of infection have been widely used in HTLV-1 research 147
5.1.2 Transgenic mouse models elucidated functions of HTLV-1 Tax and HBZ proteins in vivo 147
5.1.3 Immunocompetent mice modelled persistent HTLV-1 infection 148
5.1.4 Immunodeficient mouse models progressed understanding of ATL development and therapeutic interventions 148
5.1.5 Humanised mouse models have greatly advanced HTLV-1 research 150
5.2 Chapter five aims 151
5.3 Characterising the phenotype of HTLV-1c infected cells in humanised mice 152

5.3.1 Generating an HTLV-1c infected hu-NSG mouse model 152
5.3.2 Immunophenotype of hu-NSG mice infected with HTLV-1c mirrors human in vivo infection 154
5.3.3 HTLV-1c provirus is enriched in CCR4+ and CCR4- phenotypes of CD4+ T cells in hu-mice splenocytes 159
5.4 Characterising the proviral landscape of hu-mice infected with HTLV-1c 161
5.4.1 Envelope deletions accumulate over time in HTLV-1c-infected hu-mice 161
5.4.2 In vivo HTLV-1c proviral deletions are more extensive in humanised mice than in humans 163
5.4.3 HTLV-1c defective provirus breakpoints in hu-mice do not have common sequence motifs 167
5.4.4 Chimeric HTLV-1c-human genome proviruses detected in hu-mouse model 167
5.5 Chapter five discussion 171
5.5.1 Immunophenotyping in HTLV-1c hu-mouse model 171
5.5.2 Extensive internal proviral deletions and chimeras detected in HTLV-1c-infected humanised mice 173
5.5.3 Future plans to utilise the humanised mouse model to map integration sites and epigenetic modifications 174
5.5.4 Chapter five conclusion 175
6. General thesis discussion 177
6.1 The importance of HTLV-1c research in Australia 178
6.1.1 General comments 178
6.1.2 HTLV-1c infection is part of a complex health disparity in Central Australia 178
6.2 Novel characteristics of HTLV-1c infection in Central Australia 179
6.2.1 Chronic T cell activation in all HTLV-1c infection 179
6.2.2 HTLV-1c bronchiectasis is characterised by immune exhaustion 180
6.2.3 *hbz* gene drives HTLV-1c-associated bronchiectasis development 181
6.2.4 CCR4- CD4+ T cells may be a persistent reservoir of HTLV-1c provirus 184
6.2.5 Novel chimeric HTLV-1c-hg provirus may contribute to immune dysfunction and pathogenesis 184
6.3 Translational impacts of HTLV-1c pathogenesis research 185
6.3.1 Improvements to HTLV-1 PVL assay in Australian reference laboratories 185
6.3.2 Viral therapeutics should target *hbz* expression 185
6.3.3 The humanised mouse model of HTLV-c infection will be an important tool for pre-clinical testing of preventative and therapeutic drugs 186

6.3.4 Changes are required to improve the HTLV-1 public health strategy in Australia.....187
6.4 Concluding thesis remarks 188
References..... 189

List of Figures

Figure 1.1 - Structure of the mature HTLV-1 virion. 3
Figure 1.2 – Overview of the productive HTLV-1 replication cycle. 6
Figure 1.3 – The genomic structure of HTLV-1 and *pX* proteins role in proliferation and persistence of infection..... 16
Figure 1.4 – The balance between HTLV-1 viral spread and viral gene expression leading to cellular proliferation and dysfunction.. 25
Figure 3.1 – Flow cytometry gating strategy for HTLV-1c activation induced marker assay.. 56
Figure 3.2 – HTLV-1c infection impacts the frequency of cTFH, Tmem and Th1 CD4+ phenotypes in primary patient PBMCs. 60
Figure 3.3 – HTLV-1c infection results in significant chronic CD4+ T cell background activation in primary patient PBMCs. 61
Figure 3.4 – Tax-specific CD4+ T cell responses are present in primary HTLV-1c infection. 65
Figure 4.1 – Fluorescent activated cell sorting (FACS) gating strategy for HTLV-1c proviral reservoir phenotyping study. 104
Figure 4.2 – Effector/memory CD4+ T cells in patient PBMCs are activated in HTLV-1c infection. 109
Figure 4.3 – FOXP3+ proxy phenotype CD4+ T cells in patient PBMCs do not expand during HTLV-1c infection but are highly activated. 110
Figure 4.4 – CD4+ T cells with potential lung trafficking phenotype are significantly expanded and activated in HTLV-1c infection. 111
Figure 4.5 – CCR4- effector/memory CD4+ T cells are highly activated in HTLV-1c infection. 112
Figure 4.6 – HTLV-1c provirus is enriched in CCR4+ and CCR4- phenotypes of effector/memory CD4+ T cells in patient PBMCs. 114
Figure 4.7 – Normalised gene retention of the HTLV-1c provirus pool in effector/memory CD4+ T cell reservoirs in human PBMCs.....117
Figure 4.8 – Single provirus amplification assay with Oxford Nanopore Technologies sequencing strategy.....120

Figure 4.9 – The overall HTLV-1c proviral landscape in primary human PBMCs is highly defective. 121
Figure 4.10 – HAPD+ patients have higher overall HTLV-1c proviral genome coverage in PBMCs than HAPD- donors.....124
Figure 4.11 – The *pX* region of the HTLV-1c genome has higher retention in the total provirus pool of HAPD+ donors than HAPD-. 124
Figure 4.12 – The HTLV-1c proviral landscape is highly defective in FOXP3+ proxy CD4+ T cell phenotype reservoir of primary human PBMCs. 126
Figure 4.13 –The HTLV-1c proviral landscape is highly defective in lung homing proxy CD4+ T cell phenotype reservoir of primary human PBMCs. 127
Figure 4.14 – HTLV-1c provirus sequences in CCR4- effector/memory CD4+ T cells of primary human PBMCs are full-length. 128
Figure 4.15 – Defective HTLV-1c proviruses from primary human PBMCs display breakpoints with LTR-like motifs. 129
Figure 4.16 – Novel HTLV-1c-human genome chimeric proviruses were amplified from primary human PBMCs. 131
Figure 4.17 – Junction sequences of human genome component of novel HTLV-1c chimeric proviruses from primary patient PBMCs**.** 134
Figure 5.1 – Generation of novel HTLV-1c humanised mouse model. 153
Figure 5.2 – Lung homing proxy phenotype of effector/memory CD4+ T cells expand in humanised mice infected with HTLV-1c. 157
Figure 5.3 – All effector/memory CD4+ T cells are highly activated in humanised mice infected with HTLV-1c and correlated to proviral load. 158
Figure 5.4 – Enrichment of HTLV-1c provirus in CCR4± phenotype CD4+ T cell reservoirs in humanised mouse splenocytes. 160
Figure 5.5 – Normalised gene retention of HTLV-1c provirus in humanised mouse splenocytes shows accumulation over time of proviruses with deleted *env*.....162
Figure 5.6 – The HTLV-1c proviral landscape in humanised mice splenocytes is highly defective when compared to primary human PBMCs. 165
Figure 5.7 – HTLV-1c proviral sequences have large internal deletions in humanised mice splenocytes.....166
Figure 5.8 – Breakpoint analysis of HTLV-1c defective proviruses from hu-mice model. 168

Figure 5.9 – Novel HTLV-1c-human genome chimeric proviruses were amplified from humanised mouse splenocytes. 170

List of Tables

Table 1.1 - Summary of HTLV-1 transmission modes. 9
Table 1.2 – Summary of HTLV-1-associated diseases. 13
Table 2.1 – Oligonucleotide primers used in this thesis. 33
Table 2.2 – Fluorescent TaqMan® hydrolysis probes used in this thesis. 33
Table 2.3 – Flow cytometry antibodies used in this thesis. 34
Table 2.4 – Human cytokine/chemokine magnetic beads used in this thesis..... 35
Table 3.1 – Cohort demographics and clinical characteristics of HTLV-1c T cell response study. 55
Table 3.2 – Cohort demographics and clinical characteristics of HTLV-1c host cytokine response study. 74
Table 3.3 – Logistic regression analysis demonstrates that HTLV-1c infection disrupts the host cytokine network. 79
Table 3.4 – Demographics and clinical characteristics of HTLV-1c+ donors of host cytokine response study. 80
Table 3.5 – HTLV-1c-associated bronchiectasis shows a cytokine profile distinct from other HTLV-1a-associated diseases. 85
Table 3.6 – Summary of plasma cytokine networks in HTLV-1-associated diseases..... 95
Table 4.1 - Demographics and clinical characteristics of donors in the HTLV-1c proviral reservoir phenotyping cohort. 105
Table 4.2 – Summary and annotations of HTLV-1c-human genome chimeric proviruses amplified from primary human PBMCs. 132
Table 5.1 – Immunocompromised mouse strains commonly used in HTLV-1 xenograft and humanised mouse models 151
Table 5.2 – Summary and annotations of HTLV-1c-human genome chimeric proviruses amplified from humanised mouse splenocytes. 169
Table 6.1 – Summary of the components of HTLV-1c infection that contribute to immune exhaustion and pathogenesis in associated bronchiectasis.....183

Abbreviations

(+)ssRNA	Positive sense single-stranded RNA
AAB	Assigned at birth
Ab	Antibody
AC	Asymptomatic carrier
AICD	Activation induced cell death
AIM	Activation induced marker
AP-1	Activator protein 1
APC	Antigen presenting cell
ART	Antiretroviral therapy
ATL	Adult T cell leukaemia/lymphoma
att	Attachment
BALF	Bronchoalveolar lavage fluid
BCR	B cell receptor
Bex	Bronchiectasis
BLV	Bovine leukemia virus
bp	Base pair
BRG	BALB/c.Cg-*Rag2*$^{-/-}$-IL2Rγ$^{-/-}$
BSI	Blood stream infection
CA	Capsid protein
CADM	Cell adhesion molecule
CAHREC	Central Australian Human Research Ethics Committee
CBP	CREB binding protein
CCR	CC chemokine receptor
censat	centromeric satellite
Chr	Chromosome
CNS	Central nervous system
Co-inf	Co-infection
CREB	Cyclic AMP response element binding protein
CRM1	Chromosomal maintenance 1
CSF	Cerebrospinal fluid

CTCF	CCCTC-binding factor
cTFH	Circulating T follicular helper cells
CTL	Cytotoxic T lymphocyte
CV	Cross-validation
CXCL	CXC motif chemokine ligand
DC	Dendritic cell
DDIT3	DNA damage inducible transcript 3
ddPCR	Droplet digital PCR
DMSO	Dimethylsulfoxide
DN	Double negative
dNTP	Deoxynucleoside triphosphate
dsDNA	Double stranded DNA
E/M	Effector/memory
EDTA	Ethylenediaminetetraacetic acid
EIA	Enzyme immunoassay
Env	Envelope
FBS	Fetal bovine serum
FDR	False discovery rate
FOXP3	Forkhead box p3 protein
gDNA	Genomic DNA
GLUT1	Glucose transporter type 1
gp	Glycoprotein
HAID	HTLV-1-associated infective dermatitis (table)
HAM	HTLV-1-associated myelopathy
HAPD	HTLV-1-associated pulmonary disease
HBV	Hepatitis B virus
HBZ	HTLV-1 basic leucine zipper factor
HCV	Hepatitis C virus
HERV	Human endogenous retrovirus
hg	Human genome
HI	Heat inactivated
HIS	Human immune system
HIV-1	Human immunodeficiency virus type 1

HLA	Human leukocyte antigen
HOR	Higher order repeat
Hsat	Human satellite
HSC	Hematopoietic stem cell
HSP	Heparin sulfate proteoglycan
HTLV-1	Human T cell lymphotropic virus type 1
HU	HTLV-1-uveitis
hu	Humanised
ICAM	Intracellular adhesion molecule
IDU	Injecting drug user
IFN	Interferon
IL	Interleukin
IL2R	IL-2 receptor
IN	Integrase
IP	Interferon gamma induced protein
IQR	Interquartile range
kb	Kilobase
LFA	Lymphocyte function-associated antigen
LINE	Long interspersed element
lnc	Long non-coding
LTR	Long terminal repeat
LV	Latent variable
MA	Matrix protein
mAb	Monoclonal antibody
MCL-1	Myeloid cell leukemia 1
MCP-1	Monocyte chemoattractant protein 1
MHC	Major histocompatibility complex
mRNA	Messenger RNA
NC	Nucleocapsid
NFAT	Nuclear factor of activated T cells
NF-κB	Nuclear factor kappa B
NGS	Next generation sequencing
NHMRC	National Health and Medical Research Council

NHP	Non-human primate
NK	Natural killer
NMD	Nonsense-mediated decay
NOD	Non-obese diabetic
NOG	NODShi-*scid*-IL2Rγ$^{-/-}$
NRL	National Reference Laboratory
NRP-1	Neuropilin-1
NSG	NOD-*scid*-IL2Rγ$^{-/-}$
NTC	Non-template control
OCI	Oligoclonality index
ONT	Oxford Nanopore Technologies
OR	Odds ratio
PBL	Peripheral blood lymphocytes
PBMC	Peripheral blood mononuclear cell
PBS	Phosphate-buffered saline
PCA	Principal component analysis
PCR	Polymerase chain reaction
PD-1	Programmed cell death protein 1
PD-L1	Programmed death ligand 1
PLS-DA	Partial least squares discriminant analysis
Pro	Protease
PVL	Proviral load
QC	Quality control
RBC	Red blood cell
RF10	RPMI + 10% HI-FBS
RM	Repeated measures
RN7SL1	Signal recognition particle 70SL1
RPMI	Rosswell Park Memorial Institute
RPS29	40S ribosomal protein subunit 29
RT	Reverse transcriptase
RXRE	Rex responsive element
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SCID	Severe combined immunodeficiency

SDHC	Succinate dehydrogenase complex subunit
SEB	Staphylococcal enterotoxin B
SINE	Short interspersed element
SNP	Single nucleotide polymorphism
SPA	Single provirus amplification
Std	Standard
STI	Sexually transmitted infection
STLV	Simian T cell lymphotropic virus
SU	Surface unit
TAE	Tris-acetate-ETDA
Tax	Transcriptional activator of the *pX* region
Tc	Cytotoxic T cells
TCR	T cell receptor
TF	Transcription factor
Tg	Transgenic
TGFβ	Transforming growth factor beta
Th	Helper T cell
TIGIT	T cell immunoreceptor with immunoglobulin and ITIM domain
TM	Transmembrane
Tmem	Memory T cell
TNF	Tumor necrosis factor
Treg	Regulatory T cell
tRNA	Transfer RNA
U3	Unique 3’
UTR	Untranslated region
VL	Viral load
wpi	Weeks post infection

1. Introduction

1.1 Introduction to HTLV-1

Human T cell lymphotropic virus type 1 (HTLV-1) is a complex retrovirus that causes lifelong infection in around 10 million people worldwide [1]. HTLV-1 infects immune cells, especially T lymphocytes and myeloid cells, which can lead to overt immune damage that eventually develops as adult T cell leukaemia (ATL) [2] in around 5% of infected individuals [3, 4], or HTLV-1-associated myelopathy (HAM) [5] in a further 3% [6, 7]. Other patients sustain more subtle immune damage and develop a range of serious associated sequelae (shown to occur at a lower frequency in people without HTLV-1 than those with infection), including uveitis [8], bronchiectasis (Bex) [9, 10], infectious dermatitis and scabies [11, 12], and blood stream co-infections [13]. Even infection in people considered to be asymptomatic carriers (ACs) of HTLV-1, is strongly associated with premature mortality without a recognised pathogenesis [14]. Endemic prevalence among Central Australian First Nations communities is among the highest reported for this preventable viral infection [15-17]. The complex virology of HTLV-1, the host immune response, and host-virus interplay make effective therapy, prevention, and viral eradication more challenging than with other blood-borne and sexually transmitted persistent viruses, such as human immunodeficiency virus type 1 (HIV-1), hepatitis B virus (HBV), and hepatitis C virus (HCV).

1.2 HTLV-1 lifecycle

1.2.1 Viral particle

HTLV-1 is classified as a member of the *Retroviridae* family, and falls under the Orthoretroviridae subfamily, and the Deltaretrovirus genus, together with simian T-lymphotropic viruses (STLV) and bovine leukemia virus (BLV) [18].

Similar to other retroviruses, the HTLV-1 viral particle is comprised of key structural [19], enzymatic [20] and genomic components [21] (**Figure 1.1**). The envelope (Env) glycoproteins on the outside of the viral particle form stable trimers [22]. Each Env monomer is comprised of two domains covalently joined with a disulfide bond [23]: a surface unit (SU) glycoprotein, gp46, responsible for viral attachment and entry into target host cells [24]; and a transmembrane (TM) glycoprotein (gp21), responsible for

tethering gp46 to the lipid bilayer of the viral particle, which originates from a host cell membrane [24]. The matrix (MA) protein associates with the lipid bilayer, forming an extra inner layer on the inside of the viral particle [25]. The capsid (CA) p24 proteins form a polyhedrally shaped lattice core containing the nucleocapsid (NC) structure, which associates with the two copies of (+)ssRNA HTLV-1 genome, forming a protective coating [26]. Inside the capsid, three viral enzymes vital to the HTLV-1 lifecycle; integrase (IN), reverse transcriptase (RT) and protease (Pro) can be found [27]. Alongside these viral enzymes, the HTLV-1 virion also packages host tRNA-pro and other RNAs and proteins, some of which initiate and support productive infection [28].

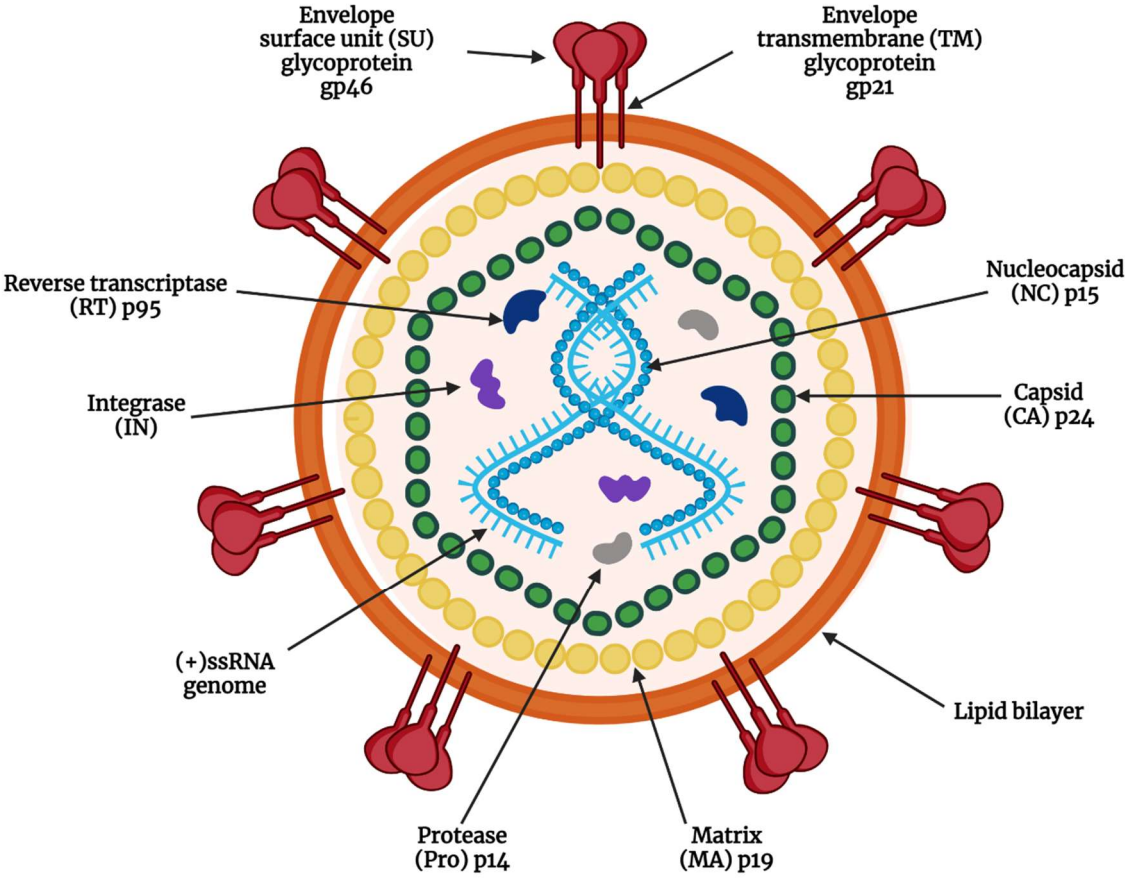


Figure 1.1 - Structure of the mature HTLV-1 virion. On the surface of the virion, envelope glycoproteins form stabilised trimers, with each subunit comprised of a surface unit (SU) glycoprotein (gp46) covalently bonded to a transmembrane (TM) glycoprotein (gp21). The envelope trimers are embedded in the outer lipid bilayer, which originates from the host cell. Associating with the lipid bilayer is the matrix (MA), formed by p19 proteins. A polyhedral-shaped lattice core is comprised of capsid (CA) proteins (p24). Within the CA is the nucleocapsid (NC) structure, which associates with and protects two copies of the (+)ssRNA HTLV-1 genome. Also contained in the CA are three viral enzymes, integrase (IN), reverse transcriptase (RT) and protease (Pro). Created with Biorender.com

1.2.2 Viral replication cycle

The HTLV-1 life cycle follows the classical retroviral pathway (**Figure 1.2**). In overview, viral replication begins with attachment of the viral particle to the target host cell. HTLV-1 preferentially infects CD4+ T cells [29], but due to the universal nature of its cellular receptors – heparin sulfate proteoglycans (HSPs) [30, 31], neuropilin-1 (NRP-1) [32] and glucose transporter 1 (GLUT1) [33], it is also known to infect other cell types expressing these receptors, including CD8+ T cells [34, 35], dendritic cells (DCs) [36], macrophages and monocytes [37]. It is not currently well understood how or why HTLV-1 exhibits this tropism without target-specific receptors, but there are host co-factors which assist in later steps of viral replication and transmission, and this will be explored in more detail in subsequent sections of this thesis.

Entry into the host cell is initiated by the binding of the HTLV-1 Env SU gp46 to HSP on the host cell [38], triggering the recruitment of NRP-1 [32, 39] (**Figure 1.2A**). This process is thought to induce a conformational change in the envelope glycoprotein, facilitating the interaction of envelope TM with GLUT1 and ultimately the fusion of the viral and cellular membranes [33, 40] (**Figure 1.2B**). The viral capsid structure is then released into the host cell cytoplasm and uncoating occurs (**Figure 1.2C**). During this time the viral (+)ssRNA genome is reverse transcribed into dsDNA by viral RT using the tRNA as a primer, that places long terminal repeats (LTRs) at either end of the linear dsDNA viral genome. A dimer of IN [41] binds to attachment (att) sites in the viral LTR dsDNA and assists in nuclear import [42] (**Figure 1.2D**). IN further mediates the integration of the viral dsDNA genome into the host genome [43], where the viral 5' and 3' LTRs flank the provirus (**Figure 1.2E**). An intact, integrated provirus is required for the expression of new infectious virions [44].

During active replication, proviral transcription occurs using the host cellular RNA polymerase II (**Figure 1.2F**), and several host transcriptional co-factors, including cyclic AMP response element binding protein (CREB) [45], and CREB binding protein (CBP)/p300 [46]. The role that viral proteins play in HTLV-1 gene expression and life cycle will be explored in later sections of this thesis chapter. Viral transcription is initiated at the 5' or 3' LTRs, which contain viral promoter and enhancer elements, and hence occurs in a bidirectional manner producing sense or antisense mRNA transcripts. The complexity of the HTLV-1 genome facilitates alternative splicing to produce various

unspliced, singly, and doubly spliced mRNA products from the one 9 kilobase (kb) proviral genome. These viral mRNAs are exported from the nucleus into the cytoplasm, where they are translated by host ribosomes [47] (**Figure 1.2G**), and regulation of this process will be discussed in later sections. The viral structural and enzymatic proteins (Gag, Pro, RT, IN and Env) along with full-length viral genomic RNAs then translocate to the host cell plasma membrane (**Figure 1.2H**), where they form immature and non-infectious virus particles [48] (**Figure 1.2I**). Full infectivity is attained after Gag and Gag-Pol is cleaved by the viral Pro into its virion structural components and three functional viral enzymes [49], causing the CA proteins to reassemble and form the capsid core during virus budding and maturation (**Figure 1.2J**).

Unlike HIV-1 infection, which results in a very high plasma viral load (VL) when left untreated with antiretroviral therapy (ART) [50], HTLV-1 produces few cell-free virions and exists mostly within infected cells as a provirus [44], which in this form is extremely stable [51]. HTLV-1 conducts an infection that largely expands in the host through stimulating proliferation of infected cells and preventing death of infected cells by apoptosis, and also through various cell-to-cell transmission mechanisms, all the while avoiding immune system detection.

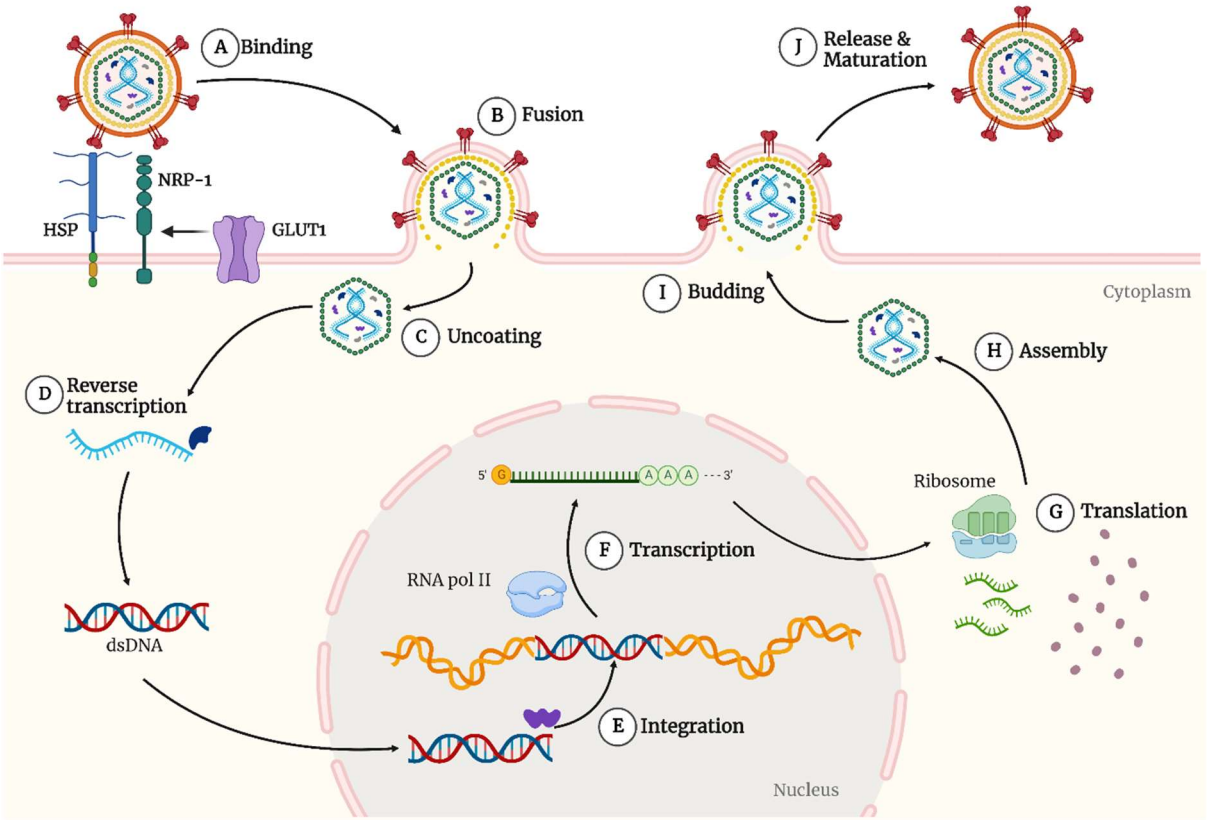


Figure 1.2 – Overview of the productive HTLV-1 replication cycle. (A) Binding of the HTLV-1 virion to the host cell: HTLV-1 Env surface glycoprotein binds to HSPs on the host cell, recruiting NRP-1. (B) A subsequent Env conformational change allows the Env transmembrane glycoprotein to interact with GLUT1 and fuse the viral particle with the host cell. (C) Uncoating of the viral capsid occurs once it is released into the cytoplasm of the host cell. (D) The HTLV-1 (+)ssRNA is reverse transcribed by viral enzyme reverse transcriptase into dsDNA. (E) Viral integrase binds to the dsDNA genome and facilitates nuclear import and integration into the host cell genome, forming the provirus. (F) Proviral transcription occurs from the 5' and 3' long terminal repeat viral promoters, with cellular RNA polymerase II (RNA pol II) and other host and viral co-factors. Alternative splicing of viral mRNAs occurs to produce different transcripts. Nuclear export of intron-containing viral mRNAs is facilitated by the HTLV-1 protein Rex. (G) Translation of viral proteins by host ribosomes occurs in the cytoplasm. (H) Viral structural and enzymatic proteins, and two full length (+)ssRNA genomes, are assembled at the cell plasma membrane. (I) Immature virus particles are formed as budding occurs. (J) The viral particle is released from the host cell and maturation occurs when gag and gag-pol are cleaved by the viral protease to form the capsid structure, whereby the virion is mature and infectious. Env, envelope; HSP, heparin sulfate proteoglycan; NRP-1, neuropilin-1; GLUT1, glucose transporter 1; RNA pol II; RNA polymerase II. Created with Biorender.com

1.3 HTLV-1 epidemiology and transmission

1.3.1 HTLV-1 subtypes

HTLV-1 originated from STLV-1 infected non-human primates (NHP), such as macaques, gorillas, chimpanzees, and orangutans. In some cases, STLV-1 infection can lead to the development of ATL in NHPs [1]. It is thought that the HTLV-1 subtypes observed today arose from several interspecies transmission events [52], followed by a period of evolution in humans [1]. Indeed, there may be further interspecies transmission events that are occurring periodically in populations with close and frequent contact with NHPs [52, 53]. However, the geographical distribution of the current HTLV-1 subtypes follows early human migration patterns [54, 55], and there are numerous endemic foci around the world.

The most common and widespread HTLV-1 subtype is Cosmopolitan subtype-A [56], which is found in Southwest Japan [57], the Caribbean [58], North and West Africa [59, 60], Central and South America [61, 62], and the Middle-East [63, 64]. While there are multiple subgroups within subtype-A [61], it is actually the least divergent subtype of HTLV-1, suggesting a relatively more recent NHP to human transmission, followed by increased intercontinental mobility of humans [1]. In Central Africa, the most common HTLV-1 subtypes are subtype-B [65] and -D [60], while rare subtypes E, F, and G have been reported [65]. The most genetically divergent subtype, subtype-C, is found in Central Australia [66, 67], the Solomon Islands [68], Vanuatu [69], and Papua New Guinea [70], and may be the most ancient subtype [1]. HTLV-1 has likely persisted in these isolated populations for tens of thousands of years.

1.3.2 Human-to-human transmission

HTLV-1 infections predominantly transmit with the transfer of infected cells. Like other blood borne viruses such as HIV-1, HTLV-1 can be transmitted through three main modes: sexual transmission, mother-to-child transmission, and transmission through infected blood products, summarised in **Table 1.1**. Transmission risk from these documented modes is linked to the relative levels of infected cells present, as outlined below, and is further addressed in subsequent chapters of this thesis.

Sexual transmission is thought to be the most common mode of horizontal transmission worldwide [71-73], with studies showing that it most frequently occurs as male-to-female

(sex assigned at birth) transmission [72, 74], although high transmission rates have also been reported for male-to-male [75]. This may be due to the higher number of infected cells present in seminal fluid, rather than in vaginal secretions, and an abundance of prostaglandins, which participate in a positive feedback loop with HTLV-1 LTR activation [76, 77]. Furthermore, a high proviral load (PVL) of HTLV-1 is one of the primary risk factors for sexual transmission [78]. Moreover, the co-infection of other sexually transmitted diseases, particularly those with genital lesions, is another risk factor for transmission [71, 75], as it not only provides an exit/entry point for infected cells to transmit, but also causes inflammatory responses in- and recruitment of infected cells to-the genital regions [79].

Mother-to-child transmission can occur through breast-feeding [80, 81], with an estimated 10-25% of babies who are breast-fed by HTLV-1-infected mothers developing the infection [82]. Risk factors for transmission through breast-feeding include PVL of infected lymphocytes and myeloid cells in the mother's milk and peripheral blood, and breast-feeding for more than 6 months, while risk may be reduced by high levels of anti-HTLV-1 antibodies in the mother's serum [83-85]. However, the risk of transplacental transmission of HTLV-1 from mother-to-child is low [83], most likely due to the limited transfer of infected cells through the placenta [86, 87].

HTLV-1 transmission through infected blood products can occur through a variety of different modes. Organ transplantation from HTLV-1-infected individuals can result in the seroconversion of recipients and rapid development of HAM [88-91]. It is postulated a rapid onset of HTLV-1-associated disease may in part be due to the immunosuppressed conditions of recipients [90], although not exclusively [92]. Similarly, the virus can be transmitted through HTLV-1+ blood transfusions, resulting in swift seroconversion of recipients and progression to HAM [93, 94]. HTLV-1 can transmit through shared needles among injecting drug users (IDUs), and risk increases with duration of practice [95, 96]. HTLV-1 is also thought to transmit during cultural rituals such as self-flagellation [97, 98], although the extent of transmission through these routes in areas of high HTLV-1 prevalence remains unknown due to the sacristy of such cultural rites and ceremonies.

Table 1.1 - Summary of HTLV-1 transmission modes.

Transmission	Main modes	Risk factors
Sexual intercourse	Predominantly transmitted by HTLV-1+ males (AAB) [72, 74, 75]	High PVL in blood [78]; Infected cells in seminal fluid [76, 77]; Unprotected sex; Other concurrent STIs (lesions) [71, 75]
Mother-to-child	Breast-feeding by HTLV-1+ mother [80, 81]	High PVL in blood and breast milk; Long duration of breast-feeding; Low levels of anti-HTLV-1 Ab in breast milk [83-85]
Infected blood products	IDU needle sharing [95, 96]; Organ transplantation [88-91]; Blood transfusion [93, 94]; Cultural practices	High PVL in blood; Immunosuppression of organ/blood donor recipients [90]

AAB, assigned at birth; IDU, injecting drug user; PVL, proviral load; STI, sexually transmitted infection; Ab, antibody.

1.4 HTLV-1 associated diseases

HTLV-1 is the etiological agent for two serious and life-threatening conditions, adult T cell lymphoma/leukemia (ATL) and HTLV-1 associated myelopathy (HAM). HTLV-1 infection is also associated with a wide range of other diseases, co-infections and co-morbidities. The categories, clinical features, viral pathogenesis, and current treatments of HTLV-1-associated diseases are summarised in **Table 1.2**. Nonetheless, it has become clear in recent years that all HTLV-1 infections result in higher all-cause mortality, even in the absence of a clearly defined associated disease [14].

1.4.1 Adult T cell Lymphoma/Leukemia (ATL)

ATL is a highly aggressive neoplasm resulting from the malignant transformation of HTLV-1-infected CD4+ T cells. It affects around 5% of all HTLV-1 infected individuals [3, 4]. The clinical features of ATL are severe and widespread, including skin lesions, hepatosplenomegaly, lymphadenopathy, and hypercalcemia, along with a plethora of more general symptoms such as pain, fever and gastrointestinal upset [99]. There are four major subtypes of ATL: chronic, smouldering, lymphoma and acute, which are defined by stringent diagnostic characteristics [100]. The chronic and smouldering subtypes may progress to lymphoma and acute ATL, while patients with the latter two have a very poor prognosis, with a median survival rate of less than 9 months, despite any treatment they may receive [101].

Factors that contribute to HTLV-1 infection leading to ATL have been extensively investigated. It is estimated that ATL takes between 30-50 years to develop, with an accumulation of multiple provirus-associated host genetic mutations. Because of this, acquiring HTLV-1 early in life is a risk factor for ATL. On a cellular level, ATL develops through combinations of viral pathogenesis mechanisms that include inhibiting apoptosis and functions of host DNA repair machinery, increasing cell proliferation, and disrupting cell cycle checkpoints, which ultimately result in malignant transformation [102].

There are currently no licensed effective treatments for ATL. The most successful strategy for prolonging the life of an individual with either chronic or smouldering ATL is to use a combination ART of zidovudine and interferon alpha (IFN-α) [103-105]. The acute and lymphoma subtypes do not respond well to ART or chemotherapy, and the prognosis remains poor or worsens further [106]. Some patients are reported to respond to allogenic hematopoietic stem cell (HSC) transplantation and have increased life expectancy [107-109]. Numerous monoclonal antibody therapies have been trialled to target proviral infected cells, including anti-CCR4 [110, 111], anti-CD25 [112-114], and anti-transferrin [115], but none have consistently reduced PVL or improved prognosis and survival [116].

1.4.2 HTLV-1 associated myelopathy (HAM)

HTLV-1-associated myelopathy (HAM) is a chronic and debilitating neurological disease that develops in approximately 3% of HTLV-1-infected individuals [117]. Clinical symptoms include, but are not limited to, progressive leg weakness and spasticity, and both bladder and bowel dysfunction [118].

HAM pathogenesis is characterised by CD4+ and CD8+ T cell infiltration into inflammatory lesions of the spinal cord [119]. Many of these infiltrating cells are infected with HTLV-1: HAM patients have higher PVL in cerebrospinal fluid (CSF) than PBMCs [120]. The immune response to HTLV-1 infected cells in the central nervous system (CNS) results in high levels of inflammatory cytokines and drives degeneration of the spinal cord [117].

Currently, anti-inflammatory steroid treatments, such as Prednisone and other corticosteroids, are used to treat HAM [121, 122]. Moreover, individual clinical trials

have seen some positive results with reduced PVL or slowed symptom progression, using strategies such as antiretrovirals [123], interferons [124], integrase inhibitors [125], and reverse transcriptase inhibitors [126], but none of them are widely replicable.

1.4.3 HTLV-1 associated pulmonary disease (HAPD)

HTLV-1 associated pulmonary disease (HAPD) encompasses several inflammatory entities including pneumonias, bronchiolitis, alveolitis, bronchitis and especially Bex. The global burden of HAPD is currently unknown due to the lack of uniform clinical diagnostic criteria and being an under-recognised complication of HTLV-1 infection [127, 128]. Unlike ATL, cases of Bex occur in many individuals without infection with HTLV-1, complicating a clear causative connection for virus.

Many HAPD abnormalities can be detected by high resolution computed chromatography, such as centrilobular nodules, ground-glass opacities, interlobular septal thickening, and bronchial dilation [129, 130]. Like HAM, HAPD development arises from the infiltration of activated T lymphocytes into lung tissues causing chronic inflammation and ultimately, irreversible structural damage [128]. Similarly, HTLV-1 DNA has been detected in bronchoalveolar lavage fluid (BALF) [131], indicating that some of the infiltrating cells are infected and contribute to the ongoing chronic inflammatory loop [132]. This infiltration is likely facilitated by the upregulation of cell adhesion molecules by lung epithelial cells [133].

Therapeutic treatments for HAPD are extremely limited and mostly consist of immunomodulatory agents to combat the host inflammatory response, with low levels of success. Only individual case studies have reported improved symptoms with the use of corticosteroid Prednisolone for HTLV-1-associated bronchiolitis [134] and organising pneumonia [135].

1.4.4 HTLV-1 Uveitis (HU)

Uveitis, although not exclusively so, is another clinical entity of HTLV-1 infection (HU) [136-138], defined as a sight-threatening inflammatory disorder affecting the intraocular tissues [139]. HU is characterised by a sudden onset of symptoms such as eye floaters and blurred vision, that are brought on by clinical manifestations including iritis, vitreous opacity and retinal vasculitis [140].

While it is currently unknown how HTLV-1 infected cells cross the ocular-blood barrier, HU involves HTLV-1-infected T lymphocytes concentrating into the eye, causing severe local inflammation through dysregulated cytokine production [141, 142]. HU individuals have higher PVL in the eyes when compared to peripheral blood mononuclear cells (PBMCs) [142], and the level of inflammation is correlated with eye PVL [143].

Clinical management of uveitis can include topical, oral or sub-Tenon's injection of corticosteroid anti-inflammatory drugs, depending on the level of ocular inflammation [138]. In many cases with treatment, ocular inflammation is reduced and sight restored, however approximately 60% of HU patients experience recurrence [140].

1.4.5 Co-infections and other associated diseases

There are numerous co-infections associated with HTLV-1 infection. *Strongyloides stercoralis*, a parasitic roundworm, more commonly and severely infects HTLV-1 infected individuals [144]. HTLV-1 associated *S. stercoralis* infection can result in disseminated strongyloidiasis and hyper-infection syndrome, which results in 90% mortality [145]. Furthermore, HTLV-1 infected individuals are more likely to not only experience a blood stream infection (BSI) with enteric organisms, but to have repeated BSI episodes [17]. HTLV-1-associated infective dermatitis (HAID) is a chronic exudative eczema that is mostly prevalent in children infected with HTLV-1 and is caused by *Staphylococcus aureus* and *beta-hemolytic streptococci* infections [146]. Crusted scabies is also associated with HTLV-1 infection [12], most likely due to the viral-induced dysfunction of the immune system. In Central Australia, HBV/HTLV-1c coinfections are commonly reported, perhaps due to the similar transmission routes and endemicity of these viral infections [147].

Several non-communicable endocrine and metabolic diseases are also reported to be associated with HTLV-1 infection [148], including chronic kidney disease [149], type 1 diabetes mellitus [150], and autoimmune thyroid diseases [148].

Ultimately, the common threads that link the various disease manifestations associated with HTLV-1 infection are dysregulated immunity and inflammation, something that will be explored in more detail in later chapters of this thesis.

Table 1.2 – Summary of HTLV-1-associated diseases.

Disease	Categories	Clinical features	Viral pathogenesis	Current treatments
ATL	Chronic; Smouldering; Lymphoma; Acute[100]	Skin lesions; Hepato- and splenomegaly; Lymphadenopathy; Hypercalcemia[99]	Inhibition of apoptosis and DNA repair; Disruption of cell cycle; Malignant transformation[102]	ART[103-105]; HSC transplantation[107-109]; mAb therapy [116]
HAM	Demyelinating myelopathy	Leg weakness; Spasticity; Bladder & bowel dysfunction[118]	Infiltration of infected cells into CNS[119]; Inflammation in spinal cord[117]	Anti-inflammatory corticosteroids[121, 122]
HAPD	Bronchiectasis; Bronchiolitis; Bronchitis; Alveolitis; Pneumonias[127, 128]	Bronchial dilation; Lobular thickening; Persistent cough; Difficulty breathing [129, 130]	Infiltration of infected cells into lung tissues; Chronic inflammation[128]	Anti-inflammatory corticosteroids[134, 135]
HU	Iritis; Vitreous opacity; Retinal vasculitis[136-138]	Blurred vision; Eye floaters[140]	Infiltration of infected cells into eye; Local inflammation[141, 142]	Anti-inflammatory corticosteroids[138]
Co-inf.	BSI[17]; *S. stercoralis*[144]; HAID[146]; Crusted scabies[12]	Recurring infections; Difficulty clearing infections	Inefficient immune response to co-inf.	Antibiotic or insecticide, as required

ATL, Adult T cell leukemia; HAM, HTLV-1-associated myelopathy; HAPD, HTLV-1-associated pulmonary disease; HU, HTLV-1 uveitis; Co-inf; Co-infections; BSI, Blood stream infection; HAID, HTLV-1-associated infective dermatitis; CNS, central nervous system; ART, anti-retroviral therapy; HSC, hematopoietic stem cell; mAb, monoclonal antibody.

1.5 HTLV-1 gene expression

1.5.1 HTLV-1 genomic structure

HTLV-1 is a genetically complex retrovirus that includes an array of genes at the three prime (3′) end of the genome that regulate viral and cellular gene function. The infectious 9kb (+)ssRNA HTLV-1 genome encodes for the quintessential retroviral structural proteins (Gag, Pro, Pol, and Env), and the 3′ *pX* region encodes for the unique regulatory and accessory proteins (transcriptional activator of the *pX* region [Tax], Rex, p21, p30, p13, p12 and its cleaved product p8; **Figure 1.3**). HTLV-1 provirus can also transcribe a negative-sense RNA for the HTLV-1 basic leucine zipper factor (HBZ) using the 3′-LTR promoter (**Figure 1.3**). The HBZ protein and even non-protein coding *hbz* RNA exert strong pathogenic effects of HTLV-1 during ATL and HAM [151]. Hence, 5′-end-deleted and defective HTLV-1 proviruses that accumulate throughout infection [152] can still cause substantial adverse health outcomes [153, 154]. The viral subtype prevalent in Australia, Papua New Guinea, and the Melanesian Islands (HTLV-1c) is genetically divergent from subtype-A [155], most prominent in the *pX* region overlapping HBZ, p12/p8, and p30, where homology between the strains is below 85% (Hirons et al., unpublished), and will be discussed further. However, most of our understanding of the functions of these important regulatory proteins comes from investigations of the HTLV-1a strains endemic in Japan, central west Africa, Iran, and the Caribbean Islands.

The interplay among expressed HTLV-1 genes effects viral spread, persistence, and pathogenesis. Below is discussed the pleotropic roles of Tax, HBZ, Rex and p12/p8. The remaining HTLV-1 accessory and regulatory proteins play a role in viral transmission, persistence and latency, and have been reviewed extensively; p30 [156], and p13 [157, 158].

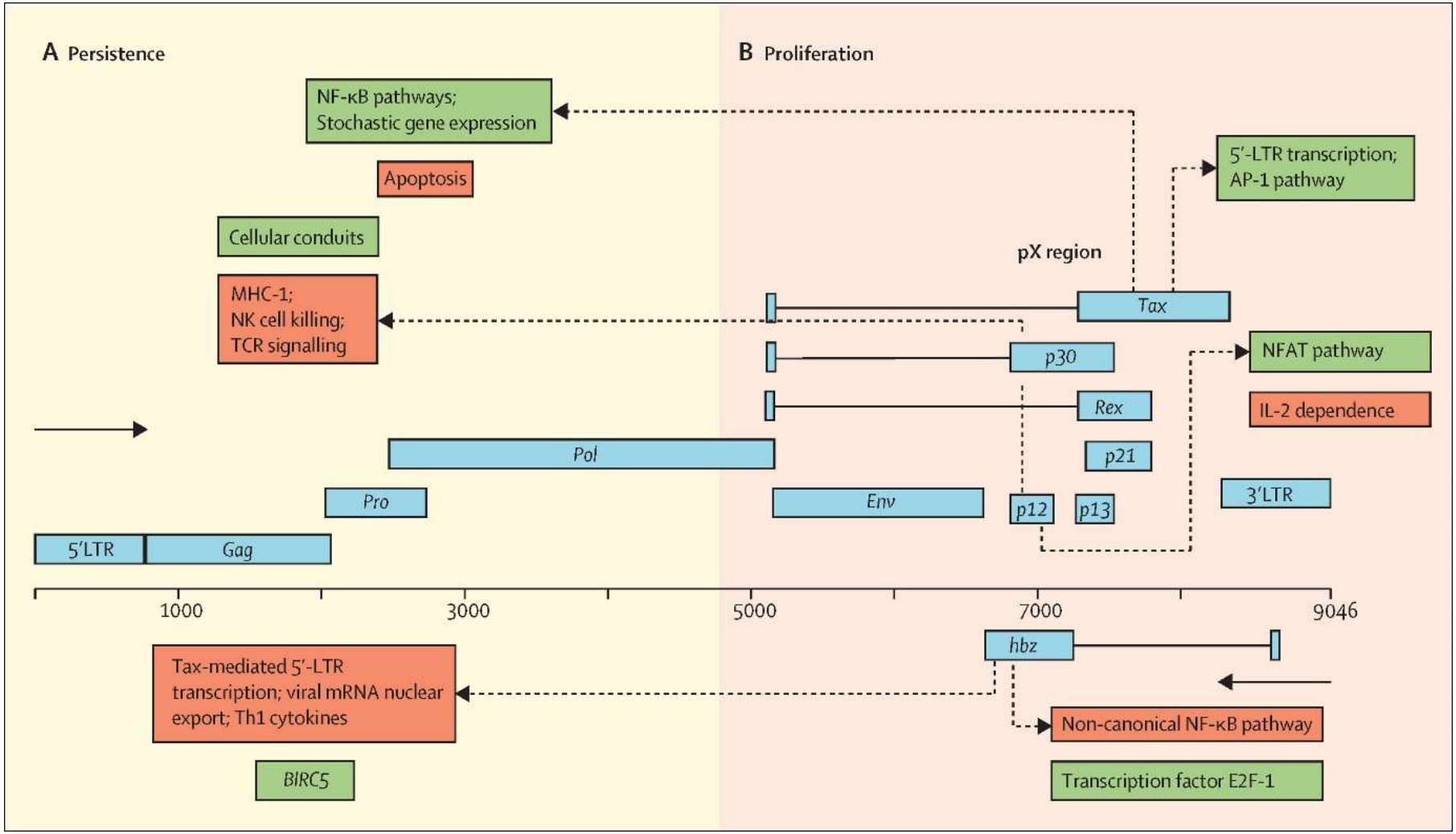

A Persistence
B Proliferation
NF-κB pathways;
Stochastic gene expression
Apoptosis
Cellular conduits
MHC-1;
NK cell killing;
TCR signalling
5′-LTR transcription;
AP-1 pathway
pX region
Tax
p30
Rex
NFAT pathway
IL-2 dependence
Pol
p21
Pro
Env
p12
p13
3′LTR
5′LTR
Gag
1000
3000
5000
7000
9046
hbz
Tax-mediated 5′-LTR
transcription; viral mRNA nuclear
export; Th1 cytokines
Non-canonical NF-κB pathway
BIRC5
Transcription factor E2F-1

Figure 1.3 – The genomic structure of HTLV-1 and *pX* proteins role in proliferation and persistence of infection. (A) Viral gene expression contributes to persistent HTLV-1 infection: Tax protein hyperactivates the NF-κB pathway leading to senescence and uses stochastic gene expression to avoid immune responses, while downregulating apoptosis. p12 protein downregulates MHC-1 expression on infected cell surfaces to prevent immune recognition and downregulates ICAM-1 and ICAM-2 expression to reduce NK cell killing. p8 protein increases cell-to-cell transmission through formation of cellular conduits, while downregulating proximal TCR signalling. HBZ protein downregulates viral transcription from the 5′-LTR, inhibits formation of viral particles by downregulating nuclear export of viral mRNAs, and decreases host immune response by downregulating the production of Th1 cytokines. *hbz* mRNA upregulates *survivin* (BIRC5) to inhibit apoptosis. (B) Viral gene expression contributes to the proliferation of HTLV-1-infected cells: Tax protein transactivates transcription from the 5'-LTR and upregulates the AP-1 pathway. p12 protein increases cytosolic calcium concentrations to activate the NFAT pathway and decreases infected cell dependence on IL-2 for proliferation. HBZ protein downregulates the non-canonical NF-κB pathway that allows cells to overcome senescence. *hbz* mRNA upregulates expression of host transcription factor E2F-1. Red-coloured boxes represent downregulation or reduction; green coloured boxes represent upregulation or activation. HBZ, HTLV-1 bZip factor. LTR, long terminal repeat. MHC-I, MHC class I molecules. NK cell, natural killer cell. NFAT, nuclear factor of activated T cells. Tax, transcriptional activator of the *pX* region. TCR, T cell receptor. Th1, type 1 helper T cell.

1.5.2 The effect of Transcriptional activator of the *pX* region (Tax) on viral expression and host immune responses

Tax, the viral oncogene-encoded protein, has a wide range of effects on HTLV-1 infection. Tax transactivates viral transcription from the 5′-LTR [159] by stabilising the binding of CREB, and CBP/p300, to Tax-responsive elements in the 5′-LTR untranslated unique 3' (U3) region [160-163] (**Figure 1.3B**). Tax can also activate many different cellular pathways, including nuclear factor kappa B (NF-κB) [164] (**Figure 1.3A**) and activator protein 1 (AP-1) pathways [165] (**Figure 1.3B**). The pleiotropic effects of Tax include enhancing T cell proliferation and cell mutagenesis, inhibiting DNA repair, disrupting the cell cycle, and downregulating stress-induced apoptosis [166] (**Figure 1.3A**).

Tax expression is high in early infection [167], but often suppressed in vivo or silent at later timepoints, because of the highly immunogenic nature of Tax [168]. Mechanisms used to inhibit Tax expression include clearance of infected cells by Tax-specific cytotoxic T lymphocytes (CTLs) [169-171], mutations in the *tax* gene [172], and

methylation [173] or deletion of the 5′-LTR [154]. Studies using single-molecule RNA fluorescence in-situ hybridisation show that Tax is expressed after intense intermittent bursts of the HTLV-1 positive sense in a small proportion of ex vivo cultures of T cell clones isolated by limiting dilution from peripheral blood of patients infected with HTLV-1 [174]. These positive sense bursts are triggered in vitro by cellular stress, such as reactive oxygen species and cisplatin-induced cytotoxic stress [175], hypoxia (1% oxygen), and glycolysis [176]. These bursts in Tax synthesis are inhibited by polycomb repressive complex 1 [177], and spontaneously switch between on and off states [174, 175]. The anti-apoptotic machinery activated by Tax expression might continue to mediate in vitro cell survival at single-cell level after Tax expression has been switched off [175]. This sporadic and minimal expression of Tax from clones of low abundance in vivo [178] enables the virus to maintain its reservoir and avoid the strong Tax CTL response [169, 171, 179, 180], which promotes persistent HTLV-1 infection (**Figure 1.3A**). However, it remains to be tested whether the transcriptional HTLV-1 positive sense bursts seen in vitro within T cell clones are preserved in vivo.

1.5.3 The role of HTLV-1 basic leucine zipper factor (HBZ) in survival and proliferation of infected cells

The viral *hbz* is the only gene encoded on the reverse strand [181], and is constitutively expressed in vivo throughout infection [34, 182], suggesting its importance in viral persistence. The retention of the *pX* region in deleted proviruses, and the persistent expression of *hbz* RNA, ensure that HTLV-1 is never truly silent during infection, even in asymptomatic individuals [154, 182]. Indeed, the accumulated concentration of *hbz* mRNA is positively correlated with HTLV-1 PVL and disease severity in infected individuals, including people who are ACs, and patients with ATL or HAM [183].

HBZ has a wide range of properties and functions that contribute to the pathogenesis and persistence of HTLV-1. HBZ opposes several functions of Tax, helping to balance pathogenesis with persistence. HBZ binds to CREB proteins [184], and thus downregulates Tax-mediated transcription at the 5′-LTR, by inhibiting the assembly of the transactivation complex (**Figure 1.3A**). Although Tax-mediated NF-κB hyperactivation leads to senescence, HBZ inhibits the non-canonical NF-κB pathway [185], allowing cells to overcome senescence and promoting proliferation of infected cells and persistent infection [186] (**Figure 1.3B**).

Several other HBZ-mediated strategies add to HTLV-1 persistence in the host. The HBZ protein has low immunogenicity, thought to be due to poor presentation of HBZ epitopes by human leukocyte antigen (HLA) class I molecules [187]. Moreover, HBZ inhibits the Rex-mediated nuclear export of viral structural mRNA transcripts, preventing production of infectious viral particles and maintaining viral persistence in the host [188] (**Figure 1.3A**). HBZ helps to impair cell-mediated immunity by downregulating transcription of type 1 helper T cell cytokines (IFN-γ, tumor necrosis factor alpha (TNF-α), and interleukin 2 (IL-2)) in HTLV-1-infected CD4+ T cells [189] (**Figure 1.3A**). HBZ also promotes HTLV-1-infected cells to convert to regulatory T cells, through activation of the transforming growth factor beta (TGF-β) signalling pathway and increased forkhead box p3 (*Foxp3)* transcription [190]. Therefore, HBZ contributes greatly to HTLV-1 evasion of immune control.

In addition to the HBZ protein functions, *hbz* mRNA transcripts can be retained in the nucleus, and are shown to have non-coding regulatory properties [168, 174]. *hbz* mRNA promotes T cell proliferation by upregulating the host transcription factor (TF) E2-F1 expression and its target genes [182] (**Figure 1.3B**). It also promotes cell survival, by upregulating the transcription of the cellular anti-apoptotic gene, *survivin* (*BIRC5)* [191] (**Figure 1.3A**). Thus, even deleted and non-protein coding HTLV-1 proviruses can have an effect on the function of infected host T cells, and this will be explored in subsequent chapters of this thesis.

1.5.4 The role of Rex in modulating viral mRNA expression and persistence

Rex, a regulatory protein which controls viral mRNA expression at the post-transcriptional level, is indispensable for viral spread and persistence in vivo [192, 193]. Rex binds to the Rex-responsive element, a specific stem-loop structure found in the 3' LTR region of viral intron-containing mRNA [194, 195]. The Rex-mRNA complex is then exported from the nucleus assisted by the cellular export factor chromosomal maintenance 1 (CRM1) [196, 197].

In addition to its role in viral mRNA export, Rex also preserves viral mRNA transcripts by interfering with the host mRNA surveillance pathway. Specifically, it suppresses the function of the cellular nonsense-mediated decay (NMD) machinery [193, 198], which

would otherwise recognise and degrade HTLV-1 mRNAs containing multiple alternate splice signals, ribosomal frameshifting signals, and multiple stop codons. In addition, this inhibition of NMD pathway by HTLV-1 may lead to the dysregulation of cellular mRNA expression and further facilitate viral pathogenesis, of which more investigation is needed [199]. Indeed, NMD dysregulation has been linked to the pathogenesis and progression of neurodegenerative diseases, such as motor neuron disease [200].

1.5.5 The role of p12/p8 in promoting T cell activation and proliferation and in avoiding immune recognition

Viral p12 is an important accessory protein in the maintenance of HTLV-1 infection, resides in the endoplasmic reticulum [201, 202], and has many functions [203]. It promotes T cell activation, viral transmission, and proliferation by increasing cytosolic calcium concentrations via release of calcium ions from the endoplasmic reticulum to the cytoplasm to activate the nuclear factor of activated T cells (NFAT) pathway [204, 205], and by binding IL-2 receptors to decrease dependence on IL-2 stimulation [206] (**Figure 1.3B**). By contrast with these functions, p12 also acts on viral pathogenesis and persistence by decreasing HTLV-1 recognition by the immune system. Viral p12 downregulates major histocompatibility complex class I molecule (MHC-I) cell surface expression, by binding newly synthesised MHC-I complexes and transporting them to the proteasome for degradation [207] (**Figure 1.3A**). Additionally, p12 promotes survival of infected CD4+ T cells by decreasing adherence to natural killer cells through downregulation of intracellular adhesion molecule 1 (ICAM-1) and ICAM-2 [208] (**Figure 1.3A**).

p8 is the cleaved product of p12 [209], and traffics to the infected cell surface following removal of the endoplasmic reticulum retention signal [210]. p8 increases viral cell-to-cell transmission by (1) increasing the number and length of cellular conduits; and (2) increasing clustering of lymphocyte function-associated antigen-1 on the cell surface, which helps transfer p8 to uninfected neighbouring cells, creating a T cell contact network [211, 212] (**Figure 1.3A**). p8 counteracts cell activation and proliferation modulated by p12, by inducing T cell anergy through downregulating proximal T cell receptor (TCR) signalling [213] (**Figure 1.3A**). p12 and p8 are both required for HTLV-1 infection in vivo [214], and their opposing and collective functions help to maintain the balance between host and HTLV-1 PVL.

1.6 The persistence and proliferation of HTLV-1 and associated pathogenesis

1.6.1 HTLV-1 cell-to-cell transmission modes avoid immune system surveillance

HTLV-1 infected cells produce few virions that are free of the cell, and HTLV-1 exists mostly as a stable provirus [44, 51]. This has proved to be an evolutionary asset for HTLV-1, which utilises a variety of other mechanisms to increase viral transmission and avoid immune recognition. HTLV-1 facilitates the polarisation of the cytoskeleton in infected cells and forms a tethered cluster of infectious virions in a virological synapse, which provides the viral genome and proteins with a transmission route into an uninfected target cell when there is cell-to-cell contact [215]. HTLV-1- infected cells can also form extracellular biofilm-like structures [216] composed of carbohydrates and linker proteins, or small membrane-bound structures known as extracellular vesicles (exosomes and microvesicles) [217], enabling both protection from the immune system and increased viral transmission. HTLV-1 can also promote the formation of conduits (i.e., filopodium-like growths that extend from host cells towards target cells) helping virions to transfer between cells [211]. Using various cell-to-cell transmission modes (**Figure 1.4A(1)**) helps HTLV-1 to infect new cells and persist, while simultaneously avoiding immune system detection.

1.6.2 The survival of HTLV-1 infected cells and non-infectious proviruses

More than 90% of HTLV-1 infected cells in vivo are memory CD4+ T cells [29, 218], allowing the HTLV-1 provirus to persist for years in this T cell reservoir (**Figure 1.4A(2)**). Studies show that patients with HAM have increased amounts of unique proviral integration sites correlating with higher PVL [219] and increased circulating HTLV-1-specific CTLs [179], suggesting that these initially non-productive infectious proviruses are reactivated and produce newly infected T cells, after years of quiescent infection. Furthermore, genomic studies suggest that patients with ATL have a high

frequency of low-abundance clones, suggesting that oligoclonal proliferation might not be the sole driver of malignant transformation [220].

The ability of HTLV-1 to infect HSCs in vivo [221] increases the HTLV-1 reservoir and contributes to the persistence of the virus (**Figure 1.4A(3)**). Identical integration sites of HTLV-1 were found in a range of cell types (CD4+ T cells, CD8+ T cells, monocytes, and B cells), showing their haematopoietic differentiation from HTLV-1-infected HSCs [221]. The HTLV-1 provirus strongly promotes HSC differentiation to T cell subsets that accumulate in breastmilk and semen, increasing the probability of mother-to-child or sexual HTLV-1 transmission through a cell-to-cell infection pathway [221]. Furthermore, both infected monocytes and DCs that differentiate from monocytes are highly efficient at seeding de novo infection in mature T cells [221]. Longitudinal studies of patients with HAM also showed identical integration sites in neutrophils [221], indicating the survival and persistence of these clones in vivo.

1.6.3 The interplay between the host genomic environment and the HTLV-1 provirus

The integration site of HTLV-1 can influence proviral expression levels and proliferation of infected cells. A provirus inserted into transcriptionally silenced DNA can lead to the establishment of long-lived T cell clones [219] (**Figure 1.4A(4)**). During chronic HTLV-1 infection, these latent proviruses survive, but those with active viral gene expression are targeted by the immune system [219].

The HTLV-1 provirus contains a non-palindromic CCCTC-binding factor (CTCF) binding site in the *pX* region of the genome [222]. CTCF serves as an RNA-transcription insulator and can bring together the provirus with distant upstream or downstream host loci up to 1.4 megabases apart, bringing host genes under the control of the HTLV-1 3′-LTR antisense promoter [222, 223] (**Figure 1.4B(1)**). Similarly, the 5′-LTR can be brought spatially close to a distal host gene and promote transcription, giving rise to potential activation of cell proliferation or survival pathways, or activation of oncogenes in ATL [224] (**Figure 1.4B(1)**). Although the frequency of this abnormal chromatin looping is unknown, there are approximately 27,000 CTCF binding sites in resting CD4+ T cells [225], and between 10,000 and 100,000 unique HTLV-1 integration sites in infected individuals [226], which might lead to widespread abnormalities in the host cell

chromatin structure and gene expression. The binding of CTCF to the HTLV-1 *pX* region blocks the enhancer activity when the provirus is inserted between an enhancer and a promoter region [222] (**Figure 1.4B(2)**). When the provirus is in cis, it has been shown to cause bidirectional transcription of the flanking host genome, and affect clone-specific transcription up to 300 kb away [223]. However, this CTCF-mediated chromatin looping also affects the HTLV-1 provirus, by downregulating post-transcriptional mRNA splicing, and physically blocking viral transcription [222].

Although the host genomic environment sometimes influences the persistence of HTLV-1, at other times it promotes the proliferation of infected cells. Higher proviral expression and clonal expansion are associated with integration sites in transcriptionally active areas of the host genome [219, 227, 228], probably owing to a combination of direct interactions with host promoters or enhancers, and open chromatin conformation allowing easier access of TFs to the provirus promoter [219] (**Figure 1.4C(1)**). The clonal expansion of the provirus is more likely if its orientation is the same sense as nearby host genes [219]. Furthermore, changes in the host epigenetic environment, such as histone modifications H3 Lys4me3, H3 Lys9ac, and H3 Lys27ac, are associated with increased transcription from the 5′-LTR [229]. However, the 5′ end of the proviral DNA is often methylated, inhibiting transcription of mRNA for the viral structural proteins targeted by adaptive immunity, but not at the 3′-LTR region that promotes persistent ongoing expression of functional hbz RNA. In many cases, transcripts from the antisense strand read through into the juxtaposed cellular genome, extending up to 50 kb downstream [230]. By contrast, the 5′-LTR promoter is hypomethylated when Tax is expressed [229].

1.6.4 The proliferation of HTLV-1 infected cells

HTLV-1-infected CD4+ T cells can undergo homoeostatic proliferation without apparent expression of virion protein, similar to the HIV-1 latent reservoir. Homoeostatic proliferation involves mitotic division, triggered by cytokine concentrations and cell cycle signalling, as opposed to an antigen-driven event (**Figure 1.4C(2)**). Proliferation of individual HTLV-1 infected T cells can be identified by expansions of clones with a common proviral integration site in the host genome, and the distribution and relative contribution of a clone to the total number of infected cells is known as the oligoclonality index (OCI) [219].

Patients infected with HTLV-1 have a T lymphocyte proliferation rate in vivo three times higher than people not infected [231]. This increase might result from continuous low viral gene expression, which activates lymphocyte proliferation of infected cells, necessary for maintaining the PVL, and also increasing proliferation of the responding immune cells. Although there are few HTLV-1 virions detected in vivo [44], ongoing de novo infection occurs at a low rate, producing between 100-200 new clones with a unique proviral integration site and T cell antigen receptor each day within a patient [232]. Longitudinal studies show that the HTLV-1 PVL remains steady over time in patients with HAM and in ACs, who typically have low PVL set points relative to patients with HAM [233]. Maintenance of a stable PVL with low rates of de novo infection might be balanced out by memory cell turnover, including productively infected cells dying from activation induced cell death (AICD), cytopathic effects, or lysis by CTLs.

Increased antigen-driven proliferation of infected cells is another mechanism that occurs with or without activation of virion protein expression. In individuals co-infected with HTLV-1 and HBV in Central Australia, HTLV-1 provirus OCIs in T cells were higher than in monoinfected individuals [147]. Importantly, the HTLV-1 provirus OCI in T cells was proportional to the concentrations of HBV surface antigen secreted by hepatocytes, indicating that the chronic antigenic stimulation during HBV infection is driving the clonal expansion of T cells infected with HTLV-1 [147] (**Figure 1.4C(3)**). Similarly, individuals co-infected with HTLV-1 and *S. sercoralis* showed a higher provirus OCI [234]. In environments where a large array of persistent antigen prevails, like Central Australia, the immunophenotype is likely to be shaped by HTLV-1 provirus positive T cells.

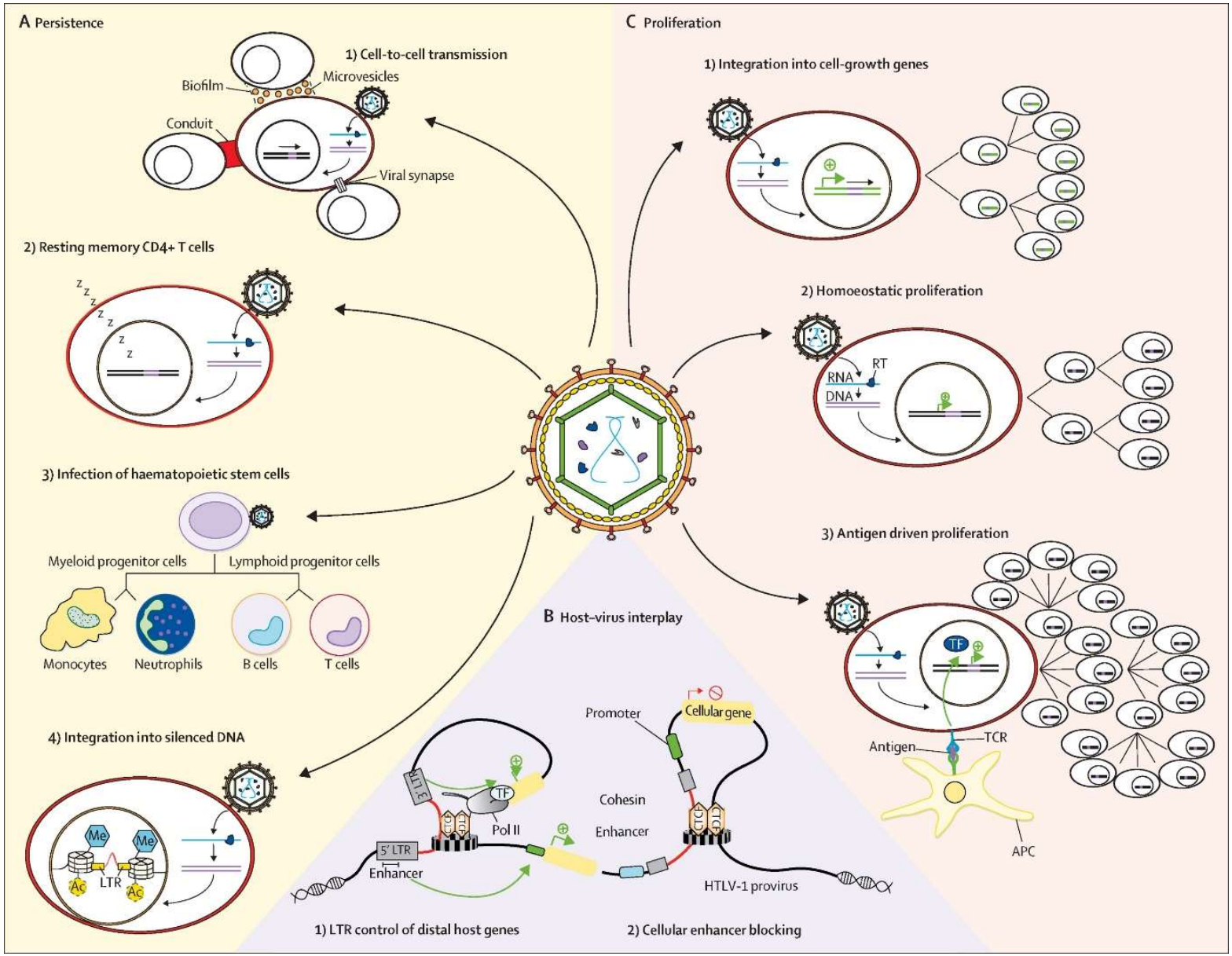

A Persistence
1) Cell-to-cell transmission
Biofilm
Microvesicles
Conduit
Viral synapse
2) Resting memory CD4+ T cells
3) Infection of haematopoietic stem cells
Myeloid progenitor cells
Lymphoid progenitor cells
Monocytes
Neutrophils
B cells
T cells
4) Integration into silenced DNA
Me
Ac
LTR
C Proliferation
1) Integration into cell-growth genes
2) Homoeostatic proliferation
RNA
RT
DNA
3) Antigen driven proliferation
TF
Antigen
TCR
APC
B Host-virus interplay
Promoter
Cellular gene
Cohesin
Enhancer
CTCF
HTLV-1 provirus
5' LTR
Enhancer
Pol II
1) LTR control of distal host genes
2) Cellular enhancer blocking

Figure 1.4 – The balance between HTLV-1 viral spread and viral gene expression leading to cellular proliferation and dysfunction. (A) Persistence of HTLV-1 infected cells. (1) Various cell-to-cell transmission modes are favoured over cell-free de novo infection through mechanisms including extracellular microvesicles or biofilm-like capture of infectious virion, formation of a conduit between cells, or use of the virological synapse formed during antigen presentation and T cell immune activation. These processes help provirus to escape detection and avoid clearance. (2) The integration of provirus directly into resting memory CD4+ T cells, or into activated cells infected with HTLV-1 that return to a resting state, promotes the long-term survival of provirus positive clones. (3) HTLV-1 infection of haematopoietic stem cells results in a wide distribution of HTLV-1-infected cells, but also drives a proliferative increase in favourable HTLV-1 proviruses integrated in responsive cell subsets. (4) The integration of provirus into DNA silenced for transcription by epigenetic modifications (eg, methylation or deacetylation) also promotes survival of long-lived clones. (B) The interplay between host and virus gene transcription through CTCF binding insulator sequence in the *pX* region of the provirus alters cellular chromatic architecture and gene expression to promote HTLV-1 proviral proliferation and persistence. (1) The 3′-LTR of HTLV-1, which has persistent activity for negative-sense RNA transcription, can drive ongoing expression of distant cell genes drawn into association through the CTCF interactions; alternatively, the 5'-LTR enhancer of HTLV-1 that is activated during elevated cell metabolism can act positively on flanking cellular genes brought into close proximity, and prompt increased transcription of the cellular gene. (2) Distal cellular gene expression can be affected when provirus is integrated between a cellular promoter and an enhancer region where the CTCF insulators alter higher order chromatin looping events to remap cellular DNA enhancer crosstalk away from cell control and in favour of viral 3'-LTR control. (C) Proliferation of HTLV-1 infected cells. (1) Integration of provirus into cell growth genes results in increased proliferation of infected cells. (2) Homoeostatic proliferation involves mitotic division of HTLV-1 infected cells triggered by cell cycle signalling, whereby two daughter cells contain the exact same proviral integration site as the parental cell. (3) Antigen-driven proliferation of infected cells is correlated with the OCI, and can be mediated by antigen-presenting cells (eg, macrophages or dendritic cells) and viral or bacterial antigens. Ac, acetylation; APC, antigen-presenting cell; CTCF, CCCTC-binding factor; LTR, long terminal repeat; Me, methylation; OCI, oligoclonality index; Pol II, DNA polymerase II; TCR, T cell receptor; TF, transcription factor.

1.7 The immune response to HTLV-1 infection

The effects of HTLV-1 infection on the immune system are complex and multifaceted. Whereas HIV-1 causes immunodeficiency by killing infected CD4+ T cells, HTLV-1 provirus alters the function of infected immune cells over several decades in the infected individual, leading to direct and immune-associated diseases. Some infected individuals develop serious or life-threatening diseases, while others may be more susceptible to co-infections, or have generally poorer health outcomes. While specific host genetics may play a role, it is not fully understood what factors determine the outcome of HTLV-1 infection. Nonetheless, it is clear that HTLV-1 significantly impacts the host immune system by impairing T cell responses and disrupting the cytokine network.

1.7.1 HTLV-1 infection impairs host T cell responses

Overall, HTLV-1 infection leads to changes in normal cell functioning, signalling pathways and transcriptional activity. The viral protein HBZ impairs the host immune response by upregulating the expression of T cell immunoreceptor with immunoglobulin and ITIM domain (TIGIT), which decreases the effectiveness of T cell responses to HTLV-1 infection [235]. The dysregulation of *Foxp3* expression and the functionality of regulatory T cells (Tregs) induced by HTLV-1 infection impacts CD8+ and CD4+ T cell responses [236, 237]. Moreover, persistent viral expression from the lifelong infection elicits a chronically activated CTL response [238], which may render it ineffective at controlling the infection and limiting damage to the host. T cell responses, particularly in the context of HTLV-1c infection and associated pulmonary disease, will be further explored in chapter three of this thesis.

1.7.2 Cytokine network is disrupted in HTLV-1 infection

HTLV-1 infection causes the dysregulation of host cytokine responses, which can contribute to the development of HTLV-1-associated diseases [239]. High levels of anti-inflammatory cytokines such as IL-10 [240, 241] and TGF-β [242] are associated with the development of ATL. IL-10 promotes cell proliferation and survival [243], and importantly for HTLV-1 infection, can also lead to CD8+ T cell exhaustion [239]. Malignant clones that produce high levels of IL-10 and TGF-β, along with the accumulation of Tregs, are also thought to trigger immunosuppression [237, 242, 244]. Contrastingly, HAM patients have a characteristic inflammatory cytokine signature

dominated by high levels of plasma IFN-γ that facilitates infiltration of infected cells into the CNS [245]. This positive feedback loop of pro-inflammatory cytokines can result in ongoing chronic inflammation in the CNS and leads to demyelination events. The host cytokine response will be further explored in chapter three of this thesis, particularly in the context of subtype-C infection and HAPD.

1.7.3 The impact of host genetics on the progression of HTLV-1 infection

Host genetics can influence the way an individuals' immune system responds to viral infections, including HTLV-1. For a lifelong infection like HTLV-1, host genetics may play a role in the long-term outcome. HLA class I genotype determines the efficacy and specificity of the CTL response, which in turn influences the HTLV-1 PVL and the risk of developing an HTLV-1 associated disease. Studies have investigated the potential protective or susceptible effects of HLA class I type on the development of HTLV-1-associated diseases in different HTLV-1 endemic regions. For example, HLA-A*02 class I allele was observed to have a protective effect in southern Japan [246], Brazil [247], and the UK (ref), however these results were not replicated in similar studies conducted in Peru [248] or the Caribbean [249]. In contrast, a study in Spain determined that HLA-DRB1*0101 and HLA-B*07 class I alleles were associated with a susceptibility to HAM development and higher PVL, possibly due to an inefficiency of the CTL response [250].

Single nucleotide polymorphisms (SNPs) in the host genome may also be determinants of susceptibility or protection against viral infections, especially if they occur in gene coding regions or nearby regulatory regions such as promoters. Some SNPs have been elucidated in HTLV-1a infection that can influence an individual's PVL and risk of HTLV-1 associated disease. A SNP in the IL-10 gene promoter (IL10-592A) was associated with a reduced risk of HAM in a study conducted in Japan [251]. Similarly, a SNP in the 3'untranslated region (UTR) of the stromal cell-derived factor 1 gene (SDF1+801A) confers reduced risk for HAM, as does polymorphism 191C of IL-15 gene [252]. On the other hand, a SNP in IL-6 promoter (IL6-634C) confers susceptibility of HTLV-1 infected individuals to develop HAM in Brazil [253]. Additionally, a promoter polymorphism in TNF (TNF-963A) was associated with HAM in Japan, particularly in individuals with high PVL [252].

1.8 HTLV-1c in Australia

1.8.1 Central Australia is endemic for HTLV-1c infection

HTLV-1c is endemic in Central Australia, with some of the highest prevalence rates in the world. In some remote communities, the prevalence of HTLV-1c infection can reach up to approximately 60% in the adult population [254]. While there have been rare cases of ATL and HAM in this region, the most common associated disease with HTLV-1c infection is HAPD, specifically bronchiectasis [9, 10, 15, 17, 128]. There are also many occurrences of BSI in HTLV-1c infected individuals in Central Australia [17, 255]. Furthermore, those infected with HTLV-1c in Central Australia have a higher predicted mortality rate than their non-infected counterparts [256], consistent with the worldwide meta-analysis showing higher mortality with HTLV-1 [14]. Moreover, HTLV-1c infected individuals report higher hospital admissions and health service utilisation [17].

1.8.2 HTLV-1c is genetically divergent from HTLV-1a

HTLV-1c is the most genetically divergent subtype, most pronounced in the *pX* region coding for regulatory and accessory proteins [155], and these nucleotide mutations are consistent across different donors (Hirons et al., unpublished). Notably, HTLV-1c has a conserved mutation in the start codon of *p12* (AUG [methionine] to ACG [threonine] at position 6840) [155, 203] (Hirons et al., unpublished). Studies found that the *p12* start codon is required for HTLV-1 infection in bone marrow-liver-thymus mouse models [257], human DCs, and macaques, but not rabbits [258, 259]. Furthermore, the coding region of HTLV-1c *p12* gene contains 21 amino acid substitutions, some of which remain in *p8* [203] (Hirons et al., unpublished). Although the full implications of these mutations are not clear, they might contribute to the development of Bex, the uniquely frequent disease association of HTLV-1c in Australia [9].

1.9 Project overview

HTLV-1c infection in Australia is associated with high mortality rates and a range of serious associated diseases, yet there is still much we do not understand about this virus. Hence, there is an urgent need to develop strategies to improve the outcomes for those already infected and prevent future transmission events.

By gaining a deeper understanding of HTLV-1c virology and pathogenesis, as well as the host immune response to this lifelong infection, we hope to pave the way for antiviral drugs, immune-activating vaccines, and passive immune therapies that could prevent new infections, and provide transformative therapeutic outcomes akin to those that have revolutionized the treatment of HIV-1, HBV and HCV.

1.9.1 Project aims

For fulfilment of these objectives, this research project had three overarching aims:

1. To characterise the T cell responses and the interwoven cytokine network during HTLV-1c infection using primary blood samples of HTLV-1c infected First Nations people from Central Australia
2. To characterise the phenotype of HTLV-1c infected T cells and determine the nature of the proviral landscape using primary blood samples and high-throughput sequencing
3. To analyse the first humanised HTLV-1c mouse model for its ability to recapitulate in vivo human infection by utilising phenotyping and sequencing strategies similarly applied to human infection characterisation

2. Materials and Methods

2.1 Materials

2.1.1 Chemicals

All chemicals used in this study, unless otherwise specified in the text, were of analytical grade or higher, from Sigma-Aldrich (St. Louis, Missouri, USA).

2.1.2 Oligonucleotides

All oligonucleotides used in this study were synthesized by Sigma-Aldrich. Stocks were received desalt-purified in H_2O at 100µM. Working solutions were prepared in molecular grade H_2O (Sigma-Aldrich) at 20µM, 10µM and 5 µM as required. A complete list of oligonucleotides is detailed in **Table 2.1**.

2.1.3 Probes

All fluorescent probes used in this study were TaqMan® hydrolysis probes synthesized by ThermoFisher Scientific (Massachusetts, USA). Stocks were received in H_2O at 100µM. Working solutions were prepared in molecular grade H_2O (Sigma-Aldrich) at 20µM. A complete list of probes used are detailed in **Table 2.2**.

2.1.4 Antibodies

All antibodies in this study were used for flow cytometry and fluorescence activated cell sorting (FACS) and are detailed in **Table 2.3**, with their respective suppliers, clones, fluorophores, isotypes and working dilution.

2.1.5 Antibody-Immobilized Magnetic beads

All antibody-immobilised magnetic beads in this study were used for a human cytokine/chemokine plasma assay on the Luminex platform and are detailed in **Table 2.4**.

2.1.6 Peptide libraries

Overlapping peptide libraries spanning full-length Tax and Env proteins were designed such that each peptide was 15 amino acids in length, overlapping the previous and subsequent peptides by 11 amino acids. Peptide libraries (PEPotex SRM) were synthesized by ThermoFisher Scientific with mass spectrometry quality control analysis. Individual peptides were reconstituted in dimethylsulfoxide (DMSO) (Merck KGaA,

Darmstadt, Germany) or 0.1% trifluoroacetic acid (TFA) (ThermoFisher Scientific) in 50% (v/v) acetonitrile (ACN)/H_2O and stored at -80°C. To create the libraries, Tax peptides were pooled to a final concentration of 0.3µg/µl/peptide and Env peptides to 0.05µg/µl/peptide and stored at -20°C for use.

Table 2.1 – Oligonucleotide primers used in this thesis.

Name	Strand	Sequence (5'-3')	Target	Amplicon size (bp)
ODP 3083	+	CAAATGAAGGACCTACAGGC	HTLV-1c *gag* region	184
ODP 3084	-	TATCTAGCTGCTGGTGATGG	HTLV-1c *gag* region	
ODP 3085	+	TCCAGGCCTTATTTGGACAT	HTLV-1c *tax* region	155
ODP 3086	-	CGTGTGAGAGTAGGACTGAG	HTLV-1c *tax* region	
ODP 3847	+	CACGTCTGACTGCCGGCTTG	Full-length HTLV-1c provirus (1st round)	8275
ODP 3848	-	CGGAACTTTCGATCGGTAGC	Full-length HTLV-1c provirus (1st round)	
ODP 3849	+	AAGCACCGGCACCCTTACT	Full-length HTLV-1c provirus (Nested round)	8221
ODP 3850	-	CGCAGAACAGAAAACGAAA	Full-length HTLV-1c provirus (Nested round)	
ODP 3986	+	CCTCTTTTTCCCGCTCTCTT	HTLV-1c *hbz* region	170
ODP 3987	-	GAGGGAGGAGGAGGAATCTG	HTLV-1c *hbz* region	
ODP 4033	+	TACCATGCCACCTATTCCCT	HTLV-1c *env* region	221
ODP 4034	-	TTCAGGCGTGAGACTTCTTG	HTLV-1c *env* region	

Table 2.2 – Fluorescent TaqMan® hydrolysis probes used in this thesis.

Name	Sequence (5'-3')	Target
ODP 3318	6FAM-CATGATTTCCGGGCCTTGC-MGBNFQ	HTLV-1c *tax* region
ODP 3321	6FAM-ACCATCCGGCTTGCAGT-MGBNFQ	HTLV-1c *gag* region
ODP 3988	6FAM-TTCGCTTTCTCTTCTCCTCG-MGBNFQ	HTLV-1c *hbz* region
ODP 4035	6FAM-ACAGGGGCCGTCTCCAGCCCC-MGBNFQ	HTLV-1c *env* region

6FAM, 5' fluorescent dye; MGBNFQ, 3' minor groove binding non-fluorescent quencher

Table 2.3 – Flow cytometry antibodies used in this thesis.

Antibody	Supplier	Clone	Fluorophore	Isotype	Dilution
Activation Induced Marker (AIM) assay					
Anti-human CD3	BD Biosciences	SK7	BUV395	Mouse BALB/c IgG_1, κ	1:300
Anti-human CD4	Biolegend	RPA-T4	BV605	Mouse IgG_1, κ	1:300
Anti-human CD8a	Biolegend	RPA-T8	BV650	Mouse IgG_1, κ	1:1500
Anti-human CD20	BD Biosciences	2H7	BV510	Mouse C57BL/6 IgG_{2b}, κ	1:300
Anti-human CD45RA	BD Biosciences	HI100	BUV737	Mouse IgG_{2b}, κ	1:600
Anti-human CD69	Biolegend	FN50	FITC	Mouse IgG_1, κ	1:600
Anti-human CD127	BD Biosciences	HIL-7R-M21	APC-R700	Mouse IgG_1, κ	1:150
Anti-human CD137	Biolegend	4B4-1	BV421	Mouse IgG_1, κ	1:300
Anti-human CD154	BD Biosciences	TRAP1	APC-Cy™7	Mouse BALB/c IgG_1, κ	1:300
Anti-human CCR6	Biolegend	G034E3	BV785	Mouse IgG_{2b}, κ	1:300
Anti-human CXCR3	Biolegend	G025H7	PE-Dazzle™594	Mouse IgG_1, κ	1:150
Anti-human CXCR5	eBioscience	MU5UBEE	PE	Mouse IgG_{2b}, κ	1:150
Anti-human OX-40	Biolegend	ACT35	PerCP-Cy5.5	Mouse IgG_1, κ	1:150
Live/dead stain	Invitrogen	N/A	Aqua	N/A	1:11000
CD4 phenotyping assay					
Anti-human CD3	BD Biosciences	SK7	PerCP-Cy5.5	Mouse BALB/c IgG_1, κ	1:15
Anti-human CD4	BD Biosciences	RPA-T4	Alexa Fluor® R700	Mouse IgG_1, κ	1:60
Anti-human CD38	BD Biosciences	HB7	PE-Cy™7	Mouse IgG_1, κ	1:60
Anti-human CD45RA	BD Biosciences	HI100	PE-CF594	Mouse IgG_{2b}, κ	1:60
Anti-human CD49d	Biolegend	9F10	BV510	Mouse IgG_1, κ	1:60
Anti-human CD127	BD Biosciences	HIL-7R-M21	BV786	Mouse IgG_1, κ	1:60
Anti-human CCR4	R&D Systems	205410	PE	Mouse IgG_{2b}, κ	1:60
Anti-human HLA-DR	BD Biosciences	G46-6	BV711	Mouse IgG_{2a}, κ	1:60
Anti-human Integrin-β7	BD Biosciences	FIB504	BV605	Rat F344 CDF IgG_{2a}, κ	1:60
Live/dead stain	Invitrogen	N/A	NIR	N/A	1:3500

Table 2.4 – Human cytokine/chemokine magnetic beads used in this thesis.

Bead Analyte	Luminex® Magnetic Bead region	Catalog number
Anti-human IL-2	48	HIL2-MAG
Anti-human Il-4	53	HIL4-MAG
Anti-human IL-6	57	HCYIL6-MAG
Anti-human IL-8	63	HCYIL8-MAG
Anti-human IL-10	27	HCYIL10-MAG
Anti-human IP-10	65	HIP10-MAG
Anti-human MCP-1	67	HCYMCP1-MAG
Anti-human IFN-γ	25	HCYIFNG-MAG
Anti-human TNF-α	75	HCYTNFA-MAG

2.2 Primary patient blood samples

2.2.1 Patient sample collection

Patients were recruited at Alice Springs Hospital, Mparntwe, Northern Territory, as part of an HTLV-1c cohort, with fully informed consent in primary language and full written consent, in accordance with the National Health and Medical Research Council (NHRMC) and Central Australian Human Research Ethics Committee (CAHREC). This study is approved by CAHREC under the ethics reference HREC-17-2930, and The University of Melbourne Human Research Ethics Committee (1442830.1). Other negative control whole blood samples were obtained from the Lifeblood Red Cross, Naarm, Melbourne, under the research agreement 17-08VIC-01.

2.2.2 PBMC and plasma isolation

PBMCs and plasma were isolated from whole blood buffy coats using Ficoll-Paque™ (GE Healthcare, Uppsala, Sweden) density centrifugation method, as per the manufacturer's instructions. Whole blood + PBS 1X mixture was added to Ficoll-Paque™ Plus (GE Healthcare) and centrifuged to separate into phases. Plasma was transferred to fresh tubes and frozen at -80°C, while PBMCs were washed three times with PBS 1X + 2mM EDTA (BioRad) pH 8.0. When required, RBC lysis was performed on PBMCs using ice cold 0.83% ammonium chloride (VWR International, Radnor, Pennsylvania, USA), and washed again with PBS 1X + 2mM EDTA (BioRad). PBMCs

were resuspended in freezing media (10% DMSO (Merck KGaA) in heat-inactivated fetal bovine serum (HI-FBS) (Bovogen Biologicals, Melbourne, Australia)) and frozen at -80°C in an isopropanol-filled (ChemSupply, Gillman, Australia) freezing container. Cells were then transferred to liquid nitrogen the following day for cryopreservation until use.

2.2.3 Diagnosis of HTLV-1 serostatus

HTLV-1c serology screening was performed by the National Reference Laboratory (NRL) in Melbourne, Australia. The presence of HTLV-1 specific antibodies was determined using the Murex HTLVI+II enzyme immunoassay (EIA) (DiaSorin, Dartford, UK) and Serodia HTLV-1 particle agglutination assay (Fujirebio, Tokyo, Japan), in accordance with manufacturer's criteria. Samples reactive on these assays were then subject to confirmation Western Blot using HTLV-I/II Blot 2.4 (MP Diagnostics, Singapore), according to manufacturer's instructions.

2.3 Genomic DNA (gDNA)

2.3.1 gDNA extraction

Genomic DNA (gDNA) extraction was carried out with the GenElute™ blood genomic DNA kit (Sigma-Aldrich) according to the manufacturer's recommendations for sorted CD4+ T cell phenotypes, bulk human PBMCs and hu-NSG mice splenocytes. Cells were treated with proteinase K (Sigma-Aldrich) and RNase A (Sigma-Aldrich), prior to cell lysis with Lysis Solution C (Sigma-Aldrich). Cells were incubated at 55°C for one hour with periodic mixing, to ensure complete lysis. For sorted cell populations with low cell recovery, UltraPure™ Herring Sperm DNA (Invitrogen) was added to the solution as carrier DNA to increase total DNA yields, prior to column binding by centrifugation. Two wash steps were carried out, followed by the addition of Elution Solution (Sigma-Aldrich) and 5 minutes of incubation at room temperature to increase subsequent elution efficiency. Eluted gDNA was centrifuged and the supernatant was transferred to a fresh tube, to remove any contaminating silica fines from the column which would inhibit downstream processes. Final gDNA products were stored at -20°C until further use.

2.3.2 gDNA ethanol precipitation

When required, gDNA was concentrated by ethanol precipitation. Firstly, sodium acetate, 100% ethanol (POCD Scientific, Sydney, Australia) and Pellet Paint NF Co-precipitant (Millipore) were added to gDNA solution and incubated at -20°C for 2 hours. gDNA was then pelleted by centrifugation and washed with 70% ethanol (POCD Scientific) twice, before being resuspended in nuclease-free H_20 (Sigma-Aldrich) and stored at -20°C.

2.3.3 gDNA quantification

gDNA concentration and A260/280 ratio was determined using the Nanodrop™ 2000 Spectrophotometer (Thermo Fisher Scientific), according to the manufacturer's instructions. Background absorbance for the blank control was measured using 1µl of Elution Solution (Sigma-Aldrich) or molecular grade H_20 (Sigma-Aldrich) as required, and then 1µl of each gDNA sample for quantification.

2.4 Droplet digital (dd) PCR

2.4.1 ddPCR primer and probe design

Droplet digital PCR (ddPCR) assay [260], including protocols, primers and probes, were designed using selection criteria detailed by BioRad, for HTLV-1c in *gag*, *env*, *hbz* and *tax* gene regions. Optimal primer design included 50-60% GC content, T_m between 50-65°C, amplification length of 60-200bp and avoiding secondary structure, 3' complementarity, and repeats of more than 3 Gs or Cs. Probes, as required by the QX200 system, were TaqMan™ hydrolysis probes. Optimal probe design included a T_m 3-10°C higher than primers, annealing to strand with higher GC content, length of <30 nucleotides, non-fluorescent quencher, and avoiding a G at the 5'end. Finally, all primers and probes were blasted to ensure no non-specific binding or alignments with other sequences of human, mouse, or herring fish genomes.

The designed HTLV-1c *hbz* and *env* gene region primers (**Table 2.1)** were validated using PCR with positive control gDNA. Amplification of 25ng gDNA was prepared with 1.25U of GoTaq® Hot Start Polymerase (Promega, Madison, Wisconsin, USA) in a 50µl

reaction mixture with final concentrations of 1X Green GoTaq® Flexi Buffer (Promega), 1.5mM $MgCl_2$ (Promega), 0.2mM dNTPs (Promega), 0.1µM forward and reverse primers. Mastercycler Nexus (Eppendorf, Hamburg, Germany) thermocycler was used for the following PCR protocol: enzymatic activation for 2 minutes at 95°C, then 30 cycles of (denaturation for 30 seconds at 95°C, annealing for 1 minute at 58°C, extension for 1 minute at 72°C), followed by a final extension of 5 minutes at 72°C. PCR products (5µl) were electrophoresed in a 2% agarose gel with GelRed 1X (Biotium, San Francisco, California, USA) in TAE buffer 1X at 100V. Bands were visualised under UV light and correct amplicon sizes validated. *Gag* and *tax* primers were previously validated for this assay [218].

2.4.2 ddPCR droplet generation and detection

In addition to HTLV-1c+ samples, each ddPCR run contained non-template controls (NTC), negative controls containing gDNA from an uninfected human or mouse donor, and positive controls containing gDNA from a confirmed HTLV-1c infected human or mouse donor, or HTLV-1c infected Jurkat cell line (Ellenberg et al., unpublished). Each sample was run in duplicate or triplicate.

ddPCR reaction mixture was prepared for gDNA (50-100ng), combining HTLV-1c forward (900nM) and reverse (900nM) primers (**Table 2.1**), HTLV-1c-specific probe (250nm) (**Table 2.2**), 2X *RPP30* CNV assay mix (BioRad) containing forward and reverse primers and probe, and 2X Supermix for probes (no dUTP) (BioRad). PCR preparation and droplet generator oil for probes (BioRad) were emulsified creating 20,000 droplets per well through the QX-200 droplet generator. PCR emulsions were transferred to a 96-well plate (Eppendorf) thermo-sealed with foil sheet (BioRad). PCR was performed with C1000 Touch™ thermocycler (BioRad) under the following conditions: 95°C for 10 mins, 40 cycles of (94°C for 30 sec, 58°C for 1 min) followed by 98°C for 10 min, with infinite hold at 10°C. Cycles had a ramp rate of 2°C/sec, lid heat was 105°C and sample volume was set to 46ul. Droplets were then analysed for two-colour detection of FAM and HEX using the QX200 droplet reader (BioRad).

2.4.3 ddPCR raw data analysis

QuantaSoft v1.7.4 (BioRad) was used to acquire and analyse the fluorescence of the droplets in the FAM and HEX channels. Manual negative thresholds were set for the

fluorescence amplitude of each primer/probe target, based on the respective NTC and HTLV-1c negative control wells. As a previously defined quality control threshold [218], NTC and negative control wells containing two or more positive droplets resulted in the assay being repeated. The number of copies of the HTLV-1c target gene and the reference gene *RPP30*, per µl of reaction was determined, and therefore the PVL, which is the number of copies of HTLV-1c gene per 10^6 cells, could be calculated.

2.5 Activation Induced Marker (AIM) Assay

2.5.1 AIM assay design

Activation induced marker (AIM) assay was designed to assess the HTLV-1c-specific T cell responses [261, 262]. In particular, the FACS panel incorporated analysis of several antigen-specific CD4+ T cell phenotypes: cTFH, Tmem, Th1, Th17, Th1Th17 cells; and CD8+ T cells.

2.5.2 HTLV-1c peptide stimulation

Cryopreserved PBMCs were thawed and washed twice with RPMI 1640 media (Gibco) + 10% HI-FBS (Bovogen Biologicals) (RF10), before being seeded into a 96-well plate in RF10 media with antibiotic-antimycotic 1X (Gibco) at 1.0 x 10^6 cells per well. Four wells were required for each donor: cells were stimulated with a final concentration of either DMSO (0.4%) (Merck KGaA), Staphylococcus Enterotoxin B (SEB) (1.25ng/ml) (Sigma-Aldrich), Tax peptide pool (1µg/ml/peptide) or Env peptide pool (1µg/ml/peptide), along with co-stimulatory reagent FastImmune Costim CD28/CD49d (BD Biosciences) in each condition. DMSO served as the negative control measuring the background levels of activation marker expression, SEB served as the positive control to ensure cells could respond to stimulation, and Tax and Env conditions served to measure the frequency of HTLV-1c specific T cells. Lastly, each condition was stained with CD154 APC-Cy7 (TRAP1) (BD Biosciences) and incubated for 20 hours at 37°C.

2.5.3 AIM staining and flow cytometry analysis

Following harvest, cells were washed with PBS 1X and then stained with Live/Dead Aqua (Invitrogen) and an incubation of 3 minutes at room temperature. Next, cells were stained with a cocktail of surface antibodies in PBS 1X + 2.5% HI-FBS (FACS wash) detailed in **Table 2.3**: CD20 BV510 (H7), CD3 BUV395 (SK7), CD45RA BUV737 (HI100), and CD127 APC-R700 (HIL-7R-M21); and incubated at 4°C for 30 minutes. PBMCs were then washed with FACS wash and fixed with BD cytofix (BD, New Jersey, USA). Single colour controls were prepared using OneComp eBeads™ (Invitrogen), and live/dead compensation controls were prepared using ArC™ Amine Reactive Compensation beads (Invitrogen). Cell suspensions were filtered through a 70µm sieve and acquired on a BD LSR Fortessa using FACSDiva™ v9.0 software (BD Biosciences). Fluorescence data was analysed in FlowJo v10 software (FlowJo, Oregon, USA).

2.6 Multiplex bead-based cytokine assay

Frozen plasma samples were thawed and incubated neat with Milliplex® Map Human Cytokine/Chemokine Magnetic Bead Panel and assay buffer (HCYTOMAG-60K) (EMD Millipore, Massachusetts, USA) with agitation on a plate shaker, overnight at 4°C. Human cytokine/chemokine standards (EMD Millipore) and human cytokine quality controls (EMD Millipore) were prepared according to the manufacturer's instructions. The following 9 bead analytes were combined using the customisable kit: (IL-2, IL-4, IL-6, IL-8, IL-10, IP-10, TNF-α, INF-γ, MCP-1) and are detailed in **Table 2.4**. Magnetic washing was carried out twice with wash buffer (EMD Millipore), before adding detection antibodies (EMD Millipore) to each sample and incubating for 1 hour at room temperature, with agitation on the plate shaker. Streptavidin-phycoerythrin (EMD Millipore) was added to each sample and incubated with agitation on plate shaker for 30 minutes at room temperature. Two final magnetic washing steps were carried out with wash buffer (EMD Millipore), before adding sheath fluid (EMD Millipore) to all wells. Samples were acquired on a Luminex FlexMap 3D machine (Luminex, Texas, USA) with xPONENT software.

2.7 Statistical Analysis and Modelling

2.7.1 Uni- and bivariate analyses

All calculations of means, medians, IQRs and univariate and bi-variate statistical analyses were performed using Graphpad PRISM software V9.0 (San Diego, California, USA). When required, data was $\log_{10}(x)$ transformed, or right-shifted $\log_{10}(x+1)$ transformed. Non-parametric statistical hypothesis testing between two groups was performed by the Mann-Whintey U test [263]. Non-parametric statistical hypothesis testing between three or more groups was performed using the Kruskal-Wallis test [264]. Non-parametrical statistical hypothesis testing between three or more groups with matched measurements was performed by using the Friedman test. Parametric statistical hypothesis testing between three groups used ordinary one-way ANOVA [265]. Parametric statistical hypothesis testing between three groups with matched measurements was performed by using the RM one-way ANOVA. When required by these tests, false discovery rate (FDR) control for multiple comparisons used two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli [266]. Correlation testing was performed using Spearman test for non-parametric data and Pearson test for parametric data.

2.7.2 Overview of multivariate analyses

All analyses used Python v3.9.6 (Python Software Foundation, Beaverton, Oregon, USA), distributed by Anaconda Inc. (Austin, Texas, USA). Data processing was performed using numpy v1.20 and pandas v1.4. Data analysis and modelling was performed using scikit-learn v1.0 [267] and statsmodels v0.12. Data visualization was performed using matplotlib v3.4 and seaborn v0.11.

All inputs were first $\log_{10}(x+1)$ transformed using pandas and numpy to reduce the impact of outliers and normalise the data, without removing any patient data. All inputs were then standardised via mean-centring and standard deviation-scaling prior to subsequent analyses, except for logistic regression where standardising was not required.

2.7.3 Correlation heatmap

Pairwise Pearson correlations were calculated using pandas v1.4 and subsequently visualized as a heatmap using seaborn v0.11.

2.7.4 Principal Component Analysis (PCA)

Principal component analysis (PCA) [268, 269] was performed using scikit-learn v1.0, generating a projection matrix with a loading for each input variable to each principal component (PC). The linear combination of all loadings and their respective inputs for a given PC and patient produced a PC score for that patient. Incremental and cumulative variance explained for each PC were calculated in scikit-learn v1.0 using the associated eigenvalues of the projection matrix. The first two PC scores for all patients were visualised using matplotlib v3.4 and seaborn v0.11, with colours used for disease state. Ninety-five percent confidence regions for disease states were created using matplotlib v3.4 and numpy v.120 by rotating, scaling and translating a unit circle into an ellipse based on the covariance matrix and means of all scores for PC1 and PC2, using 2 standard deviations as the scaling factor.

2.7.5 Elastic net logistic regression model

Elastic net logistic regression was performed using scikit-learn v1.0 for variable selection prior to subsequent multivariate analyses and modelling. A 10-fold cross-validation elastic net logistic regression model was performed with a grid search over (1) the penalty term C from 10^{-5} to 10^{5} increasing geometrically by a factor of $10^{0.5}$, and (2) the L1 ratio from 0 to 1 in increments of 0.1, where 0 represents an L2-only penalty, and 1 represent an L1-only penalty. Using a probability threshold of 0.5 for a positive disease state, the model with the highest average accuracy across all validation sets for each fold was chosen as the optimal model. All variables with non-zero coefficients were selected and kept for subsequent analysis. If the highest average accuracy across all validation sets was not greater than a baseline model that predicts the mode of the disease state in the training set, no optimal model was identified and all input variables were kept.

2.7.6 Partial Least Squares Discriminant Analysis (PLS-DA)

Partial least squares discriminant analysis (PLS-DA) was performed using scikit-learn v1.0, generating a projection matrix for both inputs and outputs, where each variable including the dependent variable was assigned a loading for 2 latent variables designed to maximise covariance between the inputs and outputs. The linear combination of all loadings and their respective inputs for a given LV and patient produced an LV score for that patient. All input loadings and therefore scores were multiplied by the sign of the

output loading for each latent variable, in order to represent the (directionless) latent variables such that more positive values were associated with positive disease states. Ninety-five percent confidence regions for disease states were created using matplotlib v3.4 and numpy v1.20 by rotating, scaling and translating a unit circle into an ellipse based on the covariance matrix and means of all scores for LV1 and LV2, using 2 standard deviations as the scaling factor.

2.7.7 Logistic regression analysis

Logistic regression was performed using statsmodels v0.12, generating a coefficient for each input variable such that the linear combination of each input and associated coefficient represents the log odds of that patient exhibiting a positive disease state. Exponentiating the coefficient represents how a one-unit change in the model input, equivalent to a 10-fold change in the original measurement, is associated with a change in the odds of a positive disease state. For the HTLV-1c positive-only cohort, the logistic regression was run again without right-shifting proviral load by 1 unit prior to a $\log_{10}$ transform, to confirm the interpretation of how a 10-fold change in proviral load was associated with the odds of HAPD. The coefficient was equal up to 4 significant figures, indicating the right-shift had limited impact on the interpretability of odds ratios for proviral load.

2.8 CD4+ T cell phenotyping

Cryopreserved PBMCs were thawed and washed with fluorescence activated cell sorting (FACS) buffer before being incubated with a cocktail of surface antibodies in FACS buffer (**Table 2.3**): CD3 PerCP-Cy5.5 (SK7), CD4 Alexa R700 (RPA-T4), CD38 PE-Cy7 (HB7), CD45RA PE-CF594 (HI100), CD127 BV786 (HIL-7R-M21), HLA-DR BV711 (G46-6), Integrin-β7 BV605 (FIB504) (BD Biosciences); CD49d BV510 (9F10) (Biolegend); CCR4 PE (205410) (R&D Systems) for 30 minutes at 4°C. Samples were washed twice: first with FACS buffer and then with PBS 1X, and then stained for live/dead (NIR) (Invitrogen) and incubated at 4°C for 15 minutes. Cells were washed with FACS buffer and filtered using a 70μM sieve. Single colour controls were prepared using OneComp eBeads™ (Invitrogen), and live/dead compensation controls were prepared using ArC™ Amine Reactive Compensation beads (Invitrogen). Samples were

processed using FACS on Aria III or Aria Fusion or acquired on Fortessa (BD Biosciences). Sorted CD4+ T cell populations were collected in RPMI media (Gibco) for further immediate processing. Data was acquired using FACSDiva™ software v9.0 (BD Biosciences) and processed using FlowJo v10.8.0 software (FlowJo).

2.9 Single provirus amplification (SPA) PCR assay

2.9.1 Design and optimisation of SPA assay

Single provirus amplification (SPA) assay is a limiting dilution touchdown nested PCR, designed to amplify single copies of integrated HTLV-1c provirus, modelled on a HIV-1 proviral sequencing study with a similar approach [270]. HTLV-1c provirus primers were designed with Primer3Plus web-interface program (version 3.2.6) [271] and SnapGene to determine the sequences with optimal length, GC content, annealing temperatures, while avoiding hairpin structures or 3' complementarity. Given the presence of LTRs at both ends of the provirus, primers had to be staggered to amplify as much of the full-length provirus as possible without unwanted amplification of an LTR section. Moreover, the amplification included more of the 3'LTR and less of the 5' LTR, given there have been many reports of 5'-deletions found in HTLV-1a infections, particularly ATL cases [153, 272, 273].

The SPA assay uses limiting dilution such that less than 30% of wells contain a copy of proviral amplification, which by Poisson distribution indicates that there is more than 80% probability that this amplification comes from a single integrated provirus. Raw ddPCR data was used to calculate limiting dilutions for the SPA assay. However, PVL is an absolute measurement of the number of *tax* gene regions detected, but does not consider proviral deletions that may be present, so concentration optimising SPA PCRS were necessary to find the exact dilution required for each individual sample.

2.9.2 SPA PCR protocol

Both rounds of PCR reaction mixture were 20µl with final concentrations of 0.5U of Platinum Taq DNA Polymerase High Fidelity (Invitrogen, Waltham, Massachusetts, USA), 1µM of both forward and reverse primers (**Table 2.1**), 2mM $MgSO_4$ (Invitrogen), 0.2mM dNTPs in 1X high fidelity (HI-FI) buffer (Invitrogen). The first PCR round with

1µl of limiting dilution gDNA, from sorted CD4+ T cell subsets, bulk PBMCs and splenocytes, used a touchdown protocol on the Mastercycler Nexus (Eppendorf) thermocycler as follows: enzymatic activation for 2 minutes at 94°C, then three cycles of each annealing temperature (denaturation for 30 seconds at 94°C, annealing for 30 seconds at 64/61/58°C, extension for 10 minutes at 68°C) followed by 21 cycles of (denaturation for 30 seconds at 94°C, annealing for 30 seconds at 55°C, extension for 10 minutes at 68°C), a final extension at 68°C for 10 minutes, finishing with an infinite hold at 10°C. The nested PCR round used 1µl of PCR product from the first round as the template and applied the same PCR touchdown thermocycler protocol as the first round, but instead with 31 cycles of annealing temperature 55°C before the final extension time.

2.9.3 Visualisation and purification of single provirus amplifications

PCR products (1µl) in Gel loading dye 1X (NEB, Ipswich, Massachusetts, USA) were electrophoresed in an 0.8% agarose gel with GelRed 1X (Biotium) in TAE buffer 1X at 100V, for visualisation of positive wells with a UV transilluminator. Positive wells for SPA were selected and all remaining PCR product was electrophoresed as above, and provirus bands were excised under UV light.

DNA was purified using NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel, Düren, Germany). Firstly, gel slices were solubilised in Buffer NTI (Macherey-Nagel) with incubation at 55°C and periodic vortexing until completely dissolved. DNA was loaded and bound to Nucleospin® column (Macherey-Nagel) and then washed twice with Buffer NT3 (Macherey-Nagel), via centrifugation. The column silica membrane was dried by centrifugation to remove residual ethanol. Buffer NE (30µl) (Macherey-Nagel) was added to the column and incubated at 70°C for 5 minutes, before elution of DNA. SPA DNA products were stored at -20°C for further analysis.

2.10 Long-read HTLV-1c provirus sequencing

2.10.1 Oxford Nanopore Technologies (ONT) sequencing

SPA DNA concentration was quantified and subject to quality control using Qubit® 3.0 fluorometer (ThermoFisher Scientific). DNA libraries were prepared for Oxford Nanopore Technologies (ONT) sequencing [274, 275] using the ligation sequencing kit

as per the manufacturer's instructions (SQK-LSK109, ONT) with native barcoding (EXP-NBD196, ONT), which ensured no nucleotide mutations were introduced to sequences, such is the possibility during PCR-based barcoding. In brief, individual amplicons were subject to end repair, dA-tailing, and barcode ligation. Amplicons were then pooled in an equimolar ratio, and sequencing adapters were ligated to the pool of DNA. Barcoded libraries were sequenced on the ONT MinION Mk1C platform, r9.4.1 chemistry. Base calling, demultiplexing and adapter trimming were performed using Guppy v1.8.5 (ONT). Reads were subset from the peak of the read length distribution generated using NanoPlot v1.40.0 and filtered for the presence of primer sequences at the ends of each fragment. Lamassemble [276] (GitLab, San Francisco, California, USA) was used to generate consensus sequences for each provirus amplicon, which were derived from 1000 full-length amplicon reads for each individual proviral sequence. Using 1000 reads to generate a single consensus sequence for each provirus amplicon mitigated the high basecaller error rates (~10%) in ONT sequencing [277].

2.10.2 Sequence alignments and analysis

Proviral consensus sequences were aligned to the full length HTLV-1c consensus sequence (Hirons et al., unpublished) using BLASTn web tool [278] (NCBI, Bethesda, Maryland, USA). To permit visualisation, reads were then mapped to the amplified region (680 - 8899 bp) with burror-wheeler aligner maximal exact match (BWA-MEM) using default parameters (algorithm unpublished). Alignments were visualised and quantified in Seqmonk v1.48.0 (Babraham Bioinformatics, Babraham Institute, Cambridge, UK). Relative representation of each nucleotide in proviral amplicons was quantified by sliding window, using 1bp windows with step size of 1bp, within 680-8898 bp of the full-length proviral sequence. Sampling between patients was normalised by scaling to the largest store. Breakpoint analysis was performed on the BLAST (NCBI) alignments, and a sequence logo of 30 bp ± breakpoints in each amplicon was generated using WebLogo [279] (University of California, Berkeley, California, USA). Defective proviral sequences were identified at the alignment stage. Amplicons containing large internal deletions were defined as gaps > 100 nucleotides. Inversions were identified as unmappable regions that were able to be mapped to the reference sequence following reverse complementation. Chimeric HTLV-1c proviral amplicons that contained internal human genetic sequences were identified as regions that did not map to the HTLV-1c

consensus, but which did map to the HS1 human genome assembly. These regions were extracted identified using BLAT web tool [280] and visualised on Seqmonk v1.48.0.

2.11 HTLV-1c-infected humanised mouse model

2.11.1 Maintenance of hu-NSG mouse strain

All work involving the humanised (hu) mouse model was approved by the Walter and Eliza Hall Institute of Medical Research (WEHI) Animal Ethics Committee (2017.016).

NOD-*scid* IL2Rgamma^{null} (NSG) mice were bred and maintained in the PC3 animal facility at WEHI, Melbourne. In each cage, mice were subjected to the following conditions: 12/12 hour light/dark cycles, 50 air exchanges per hour, 18-22°C air temperature and 50-70% humidity. Mice were provided sterilised water, a standardised diet, and corn cob granulate bedding (SAFE-lab, Rosenberg, Germany). Monitoring mice health involved daily physical health examinations and weekly weigh-ins (Cooney et al., 2023, Cell, in revision).

2.11.2 Generation of humanised mice

Twenty-four hours old mice pups were sub-lethally irradiated (150rad) using Cobalt-60 (Co-60) teletherapy machine (Best Theratronics Ltd. Ontario, Canada). After two hours, pups were intra-facially injected with $5x10^4$ – $1x10^5$ CD34+ human cord blood stem cells (Lonza, Basel, Switzerland) in PBS 1X. Mice were monitored for 16 weeks, at which point they were analysed for reconstitution of a human immune system (HIS). Peripheral blood obtained from submandibular bleeds were analysed for human CD45 frequency by flow cytometry. All mice with CD45 positivity between 20-90% were selected for HTLV-1c experimentation, while mice with a frequency falling outside these values were excluded from the study (Cooney et al., 2023, Cell, in revision).

2.11.3 Infection of hu-NSG mice with HTLV-1c

The first HTLV-1c transmission event into the hu-NSG mouse model was achieved using HTLV-1c-infected primary human patient PBMCs. Cryopreserved PBMCs were thawed and lethally irradiated (77Gy, Co-60) and $1x10^6$ PBMCs were intraperitoneally injected into each hu-NSG mouse. Successful HTLV-1c infection was determined at 2

weeks post-infection (wpi) with ddPCR as described in section **2.4**, by using peripheral blood obtained from a submandibular bleed and extracting gDNA as previously described in section **2.3**. Hu-NSG mice were euthanised by asphyxiation and harvested spleens were homogenised and filtered through a 40µm sieve. HTLV-1c infected splenocytes were cultured ex vivo in RF10, supplemented with 20U/ml human recombinant interleukin-2 (Lonza, Basel Switzerland), and then cryopreserved until further hu-NSG mouse infections (Cooney et al., 2023, Cell, in revision).

All subsequent HTLV-1c infections of hu-NSG mice used $5x10^5$ lethally irradiated (77Gy, Co-60) HTLV-1c infected splenocytes, through intraperitoneal injection. At the conclusion of the experiment, hu-NSG mice were euthanised by asphyxiation, or earlier if mouse weight loss exceeded 20%. Spleens were homogenised and filtered using a 40µm strainer, and either immediately analysed using downstream applications, or cryopreserved until further use (Cooney et al., 2023, Cell, in revision).

3. Characterising the host immune response to HTLV-1c infection

3.1 Chapter three background

HTLV-1 has infected humans for tens of thousands of years and maintains a lifelong infection in the host and therefore has developed many mechanisms with which to modulate immune responses and persist. Nevertheless, individuals infected with HTLV-1 have a significant higher all-cause mortality [14], with many developing debilitating or life-ending associated diseases. It is clear that host immune dysfunction accompanying HTLV-1 infection contributes to these poor outcomes. The host immune response to HTLV-1a infection, including factors that influence disease progression, have been extensively studied and are well defined. However, given the distinctive features of subtype-C, a detailed analysis surrounding the host response to the lesser studied HTLV-1c infection is still required.

3.1.1 CD4+ T cell responses

During any viral infection CD4+ T cell responses are a key component of the host-virus interplay and contribute crucially to the ability of the host to either control and eliminate the infection, or sustained persistence of the virus, immune dysfunction and development of associated disease. Given CD4+ T cells are the primary reservoir for HTLV-1 infection, yet play a cornerstone role for infection progression or control, we investigated the HTLV-1c specific responses of several key CD4+ T cell phenotypes, including circulating T follicular helper cells (cTFH), memory T (Tmem) cells and T helper (Th) memory cells.

Tmem cells are crucial to orchestrating antigen-specific responses; they are selected after prior exposure to antigens and respond rapidly upon re-challenge to induce early effector cytokines, to provide B cell and CD8+ T cell help, or even directly kill infected cells, to combat infection [281]. However, dysregulated responses can lead to chronic activation and detrimental inflammation [282]. Within the Tmem subset, we explored Th cells as they are one of the key instigators and regulators of the adaptive immune response, activating the CTL, B cell and macrophage responses. Th cells can differentiate into several effector subsets upon activation, including Th1, Th2 and Th17 cells. Given that Th2 cells are commonly involved in the defence against extracellular pathogens and chronic inflammatory allergic responses [283], this chapter focuses on Th1 and Th17 cells. Th1 cells are known for the production of IFN-γ as their signature cytokine, and

promote tissue infiltration of macrophages at sites of inflammation [284]. Th17 cells are known for their production of IL-17, but also promote the production of a plethora of other cytokines, including TNF-α and monocyte chemoattractant protein 1 (MCP-1), and recruit neutrophils to infiltrate affected tissues [284]. Th1 and Th17 cells may display antagonistic or collaborative effector functions, depending on the tissue type, and concurrent inflammation and infection [284]. Th17 cells have the capability to convert to Th1 phenotype, and there is a transitional state in which the cell shows characteristics of both [285]. Th1Th17 cells may be important in promoting inflammation, and further demonstrate the tightly regulated balance between various T helper cells and cytokines.

TFH cells are fundamental in generating antibody responses by promoting class-switching and differentiation of B cells in germinal centres. cTFH cells, also crucial in the humoral immune response, provide supplementary B cell help in the blood [286] and their dysregulation is linked to several inflammatory and autoimmune diseases [287]. They provide a proxy measurement of the TFH response in secondary lymphoid organs [286].

The CD4+ T cell response in HTLV-1a infection is well characterized. There are significantly higher levels of HTLV-1a-specific CD4+ T cells, more specifically Th1 phenotype, in HAM donors than AC [288]. HTLV-1a-specific CD4+ T cells secrete higher levels of IFN-γ and IL-2 in HAM donors when compared to AC [289]. Furthermore, these HTLV-1a-specific cells are preferentially infected with HTLV-1a [290]. Together, this suggests that the HTLV-1a specific CD4+ T cell response is associated with HAM disease outcome, and not merely due to higher proviral loads in these donors [290].

3.1.2 CD8+ T cell responses

CD8+ T cells are key effectors of the adaptive immune response in viral infections for destruction of infected cells. Cytotoxic T cells (Tc) recognise foreign antigen presented on MHC-I and become functional effector CTLs though signalling pathways [291]. They can induce apoptosis of infected cells through releasing perforins and granzymes at the infected cell membrane or through the Fas ligand pathway [292].

The CD8+ T cell response to HTLV-1a infection has been well characterised, as one that contributes to concurrent protection and pathogenesis. HTLV-1a-specific CD8+ T cell

responses in HAM patients correlate with high PVL [293], suggesting that the robust CTL response is ineffective at controlling infection. Indeed, HAM donors were shown to have lower levels of intracellular perforin-positive CD8+ cells, and exhibited a CD28-CD27- phenotype, as opposed to ACs mostly showing CD28-CD27+ phenotype CD8+ T cells [294]. Consistently, HTLV-1a Tax-specific CD8+ T cells were shown to be partially exhausted and contribute to systemic inflammation and progression to HAM [295].

3.1.3 Double negative (DN) T cell responses

Double negative (DN) T cells (CD3+CD4-CD8-) are a somewhat rare subset of T cells, making up between 1-5% of T cells in peripheral blood, and express either αβ- or γδ-TCRs [296, 297]. DN T cells display heterogeneous immune functions depending on context of activation, inflammation and infection, such as immunomodulatory responses or effector functions [298]. Studies on DN T cells in HTLV-1 infection are currently limited, so we investigated this phenotype to elucidate any potential impacts on pathogenesis, immune dysfunction and disease development that this subset of cells may have.

3.1.4 Cytokines and chemokines

Cytokines are produced by a plethora of immune cells and are important for the body to combat infection: they cause inflammation by directing immune cells where they are needed. The complex cytokine networks can trigger many downstream signalling pathways and help create the optimal host defence against invading pathogens [299]. However, when the optimal levels of cytokines become unbalanced during a response, it can cause a plethora of issues for the host, from a cytokine storm like the one observed during severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection [300], to immune exhaustion and inflammation from chronic viral infections [301].

The cytokine signature of HAM patients with HTLV-1a infection is pro-inflammatory. Plasma levels of TNF-α, IFN-γ, IL-4, IL-6, and IL-8 are elevated in HAM patients when compared to ACs and healthy donors [302-304]. CSF and serum levels of IFN-γ-induced protein 10 (IP-10)/CXC motif chemokine ligand 10 (CXCL10) and CXCL9 are elevated in HAM donors [305-307], while there are significantly lower levels of MCP-1 [306].

HTLV-1a infected donors with ATL show a contrasting response dominated by immunosuppressive cytokines, contributing to the rapid proliferation of infected cells. ATL donors display elevated levels of IL-10 [240], IL-6 [308] and TGF-β [309], concurrent with significantly lower levels of IFN-γ [310], when compared to HAM and AC donors.

3.2 Chapter three aims

This study aims to characterise host immune response to HTLV-1c infection in a cohort of First Nations peoples from remote communities in Central Australia. In particular, we will characterise the HTLV-1c specific CD4+ and CD8+ T cell responses and cytokine profiles of HAPD- and HAPD+ donors. We hypothesize there are increased levels of HTLV-1c-specific CD4+ and CD8+ T cells in HAPD+ donors when compared to HAPD-. We hypothesize there is elevated activation frequencies of CD4+ and CD8+ T cell phenotypes in HTLV-1c-infected donors when compared to uninfected donors, and further differences between HAPD+ and HAPD- donors.

We hypothesize that HTLV-1c infection has a signature cytokine response, and much like subtype-A associated diseases, there is a distinct difference between HAPD- and HAPD+ donors. As bronchiectasis is characterised by chronic inflammation and damage to the bronchial tubes, we hypothesise that there are elevated levels of pro-inflammatory cytokines, like in HAM, as opposed to immune suppressing cytokines as seen in ATL. Furthermore, we hypothesise the PVL is correlated to the levels of pro-inflammatory cytokines.

3.3 Characterising the HTLV-1c-specific T cell responses

The HTLV-1a specific CD4+ and CD8+ T cell responses have been well characterised in HAM and ATL patients, however it is yet to be elucidated in HTLV-1c. Envelope-specific CD4+ T cell responses [290] and Tax-specific CD8+ T cell responses [171] are the most prominent in HTLV-1a, therefore we designed 15-mer overlapping peptide libraries spanning the length of HTLV-1c Env and Tax protein coding regions. PBMCs,

including uninfected controls from Naarm (n, 4), HTLV-1c-infected donors who were HAPD negative from Mparntwe (n, 8) and HTLV-1c-infected HAPD positive donors from Mparntwe (n, 8), were stimulated for 20 hours with DMSO (background control), Env peptides, Tax peptides and SEB (positive control), and analysed via flow cytometry to determine the HTLV-1c-specific T cell responses (**Figure 3.1**). We used negative control samples from Naarm in this assay because there were insufficient numbers of cryopreserved seronegative PBMCs received of donors also from Mparntwe, to test all four conditions in each sample. A summary of the donor demographics and clinical characteristics used in this investigation is described in **Table 3.1**. To investigate the host-virus relationship further, we determined the PVL for all HTLV-1c-infected donors, shown in **Table 3.1**, using a previously established HTLV-1c *tax* droplet digital PCR (ddPCR) assay [218]. The only significant clinical and demographic difference between the two HTLV-1c seropositive test groups was the HAPD status (p, 0.0002). We determined the antigen-specific T cell responses, baseline levels of activation and phenotype frequency, and correlations with PVL.

Table 3.1 – Cohort demographics and clinical characteristics of HTLV-1c T cell response study.

	HTLV-1c+ HAPD- (n, 8)	HTLV-1c+ HAPD+ (n, 8)	p-value
Demographics, n (%)			
Median Age	61.5	54.0	0.88
Female at birth	7/8 (87.5%)	5/8 (62.5%)	0.57
Male at birth	1/8 (12.5%)	3/8 (37.5%)	
Lifestyle, n (%)			
Smoker (current or previous)	1/8 (12.5%)	2/8 (25%)	>0.99
Harmful alcohol consumption	3/8 (37.5%)	1/8 (12.5%)	0.57
Comorbidities, n (%)			
Chronic Liver Disease	2/8 (25%)	0/8 (0%)	0.47
Diabetes	7/8 (87.5%)	4/8 (50%)	0.28
Chronic Kidney Disease	7/8 (87.5%)	4/8 (50%)	0.28
Malignancy	0/8 (0%)	1/8 (12.5%)	>0.99
HTLV-1 associated diseases, n (%)			
HAPD (Bex)	0/8 (0%)	8/8 (100%)	**0.0002*** **
Strongyloidiasis, current	0/8 (0%)	2/8 (25%)	0.47
BSI, current or previous	3/8 (37.5%)	2/8 (25%)	>0.99
HAM	0/8 (0%)	0/8 (0%)	>0.99
ATL	0/8 (0%)	0/8 (0%)	>0.99
HTLV-1 burden, log_{10}(copies per 10^6 cells)			
Median PVL	3.31	2.98	0.96

HTLV-1c-infected participants consented to whole blood being taken and used for HTLV-1c pathogenesis research following written and verbal consent in primary language. Information for HTLV-1c+ donors was obtained through clinician and hospital records at Alice Springs Hospital, Mparntwe, NT, Australia. PVL was determined with gDNA extracted from PBMCs and quantified using ddPCR. HTLV-1c-uninfected samples were obtained from Lifeblood Red Cross, Naarm, Melbourne, Australia, and were routinely screened at blood donation for infectious diseases, but no other demographic or clinical details were provided. n, 20 (4 HTLV-1c- (Naarm), 8 HTLV-1c+ HAPD- (Mparntwe), 8 HTLV-1c+ HAPD+ (Mparntwe)). Statistical significance was determined by using Mann-Whitney test for continuous variables, and Fisher's exact test for categorical variables. *p<0.05; **p<0.01; ***p<0.001.

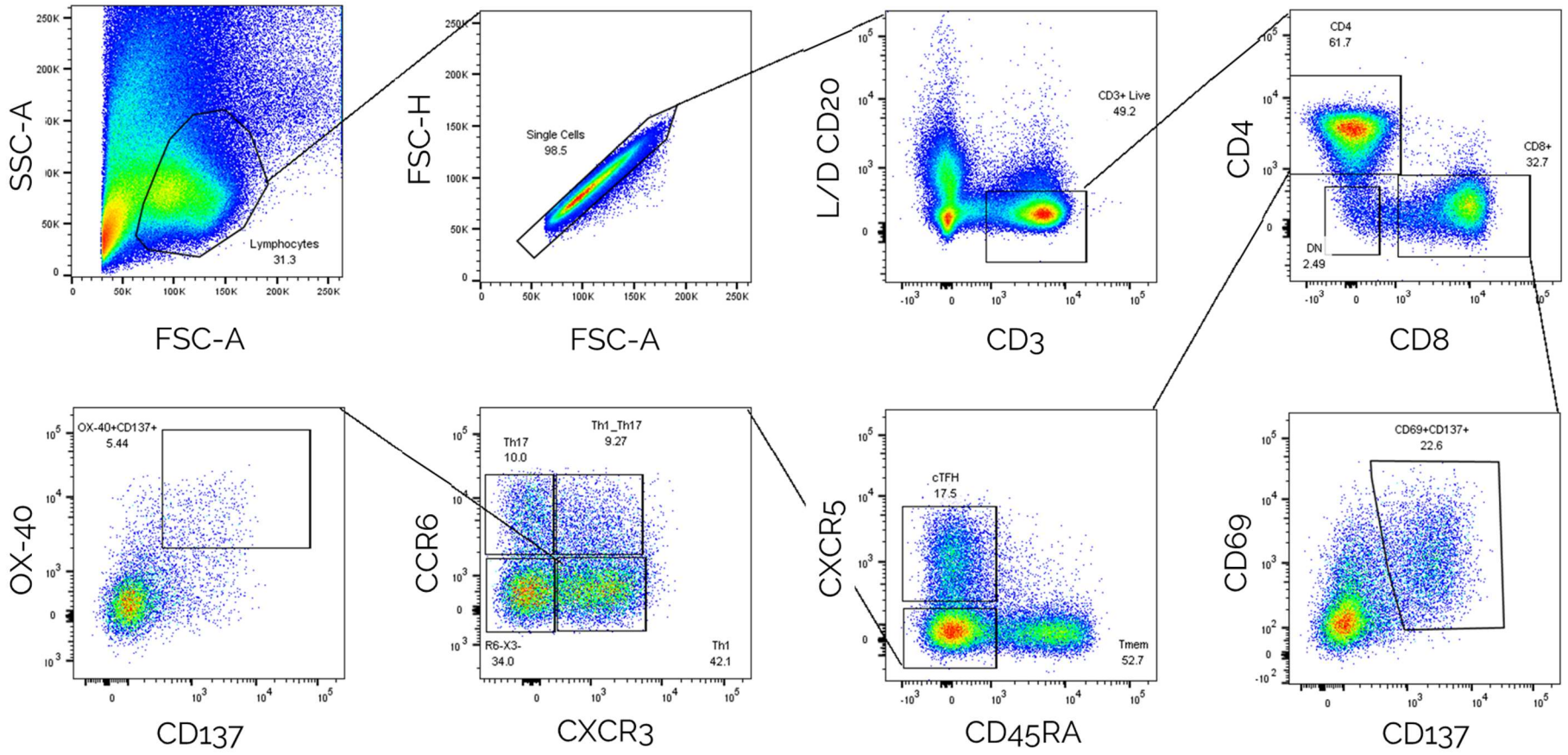


Figure 3.1 – Flow cytometry gating strategy for HTLV-1c activation induced marker assay. PBMCs were stimulated for 20 hours with either DMSO, Tax or Env peptide pools, or SEB and then incubated with a cocktail of surface mAbs. Cells were acquired by flow cytometry. CD4+ T cell phenotypes of interest were cTFH, Tmem, Th1, Th17, Th1Th17 and CD8+ T cells, alongside activation markers OX-40, CD137, CD69.

3.3.1 HTLV-1c infection causes expansion of CD4+ Tmem and cTFH phenotypes and dysregulation of Th1 frequency

To determine if HTLV-1c infection is associated with the expansion of various CD4+ T cell subsets, we first measured the total frequencies of unstimulated (DMSO condition) CD4+ T cell populations. The total frequencies of unstimulated cTFH cells were significantly higher in HAPD- (median, 9.04%; IQR, 7.61-15.10; p, 0.05) and HAPD+ (median, 15.00%; IQR, 10.09-16.45; p, 0.005) donors when compared to uninfected Naarm controls (median, 3.60%; IQR, 2.68-4.43) (**Figure 3.2A**). However, there were no significant differences between the two HTLV-1c infected groups (p, 0.14) (**Figure 3.2A**). Furthermore, the frequency of cTFH cells did not correlate with PVL for the HTLV-1c infected individuals (r, -0.07; p, 0.81) (**Figure 3.2B**).

The total frequencies of unstimulated conventional Tmem cells (CD3+CD4+CD45RA-CXCR5-) were significantly higher in the HAPD- (median, 54.55%; IQR, 35.45-69.20; p, 0.034) and HAPD+ donors (median, 55.30%; IQR, 47.83-59.78; p, 0.02) than negative controls (median, 27.60%; IQR, 21.93-37.55), however between the two HTLV-1c-infected groups there was no difference (p, 0.45) (**Figure 3.2A**). Moreover, the unstimulated Tmem frequency of HTLV-1c seropositive donors did not correlate to PVL (r, -0.09; p, 0.75) (**Figure 3.2B**).

Within the Tmem subset, we then investigated the frequency of Th subsets. Firstly, there were no significant differences in the unstimulated Th1 frequency of HAPD- (median, 31.90%; IQR, 23.38-39.88) and HAPD+ (median, 35.60%; IQR, 32.05-38.18) donors when compared to each other (p, 0.60) or to uninfected controls (median, 34.30%; IQR, 28.48-43.13; p, 0.62, 0.62, respectively) (**Figure 3.2A**). Interestingly however, the Th1 frequency of all HTLV-1c seropositive individuals was significantly negatively correlated to PVL (r, -0.70; p, 0.003) (**Figure 3.2B**). Next, our data indicated there were no differences in the unstimulated Th17 frequency between uninfected donors (median, 7.26%; IQR, 7.02-8.24) and HAPD- (median, 10.60%; IQR, 8.43-16.80; p, 0.17) or HAPD+ patients (median, 7.99%; IQR, 5.42-12.93; p, 0.82), nor between the two HTLV-1c-infected groups (p, 0.17) (**Figure 3.2A**). Furthermore, the Th17 frequency in HTLV-1c+ donors did not correlate to PVL (r, 0.26; p, 0.34) (**Figure 3.2B**). Finally, we investigated the transitional phenotype of Th1Th17 and found that the frequency in HAPD- (median, 9.20%; IQR, 5.33-13.15; p, 0.13) and HAPD+ (median, 6.02%; IQR,

4.40-10.28; p, 0.34) donors was trending higher, however not significantly different from negative controls (median, 4.52%; IQR, 4.36-4.89), nor between each other (p, 0.35) (**Figure 3.2A**). Likewise, the PVL of HTLV-1c infected donors was not correlated to the respective Th17 frequency (r, -0.25; p, 0.34) (**Figure 3.2B**).

3.3.2 HTLV-1c infection causes significant background CD4+ T cell activation

To further investigate the CD4+ T cell responses to HTLV-1c infection, we next measured the background level of activation of CD4+ T cells in DMSO treated condition of uninfected, and HTLV-1c seropositive HAPD- and HAPD+ donors. Overall, there was a significantly high level of background activation in all CD4+ phenotypes, as shown by the high percentage of OX-40+CD137+ cells, for both HAPD- and HAPD+ donors, when compared to negative controls (**Figure 3.3A**), indicating extensive chronic inflammation in all HTLV-1c-infected individuals. Similar results were also observed measuring OX-40+CD154+ expression (data not shown).

Firstly, the frequency of cTFH background activation was significantly higher in both HAPD- (median, 0.675%; IQR, 0.270-0.725; p, 0.01) and HAPD+ donors (median, 1.580%; IQR, 0.443-3.310; p, 0.001) when compared to uninfected controls (median, 0.042%; IQR, 0.003-0.083) (**Figure 3.3A**). Between the HTLV-1c seropositive groups, HAPD+ donors exhibited a higher activation trend than HAPD-, however this difference was not significant (p, 0.08) (**Figure 3.3A**). Furthermore, the proportion of background activated cTFH did not correlate to PVL for HTLV-1c infected individuals (r, 0.25; p, 0.34) (**Figure 3.3B**).

Similarly, our data indicates the frequency of background activation in the Tmem reservoir was significantly higher in HAPD- (median, 0.380%; IQR, 0.208-0.983; p, 0.01) and HAPD+ donors (median, 0.965%; IQR, 0.345-1.788; p, 0.002) than negative controls from Naarm (median, 0.015%; IQR, 0.002-0.022). Between the HTLV-1c+ groups themselves, HAPD+ donors tended to display higher levels of Tmem background activation than HAPD- individuals, but this was not significantly different (p, 0.11) (**Figure 3.3A**). Interestingly, the PVL is significantly positively correlated to background levels of Tmem activation for all HTLV-1c infected individuals (r, 0.53; p, 0.04) (**Figure 3.3B**).

Within the Tmem subset, we then measured background activation levels in Th1, Th17 and Th1Th17 subsets. The frequency of Th1 background activation was significantly higher in HAPD- (median, 0.150%; IQR, 0.062-0.475; p, 0.007) and HAPD+ groups (median, 0.315%; IQR, 0.080-1.300; p, 0.003) than uninfected controls (median, 0.001%; IQR, 0-0.006), but not between each other (p, 0.17), although overall HAPD+ donors showed higher activation frequencies (**Figure 3.3A**). In particular, one Bex individual had 6.77% of Th1 cells activated in the absence of any ex vivo stimuli. Strikingly, the PVL of all HTLV-1c infected donors was significantly positively correlated to the background frequency of activated Th1 cells (r, 0.76; p, 0.001) **(Figure 3.3B)**.

Similarly, our data indicates the total frequencies of Th17 baseline activation were significantly higher than uninfected controls (median, 0.050%; IQR, 0.008-0.107) for both HAPD- (median, 1.380%; IQR, 0.688-2.793; p, 0.03) and HAPD+ (median, 3.115%; IQR, 2.180-5.528; p, 0.002) donors (**Figure 3.3A**). Consistently, there was a trend for HAPD+ donors to display higher Th17 baseline activation than HAPD-, but no significance was observed (p, 0.09) (**Figure 3.3A**). Moreover, the PVL of HTLV-1c seropositive donors did not correlate with levels of background Th17 activation (r, 0.27; p, 0.32) (**Figure 3.3B**).

The transitional phenotype Th1Th17 exhibited similar trends. HAPD- (median, 1.030%; IQR, 0.470-3.068; p, 0.01) and HAPD+ donors (median, 2.605%; IQR, 1.480-4.973; p, 0.001) showed significantly higher levels of Th1Th17 background activation when compared to uninfected donors (median, 0.126%; IQR, 0.018-0.233) (**Figure 3.3A**). Consistent with the other CD4+ T cell phenotypes, while HAPD+ donors displayed higher Th1Th17 activation frequencies than HAPD-, these differences were not significant (p, 0.07) (**Figure 3.3A**). Finally, the baseline activation frequency of Th1Th17 phenotype did not correlate to PVL for HTLV-1c infected individuals (r, 0.42; p, 0.11) (**Figure 3.3B**).

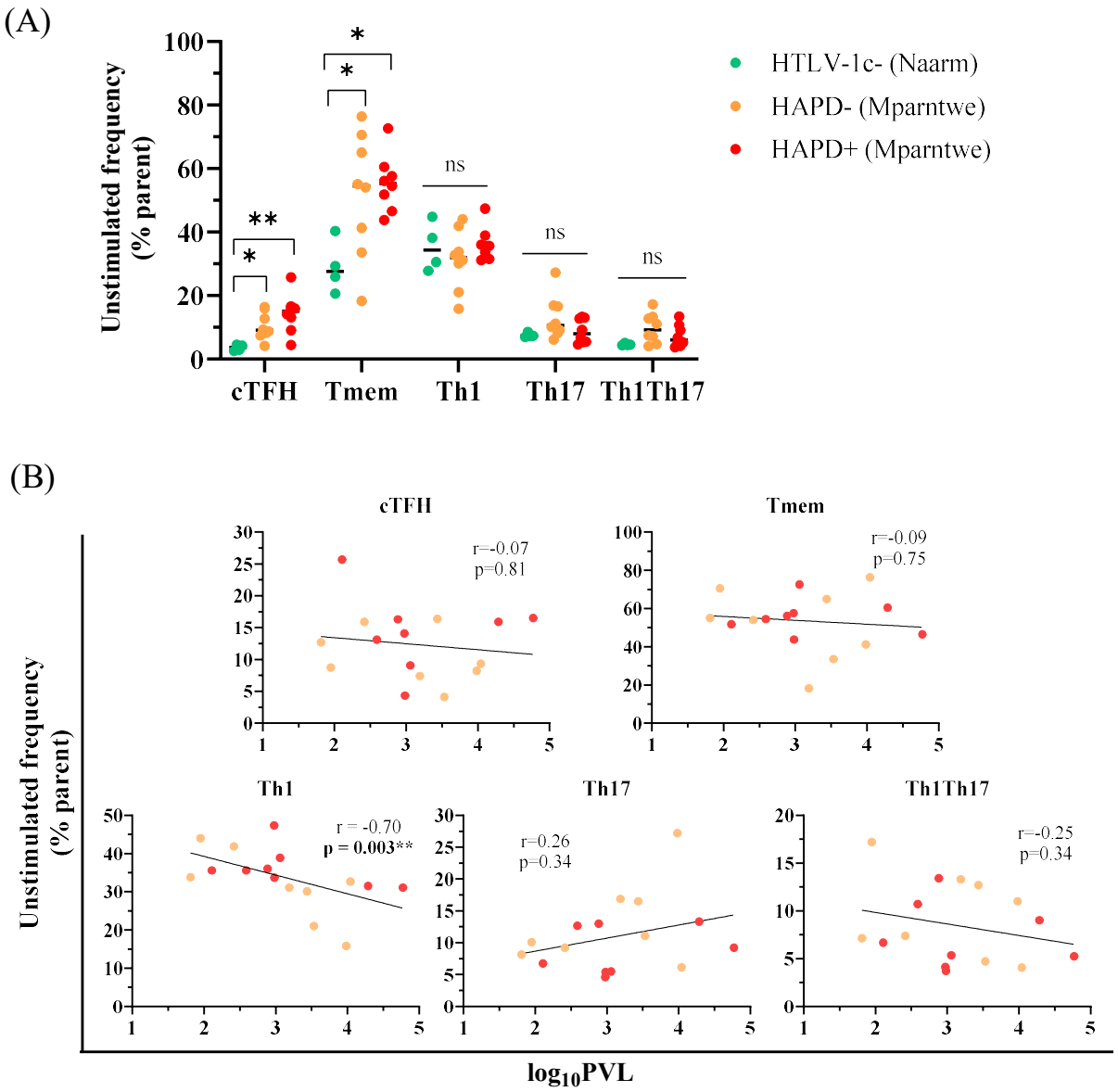


Figure 3.2 – HTLV-1c infection impacts the frequency of cTFH, Tmem and Th1 CD4+ phenotypes in primary patient PBMCs. PBMCs were stimulated for 20 hours with either DMSO, Env or Tax peptide libraries, or SEB, prior to surface staining with a cocktail of antibodies and processing via flow cytometry. gDNA was extracted from PBMCs and PVL was determined by ddPCR. (A) CD4+ T cell phenotype frequencies in DMSO treated condition. (B) Correlation of CD4+ T cell phenotype frequencies in DMSO condition and log_{10}PVL for HTLV-1c seropositive donors. n, 20: negative controls from Naarm (n, 4) are shown in green; HTLV-1c+ HAPD- from Mparntwe (n, 8) are shown in orange; and HTLV-1c+ HAPD+ donors from Mparntwe (n, 8) are shown in red. Median frequency of each test group is shown by black line. Statistical significance between groups was assessed using Kruskal-Wallis test with two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli as an FDR control for multiple comparisons, adjusted p values (q values) shown. Black trend line represents nonlinear regression straight line of best fit. Correlation significance was assessed using Spearman's correlation test. ns, not significant, $p>0.05$; * $p<0.05$; **$p<0.01$.

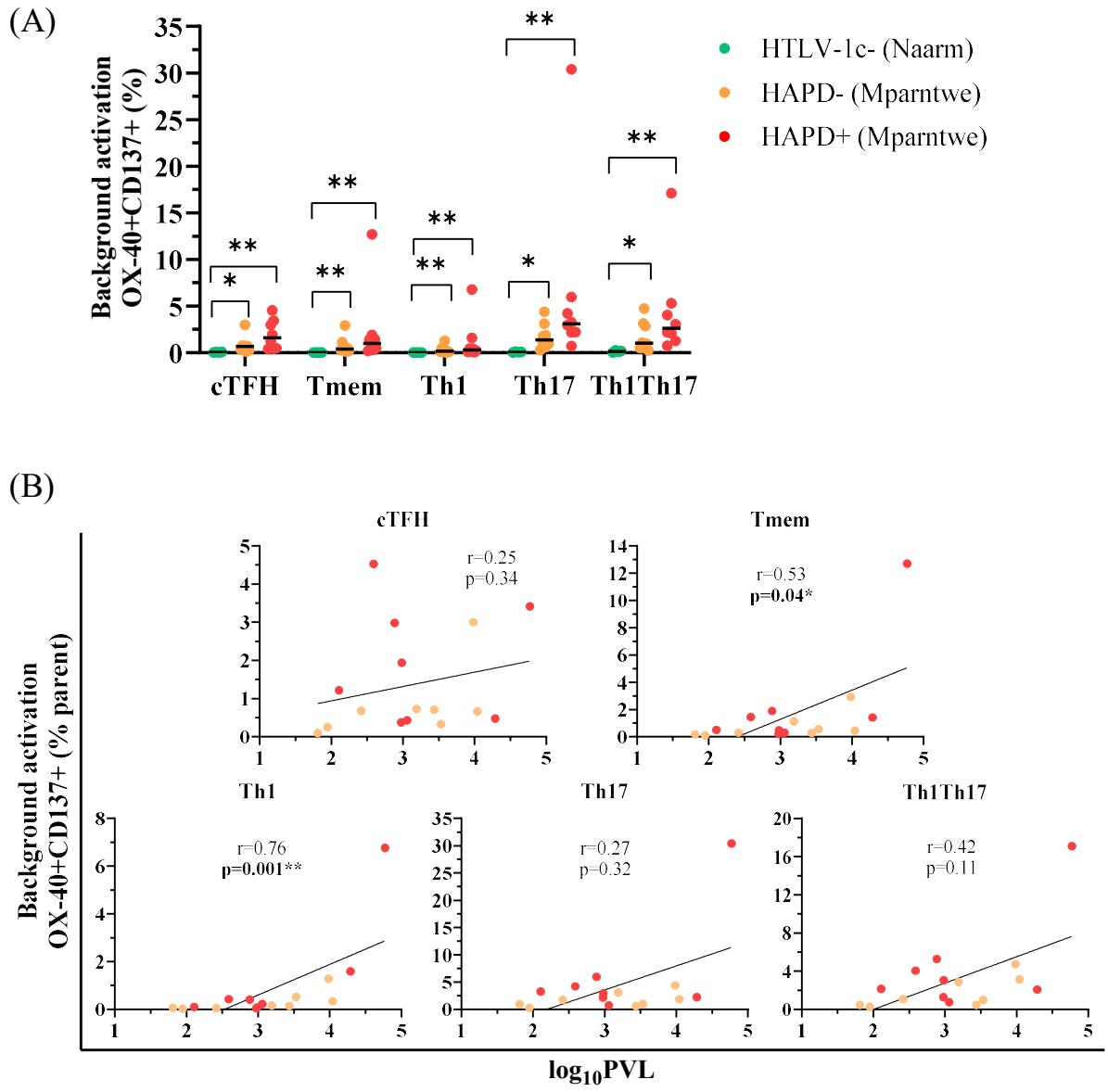


Figure 3.3 – HTLV-1c infection results in significant chronic CD4+ T cell background activation in primary patient PBMCs. PBMCs were stimulated for 20 hours with either DMSO, Env or Tax peptide libraries, or SEB, prior to surface staining with a cocktail of antibodies and processing via flow cytometry. gDNA was extracted from PBMCs and PVL was determined by ddPCR. (A) Frequency of activation surface marker expression (OX-40+CD137+) in CD4+ T cell phenotypes in DMSO treated condition. (B) Correlation of CD4+ T cell phenotype background activation (OX-40+CD137+) frequencies and $\log_{10}$PVL for HTLV-1c seropositive donors. n, 20: negative controls from Naarm (n, 4) are shown in green; HTLV-1c+ HAPD- from Mparntwe (n, 8) are shown in orange; and HTLV-1c+ HAPD+ donors from Mparntwe (n, 8) are shown in red. Median frequency of each test group is shown by black line. Statistical significance between groups was assessed using Kruskal-Wallis test with two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli as an FDR control for multiple comparisons, adjusted p values (q values) shown. Black trend line represents nonlinear regression straight line of best fit. Correlation significance was assessed using Spearman's correlation test. * $p<0.05$; **$p<0.01$.

3.3.3 HTLV-1c infection induces Tax- and Env-specific CD4+ T cells

Finally, to assess the HTLV-1c-specific responses of CD4+ T cells, we measured the frequency of activation marker expression (OX-40+CD137+) after stimulation with overlapping Env or Tax peptide libraries, and background subtracted from expression frequency in the DMSO treated condition.

Firstly, we detected HTLV-1c specific cTFH cells in both seropositive test groups. While the overall frequency of cTFH Tax-specific cells were higher in HAPD- (median, 0.270%; IQR, 0.062-0.270; p, 0.14) and HAPD+ (median, 0.0%; IQR, 0-0.578; p, 0.61) donors when compared to uninfected controls (median, 0.006%; IQR, 0-0.042), the differences were not statistically significant (**Figure 3.4A**). Similarly, there were no significant differences when comparing Tax-cTFH frequencies between seropositive test groups (p, 0.14) (**Figure 3.4A**). Furthermore, the frequency of Tax-specific cTFH in HTLV-1c infected donors did not correlate to PVL (r, 0.25; p, 0.36) (**Figure 3.4B**). Likewise, while there were Env-specific cTFH responses detected in HTLV-1c infected individuals, there were no significant differences between negative controls (median, 0.0%; IQR, 0-0.033) and HAPD- (median, 0.035%; IQR, 0-0.340; p, 0.63) and HAPD+ (median, 0.0%; IQR, 0-0.378; p, 0.63) donors, nor between the two HTLV-1c infected groups (p, 0.87) (**Figure 3.5A**). Moreover, the PVL of seropositive patients did not correlate with the respective Env-specific cTFH frequency (r, 0.43; p, 0.09) (**Figure 3.5B**).

Our data indicates HTLV-1c-specific Tmem cells were elevated in many infected donors. The frequency of Tax-Tmem cells were higher in HAPD- donors (median, 0.195%; IQR, 0.013-0.245; p, 0.06) than uninfected Naarm-based controls (median, 0.0%; IQR, 0-0.002), but there was little difference between controls and HAPD+ donors (median, 0.0%; IQR, 0-0.118; p, 0.47) (**Figure 3.4A**). When comparing the HTLV-1c+ groups, HAPD- donors tended to have higher Tax-specific Tmem frequency than HAPD+ donors, but significance was not observed (p, 0.09) (**Figure 3.4A**). Additionally, the PVL of HTLV-1c seropositive donors did not correlate to the frequency of Tax-specific Tmem cells (r, 0.15; p, 0.58) (**Figure 3.4B**). Next, our results demonstrated there were many HTLV-1c infected donors with Env-specific Tmem cells (**Figure 3.5A**). While the Env-Tmem frequencies were elevated in HAPD- (median, 0.015%; IQR, 0-0.378; p, 0.44) and HAPD+ (median, 0.130%; IQR, 0-0.768; p, 0.36) donors when compared to negative

controls (median, 0.0%; IQR, 0-0.011) they were not significantly different (**Figure 3.5A**). Moreover, the HAPD+ group showed slightly higher average Env-Tmem frequencies than HAPD-, but no significance was determined (p, 0.57) (**Figure 3.5A**). Notably, one Bex donor showed a robust response with 2.7% Env-specific Tmem cells (**Figure 3.5A**). Interestingly, the PVL of all HTLV-1c-infected donors positively correlated to the frequency of Env-specific Tmem cells (r, 0.58; p, 0.02) **(Figure 3.5B**).

Within the T helper subsets, we detected both Tax- and Env-specific cells across both HTLV-1c-infected donor groups. HAPD- donors had significantly higher levels of Tax-Th1 cells (median, 0.041%; IQR, 0.027-0.108; p, 0.05) than uninfected donors (median, 0.0%, IQR 0-0), and while some HAPD+ donors also showed elevated frequencies (median, 0.0%; IQR, 0-0.283; p, 0.21), this was not significantly higher than the control group (**Figure 3.4A**). Furthermore, there was no distinct difference in the Tax-Th1 frequency between HAPD- and HAPD+ donors (p, 0.21) (**Figure 3.4A**), nor did the PVL of these seropositive donors correlate to the proportion of Tax-specific Th1 cells (r, 0.36; p, 0.16) (**Figure 3.4B**). We detected higher frequencies of Env-specific Th1 cells in both HAPD- (median, 0.075%, IQR, 0-0.590; p, 0.26) and HAPD+ (median, 0.131%; IQR, 0.074-0.588; p, 0.11) groups above uninfected controls (median, 0.002%; IQR,0-0.012), however these were not significantly different (**Figure 3.5A**). Moreover, no significant difference was observed between the HTLV-1c+ groups (p, 0.39) (**Figure 3.5A**). However, most strikingly, the PVL of HTLV-1c seropositive donors was significantly positively correlated to the frequency of circulating Env-specific Th1 cells (r, 0.79; p, 0.0005) (**Figure 3.5B**).

Next, we determined the proportion of Tax-Th17 cells was significantly higher in HAPD- donors (median, 0.720%; IQR, 0.190-0.880; p, 0.03) when compared to negative controls (median, 0.0%; IQR, 0-0), and generally higher in HAPD+ donors (median, 0.0%; IQR, 0-2.743; p, 0.14), but not statistically significant (**Figure 3.4A**). Of the HTLV-1c+ groups, HAPD- displayed higher frequencies of Tax-specific Th17 cells than HAPD+ donors, but the observed trend was not significant (p, 0.15) (**Figure 3.4A**). In addition, the PVL of HTLV-1c seropositive donors did not correlate to Tax-Th17 frequency (r, -0.19; p, 0.49) (**Figure 3.4B**). Our data demonstrated elevated Env-specific Th17 frequencies in HAPD+ donors (median, 0.130%; IQR, 0-1.513) when compared to negative controls (median, 0.0%; IQR, 0-0.034), however significance was not observed

(p, 0.55). Moreover, there was no significant differences in frequency of Env-specific Th17 cells when comparing the uninfected control group to HAPD- donors (median, 0.0%; IQR, 0-0.368; p, 0.55) nor between the two HTLV-1c seropositive groups (p, 0.55) (**Figure 3.5A**). Notably, one Bex donor showed a robust response to Env peptides with 4.9% of Th17 cells activated above baseline levels. However, the proportion of Env-Th17 cells did not correlate to PVL for HTLV-1c infected individuals (r, 0.41; p, 0.12) (**Figure 3.5B**).

Finally, we showed that Tax-specific Th1Th17 cells were present in both HAPD- (median, 0.720%; IQR, 0.190-0.880) and HAPD+ (median, 0.0%; IQR, 0-2.743) donors, however the frequency observed was not significantly greater than negative control donors from Naarm (median, 0.0%; IQR, 0-0; p, 0.08, 0.22, respectively), nor between each other (p, 0.38) (**Figure 3.4A**). One HAPD- donor, however, displayed a very strong response to Tax peptides, with 4.65% of the Th1Th17 population being Tax-specific (**Figure 3.4A**). Meanwhile, the PVL of HTLV-1c infected donors was not correlated to the frequency of Tax-Th1Th17 phenotype (r, 0.13; p, 0.63) (**Figure 3.4B**). The frequency of Env-specific Th1Th17 cells trended higher in both HAPD- (median, 0.0%; IQR, 0-0.368) and HAPD+ (median, 0.130%; IQR, 0-1.513) groups when compared to seronegative donors (median 0.0%; IQR, 0-0.034), however no significance was observed (p, 0.19, 0.19, respectively) (**Figure 3.5A**). Moreover, there was no discernible difference in the frequency of Env-Th1Th17 cells between the HTLV-1c seropositive groups (p, 0.99) nor was the frequency correlated to PVL (r, 0.35; p, 0.18) (**Figure 3.5B**).

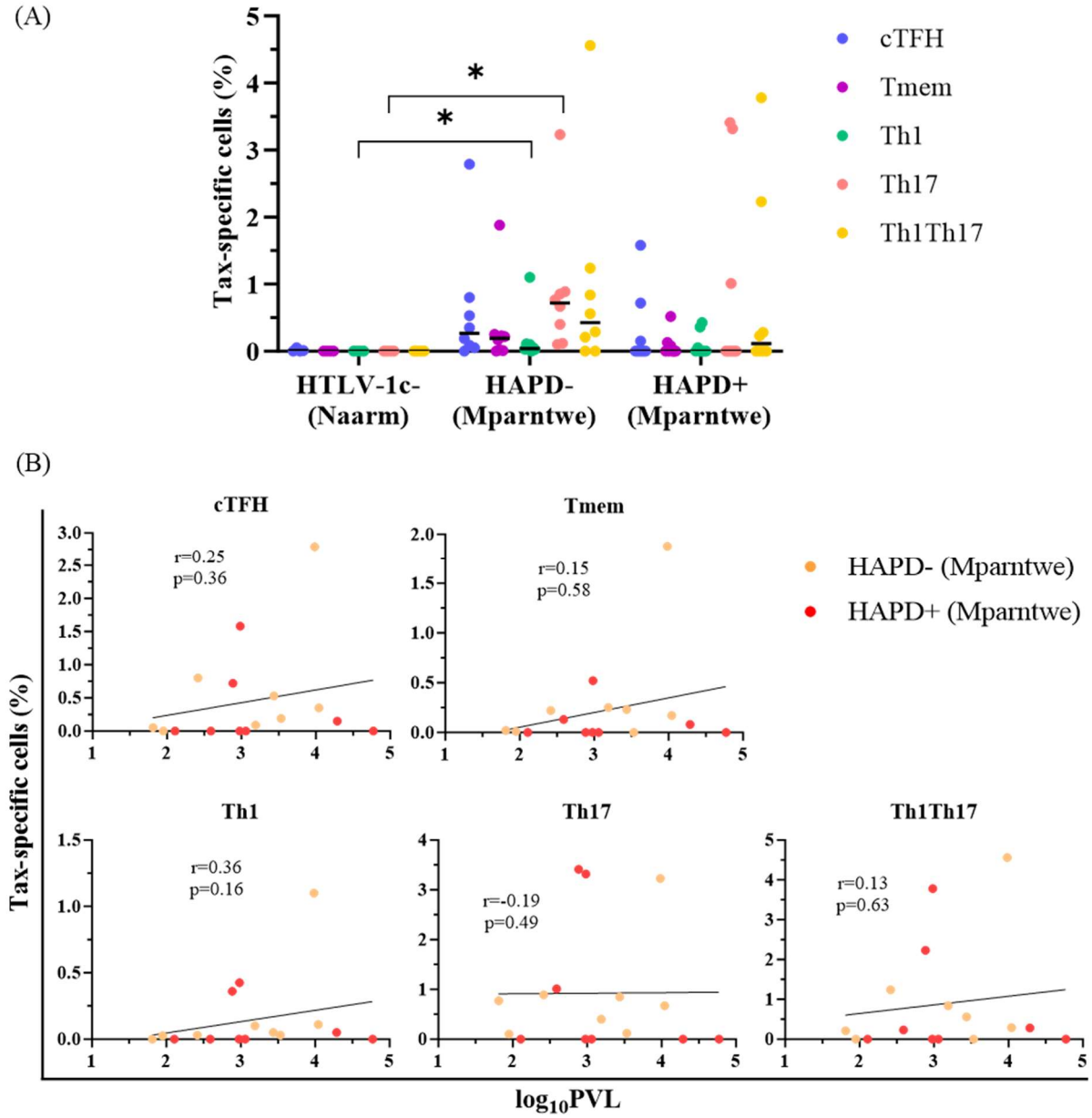


Figure 3.4 – Tax-specific CD4+ T cell responses are present in primary HTLV-1c infection. PBMCs were stimulated for 20 hours with DMSO or Tax peptide library prior to surface staining with a cocktail of antibodies and processing via flow cytometry. gDNA was extracted from PBMCs and PVL was determined by ddPCR. (A) Frequency of Tax-specific cells in CD4+ T cell phenotypes (DMSO background subtracted OX40+CD137+ frequency following stimulation with Tax peptides); (B) Correlation of Tax-specific CD4+ T cell phenotype frequencies and log_{10}PVL for HTLV-1c seropositive donors. n, 20: negative controls from Naarm (n, 4); HTLV-1c+ HAPD- from Mparntwe (n, 8); and HTLV-1c+ HAPD+ donors from Mparntwe (n, 8). Median frequency of each test group is shown by black line. Statistical significance between groups was assessed using Kruskal-Wallis test with two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli as an FDR control for multiple comparisons, adjusted p values (q values) shown. Black trend line represents nonlinear regression straight line of best fit. Correlation significance was assessed using Spearman's correlation test. * $p<0.05$.

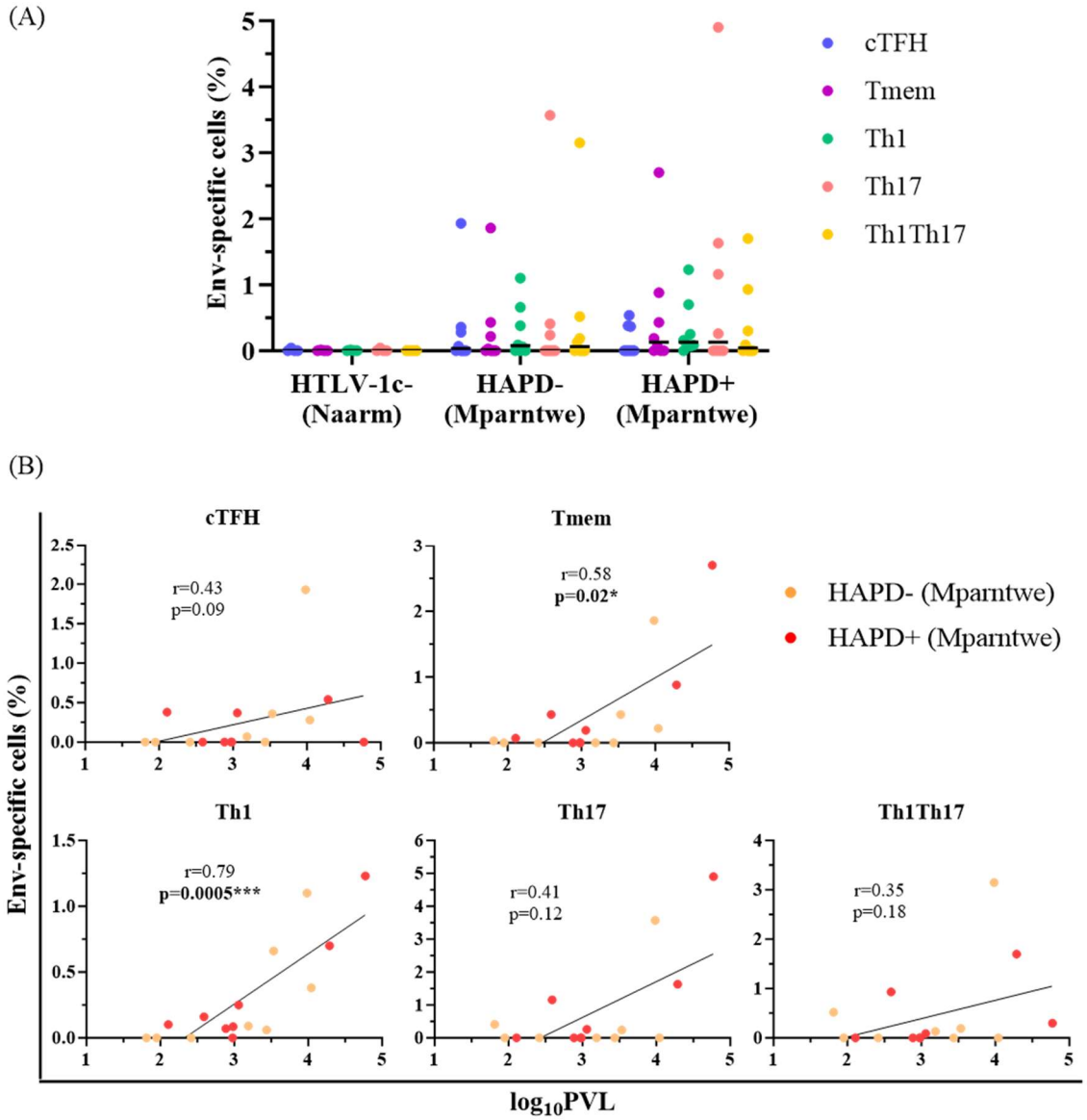


Figure 3.5 – CD4+ T cell Env-specific responses are present in primary HTLV-1c infection. PBMCs were stimulated for 20 hours with DMSO or Env peptide library prior to surface staining with a cocktail of antibodies and processing via flow cytometry. gDNA was extracted from PBMCs and PVL was determined by ddPCR. (A) Frequency of Env-specific cells in CD4+ T cell phenotypes (DMSO background subtracted OX40+CD137+ frequency following stimulation with Env peptides); (B) Correlation of Env-specific CD4+ T cell phenotype frequencies and log_{10}PVL for HTLV-1c seropositive donors. n, 20: negative controls from Naarm (n, 4); HTLV-1c+ HAPD- from Mparntwe (n, 8); and HTLV-1c+ HAPD+ donors from Mparntwe (n, 8). Median frequency of each test group is shown by black line. Statistical significance between groups was assessed using Kruskal-Wallis test with two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli as an FDR control for multiple comparisons, adjusted p values (q values) shown. Black trend line represents nonlinear regression straight line of best fit. Correlation significance was assessed using Spearman's correlation test. * $p<0.05$; **$p<0.01$; ***$p<0.001$

3.3.4 HTLV-1c infection elicits a robust CD8+ T cell response

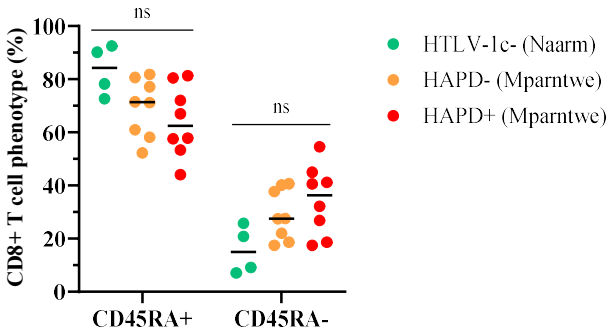
CD8+ T cells play an essential role in controlling viral infection by killing infected cells and producing cytokines [311]. To assess the CD8+ T cell responses following HTLV-1c infection, we first measured the unstimulated frequencies of circulating naïve (CD45RA+) and effector/memory (E/M) (CD45RA-) CD8+ T cells. The frequency of naïve CD8+ T cells was lower than uninfected Naarm-based controls (median, 84.25%; IQR, 74.10-91.93) in both HAPD- (median, 71.35%; IQR, 58.90-79.83) and HAPD+ (median, 62.45%; IQR, 54.45-78.38) donors, although no significance was observed (p, 0.14, 0.07, respectively) (**Figure 3.6A**). Between the HTLV-1c seropositive groups, there was no discernible difference in the naïve CD8+ T cell frequency (p, 0.50) (**Figure 3.6A**). Similarly, we observed higher frequencies of E/M CD8+ T cells in HAPD- (median, 27.55%; IQR, 19.53-39.60) and HAPD+ (median, 36.40%; IQR, 20.75-44.05) donors when compared to negative controls (median, 15.05%; IQR, 7.65-24.58), however these were not significant (p, 0.15, 0.07, respectively) (**Figure 3.6A**). Moreover, there were no differences in the proportion of E/M CD8+ T cells between the two HTLV-1c-infected groups (p, 0.44) (**Figure 3.6A).** Finally, the PVL of HTLV-1c seropositive donors did not correlate to the frequency of naïve (r, 0.29; p, 0.27) or E/M (r, -0.25; p, 0.36) CD8+ T cells (**Figure 3.6B**).

To assess the background activation of CD8+ T cells in HTLV-1c infection, we measured the proportion of CD69+CD137+ expressing cells in the DMSO treated condition. Like the ubiquitous activation of CD4+ T cells presented earlier in this chapter, significant background activation of CD8+ T cells was identified in both HAPD- (median, 0.340%; IQR, 0.168-1.568; p, 0.003) and HAPD+ (median, 0.28%; IQR, 0.092-1.910; p, 0.003) donors, when compared to uninfected controls (median, 0.006%; IQR, 0-0.018) (**Figure 3.6C**). Further consistent with CD4+ T cells, there was no distinguishable difference in the proportion of CD8+ T cell baseline activation between the two HTLV-1c seropositive groups (p, 0.35) (**Figure 3.6C**). However interestingly, the frequency of background CD8+ T cell activation was significantly positively correlated to PVL for all HTLV-1c-infected donors (r, 0.55; p, 0.03) (**Figure 3.6D**).

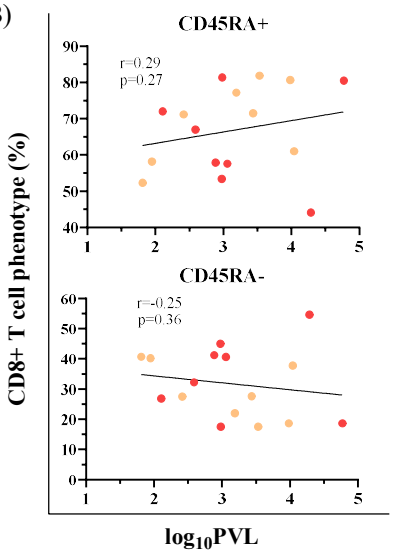
Further characterisation of the HTLV-1c-specific CD8+ T cell response was conducted by measuring the frequency of activation marker expression (CD69+CD137+) following

Tax or Env peptide stimulation, and background subtracting the frequency in the DMSO treated condition. The proportion of Env-specific CD8+ T cells was elevated in HAPD- (median, 0.104%; IQR, 0-0.330) and HAPD+ (median, 0.078%; IQR, 0.042-0.308) donors when compared to uninfected controls (median, 0.002%; IQR, 0-0.009), however no significance was observed (p, 0.22, 0.08, respectively) (**Figure 3.6E**). Moreover, no difference in the proportion of Env-CD8+ T cells was determined between the HTLV-1c-infected groups (p, 0.55) (**Figure 3.6E**). However, one HAPD- donor exhibited 1.14% of Env-specific cells from the total CD8+ T cell reservoir (**Figure 3.6E**). Overall, the level of Tax-specific CD8+ T cell responses were considerably higher than the Env-specific responses. Strikingly, HAPD- donors elicited an incredibly robust Tax-specific CD8+ T cell response (median, 0.995%; IQR, 0.443-2.825; p, 0.005) compared to uninfected control donors (median, 0.007%; IQR, 0.002-0.009) (**Figure 3.6E**). In fact, two HAPD- individuals displayed a frequency greater than 3% Tax-specific cells of their circulating CD8+ T cells. While the Tax-specific CD8+ T cell response was greater in HAPD- donors than HAPD+ donors, significance was not observed between the HTLV-1c seropositive groups (p, 0.09) (**Figure 3.6E**). Similarly, HAPD+ donors displayed elevated Tax-CD8+ T cell frequency (median, 0.152%; IQR, 0.078-1.553; p, 0.08) above controls (**Figure 3.6E**). Indeed, one Bex donor exhibited 2.27% of total CD8+ T cells as Tax-specific (**Figure 3.6E**). Finally, the PVL of HTLV-1c seropositive individuals was not correlated to the frequency of circulating Env- (r, 0.16; p, 0.55) or Tax-CD8+ T cells (r, -0.14; p, 0.61) (**Figure 3.6F**).

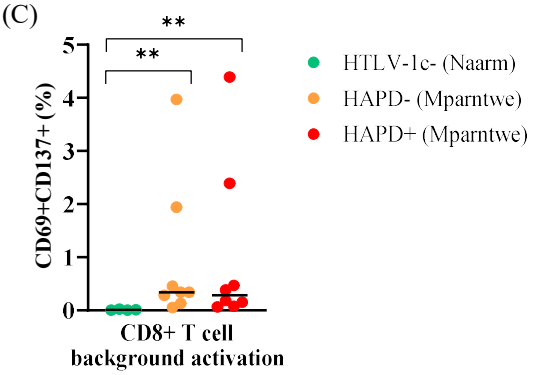
(A)


ns
ns
CD8+ T cell phenotype (%)
100
80
60
40
20
0
CD45RA+
CD45RA-
HTLV-1c- (Naarm)
HAPD- (Mparntwe)
HAPD+ (Mparntwe)

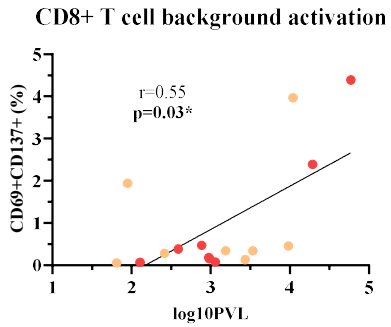
(B)


CD45RA+
r=0.29
p=0.27
CD45RA-
r=-0.25
p=0.36
CD8+ T cell phenotype (%)
log10PVL

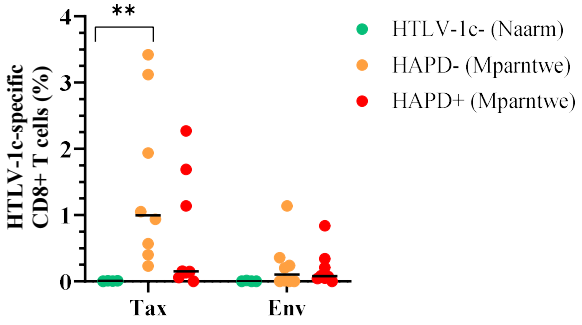
(C)


**
**
CD69+CD137+ (%)
CD8+ T cell
background activation
HTLV-1c- (Naarm)
HAPD- (Mparntwe)
HAPD+ (Mparntwe)

(D)


CD8+ T cell background activation
r=0.55
p=0.03*
CD69+CD137+ (%)
log10PVL

(E)


**
HTLV-1c-specific
CD8+ T cells (%)
Tax
Env
HTLV-1c- (Naarm)
HAPD- (Mparntwe)
HAPD+ (Mparntwe)

(F)

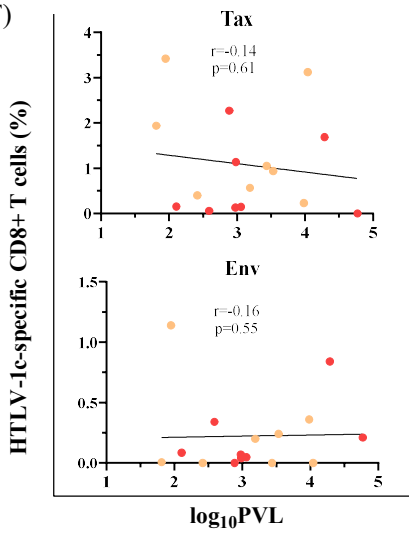

Tax
r=-0.14
p=0.61
Env
r=-0.16
p=0.55
HTLV-1c-specific CD8+ T cells (%)
log10PVL

Figure 3.6 – Significant CD8+ T cell response in primary HTLV-1c infection. PBMCs were stimulated for 20 hours with either DMSO, Env or Tax peptide libraries, or SEB, prior to surface staining with a cocktail of antibodies and processing via flow cytometry. gDNA was extracted from PBMCs and PVL was determined by ddPCR. (A) CD8+ T cell phenotype frequencies in DMSO treated condition. (B) Correlation of CD8+ T cell phenotype frequencies and log_{10}PVL for HTLV-1c seropositive donors. (C) CD8+ T cell activation frequencies (CD69+CD137+) in DMSO treated condition. (D) Correlation of CD8+ T cell background activation frequencies and log_{10}PVL for HTLV-1c+ donors. (E) Frequency of HTLV-1c-specific CD8+ T cells (background subtracted CD69+CD137+ frequency following stimulation with Tax and Env peptides). (F) Correlation of Tax- and Env-specific CD8+ T cell frequencies and log_{10}PVL for HTLV-1c+ donors. n, 20: negative controls from Naarm (n, 4); HTLV-1c+ HAPD- from Mparntwe (n, 8); and HTLV-1c+ HAPD+ donors from Mparntwe (n, 8). Median frequency of each test group is shown by black line. Statistical significance between groups was assessed using Kruskal-Wallis test with two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli as an FDR control for multiple comparisons, adjusted p values (q values) shown. Black trend line represents nonlinear regression straight line of best fit. Correlation significance was assessed using Spearman's correlation test. ns, not significant, $p>0.05$; * $p<0.05$; ** $p<0.01$

3.3.5 DN T cells are activated and virus-specific in HTLV-1c infection

Next, we evaluated the DN T cell response to HTLV-1c infection. Firstly, we examined the frequency of DN T cells in the DMSO treated condition. The total frequency of DN T cells was not significantly different when comparing HAPD- (median, 3.525%; IQR, 2.473-4.148; p, 0.39) and HAPD+ (median, 3.815%; IQR, 3.215-4.080; p, 0.29) donors to negative controls (median, 2.750%; IQR, 2.075-3.673), nor when comparing to each other (p, 0.54) (**Figure 3.7A**). Moreover, the DN T cell frequency of HTLV-1c infected donors did not correlate to PVL (r, -0.04; p, 0.90) (**Figure 3.7B**).

We measured the background activation of DN T cells by determining the frequency of CD69+CD137+ surface expression in the unstimulated DMSO condition. Consistent with the robust CD4+ and CD8+ T cell activation demonstrated earlier in this chapter, we detected elevated levels of baseline DN T cell activation in HTLV-1c infection. HAPD- (median, 0.270%; IQR, 0.110-2.098; p, 0.01) and HAPD+ (median, 0.086%; IQR, 0.030-0.410; p, 0.05) test groups had significantly higher frequencies of activated DN T cells than negative control donors (median, 0.0%; IQR, 0-0.023) (**Figure 3.7C**). Strikingly, one HAPD- donor exhibited a very high proportion of 6.7% of DN T cells expressing activation markers in the absence of ex vivo stimuli. (**Figure 3.7C**). However, no significant difference in the DN T cell baseline activation frequency was observed

between the two HTLV-1c+ groups (p, 0.21) (**Figure 3.7C**). Moreover, the PVL of HTLV-1c seropositive donors did not correlate to proportion of activated of DN T cells (r, 0.02; p, 0.94) (**Figure 3.7D**). We also examined activation markers OX-40+CD137+ in DN T cells (data not shown), however there was limited surface expression of these markers detected in this cell population.

We examined the HTLV-1c-specific DN T cell response by background subtracting the frequency of CD69+CD137+ expression in total DN T cells, after stimulation with either Env or Tax peptide libraries. The frequency of Env-specific DN T cells were significantly higher than negative controls (median, 0.007%; IQR, 0-0.018) in both HAPD- (median, 0.145%; IQR, 0.013-0.308; p, 0.01) and HAPD+ (median, 0.079%; IQR, 0.038-0.276; p, 0.01) donors, however there was no discernible difference between the HTLV-1c-infected groups (p, 0.33) (**Figure 3.7E**). Moreover, the PVL of seropositive individuals did not correlate to Env-DN T cell frequency (r, 0.24; p, 0.36) (**Figure 3.7F**). The proportion of Tax-specific DN T cells was significantly higher than uninfected controls (median, 0.0%; IQR, 0-0.007) in HAPD- donors (median, 0.125%; IQR, 0.069-0.263; p, 0.03) (**Figure 3.7E**). Meanwhile, HAPD+ donors tended to display higher Tax-DN T cell frequencies (median, 0.053%; IQR, 0.006-0.215; p, 0.07) than controls, but significance was not achieved (**Figure 3.7E**). Additionally, the Tax-specific proportion of DN T cells was not significantly different between the two HTLV-1c infected groups (p, 0.34) (**Figure 3.7E**), nor did the frequency correlate to PVL (r, 0.08; p, 0.78) (**Figure 3.7F**).

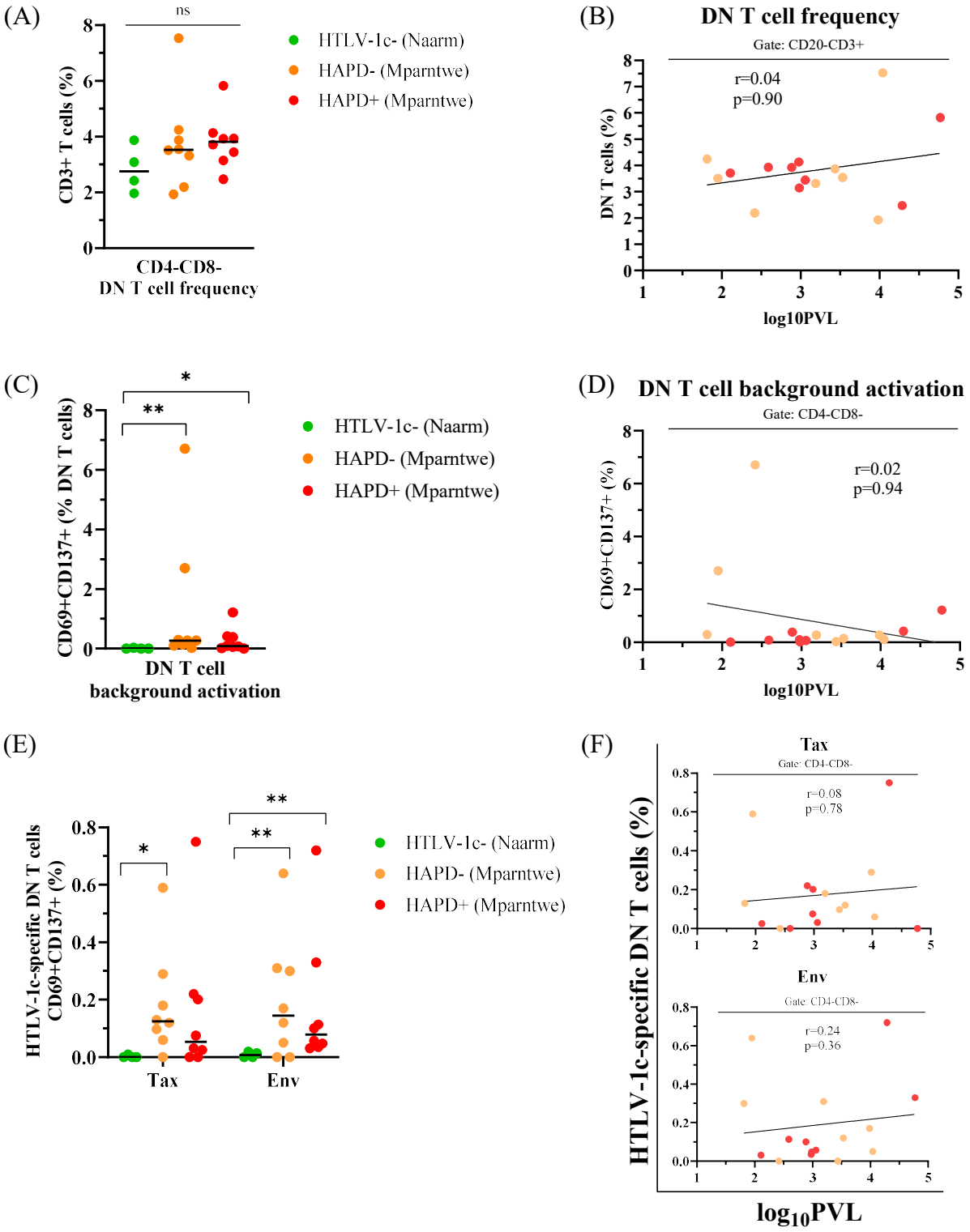

(A)
ns
CD3+ T cells (%)
CD4-CD8-
DN T cell frequency
HTLV-1c- (Naarm)
HAPD- (Mparntwe)
HAPD+ (Mparntwe)
(B)
DN T cell frequency
Gate: CD20-CD3+
r=0.04
p=0.90
DN T cells (%)
log10PVL
(C)
*
**
CD69+CD137+ (% DN T cells)
DN T cell
background activation
HTLV-1c- (Naarm)
HAPD- (Mparntwe)
HAPD+ (Mparntwe)
(D)
DN T cell background activation
Gate: CD4-CD8-
r=0.02
p=0.94
CD69+CD137+ (%)
log10PVL
(E)
*
**
**
HTLV-1c-specific DN T cells
CD69+CD137+ (%)
Tax
Env
HTLV-1c- (Naarm)
HAPD- (Mparntwe)
HAPD+ (Mparntwe)
(F)
Tax
Gate: CD4-CD8-
r=0.08
p=0.78
Env
Gate: CD4-CD8-
r=0.24
p=0.36
HTLV-1c-specific DN T cells (%)
log10PVL

Figure 3.7 – Double negative T cells are activated and HTLV-1c-specific in primary human infection. PBMCs were stimulated for 20 hours with either DMSO, Env or Tax peptide libraries, prior to surface staining with a cocktail of mAb and processing via flow cytometry. (A) DN (CD4-CD8-) T cell frequencies in DMSO treated condition. (B) Correlation of DN T cell frequencies and log_{10}PVL for HTLV-1c+ donors. (C) DN T cell activation frequencies (CD69+CD137+) in DMSO treated condition. (D) Correlation of DN T cell background activation frequencies and log_{10}PVL for HTLV-1c+ donors. (E) Frequency of HTLV-1c-specific DN T cells (background subtracted CD69+CD137+ frequency following stimulation with Tax and Env peptides). (F) Correlation of Tax- and Env-specific DN T cell frequencies and log_{10}PVL for HTLV-1c+ donors. n, 20: negative controls from Naarm (n, 4); HTLV-1c+ HAPD- from Mparntwe (n, 8); and HTLV-1c+ HAPD+ donors from Mparntwe (n, 8). Median frequency of each test group is shown by black line. Statistical significance between groups was assessed using Kruskal-Wallis test with two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli as an FDR control for multiple comparisons, adjusted p values (q values) shown. Black trend line represents nonlinear regression straight line of best fit. Correlation significance was assessed using Spearman's correlation test. ns, not significant, $p>0.05$; * $p<0.05$; ** $p<0.01$

3.4 Determining the cytokine response to HTLV-1c infection

With striking levels of T cell activation in HTLV-1c infection demonstrated, we next wanted to investigate the cytokine response generated by these immune cells. Participants for this study were recruited at the Alice Springs Hospital, Mparntwe, Australia, including 30 HTLV-1c seronegative and 44 HTLV-1c seropositive donors. A summary of each groups' demographics is listed in **Table 3.2**. To assess the host immune response to HTLV-1c infection, it was important to include a control group of HTLV-1c-uninfected First Nations donors from the same region as seropositive donors, given the high frequency of comorbidities which could also impact immune profile and function. We obtained community seronegative controls in this setting of matching plasma and PBMCs.

Plasma was obtained by centrifugation separation from whole blood samples, and frozen at -20°C prior to analysis. Upon thawing samples, plasma cytokine levels were measured for all 77 donors using a 9-plex bead-based assay. Analyte panel was chosen based on studies investigating the immune response in subtype-A infection and associated diseases [239]. Additionally, PBMCs were isolated by Ficoll centrifugation method from each

whole blood sample to match the plasma, and gDNA was extracted. The HTLV-1c PVL was determined using an established droplet digital PCR assay with primers and probe targeting the *tax* region, measured against the reference gene *RPP30* [218]. We transformed all cytokine and PVL data using $\log_{10}(x+1)$ to ensure the analysis was robust to outliers without losing any sample data, and for consistency between all uni- and multivariate analyses.

Table 3.2 – Cohort demographics and clinical characteristics of HTLV-1c host cytokine response study.

	HTLV-1c– Mparntwe (n, 33)	**HTLV-1c+ Mparntwe (n, 44)**	**p-value**
Demographics, n (%)			
Median Age	45.0	54.5	**0.005 ****
Female at birth	20/33 (61%)	29/44 (66%)	0.64
Male at birth	13/33 (39%)	15/44 (34%)	
Lifestyle, n (%)			
Smoker (current or previous)	25/33 (76%)	14/44 (32%)	**0.0002 *****
Harmful alcohol consumption	6/33 (18%)	16/44 (36%)	0.13
Comorbidities, n (%)			
Chronic Liver Disease	1/33 (3%)	8/44 (18%)	0.07
Diabetes	20/33 (61%)	29/44 (66%)	0.64
Chronic Kidney Disease	8/33 (24%)	28/44 (64%)	**0.001 ****
Malignancy	0/33 (0%)	5/44 (11%)	0.07
HTLV-1 associated diseases, n (%)			
HAPD (Bex)	0/33 (0%)	14/44 (32%)	**0.0002 *****
Strongyloidiasis, current	0/33 (0%)	12/44 (27%)	**0.0008 *****
BSI, current or previous	0/33 (0%)	18/44 (41%)	**<0.0001 ******
HAM	0/33 (0%)	0/44 (0%)	>0.99
ATL	0/33 (0%)	0/44 (0%)	>0.99

Participants consented to whole blood being taken and used for HTLV-1c pathogenesis research following written and verbal consent in primary language. Information was obtained through clinician and hospital records at Alice Springs Hospital, Mparntwe, NT, Australia. n, 77 (33 HTLV-1c- (Mparntwe), 44 HTLV-1c+ (Mparntwe)). Statistical significance was determined by using Mann-Whitney test for continuous variables, and Fisher's exact test for categorical variables. *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001.

3.4.1 Univariate analysis shows elevated plasma cytokines in HTLV-1c infection

To evaluate the cytokine response to HTLV-1c infection, we first analysed plasma cytokine levels in HTLV-1c+ and - donors. Our results show that HTLV-1c seropositive donors from Mparntwe had significantly elevated plasma concentrations of TNF-α (log_{10}(x+1) median, 1.267 pg/ml; IQR, 1.035-1.433; p, 0.01) when compared to seronegative controls log_{10}(x+1) median, 1.048 pg/ml; IQR, 0.939-1.191) (**Figure 3.8**). Furthermore, IL-10 plasma levels were significantly higher in HTLV-1c+ donors (log_{10}(x+1) median, 1.048 pg/ml; IQR, 0.823-1.204; p, 0.02) than negative controls (log_{10}(x+1) median; 0.783 pg/ml; IQR, 0.703-1.012) (**Figure 3.8**). Finally, IP-10 plasma concentration was significantly higher in HTLV-1c+ group (log_{10}(x+1) median, 2.715 pg/ml; IQR, 2.527-3.072; p, 0.02) when compared to uninfected controls (log_{10}(x+1) median, 2.539 pg/ml; IQR, 2.410-2.730) (**Figure 3.8**). On the other hand, our data did not indicate any significant differences between HTLV-1c seropositive and seronegative donors in all other cytokine analytes (IL-2, IL-4, IL-6, IL-8. IL-10, IFN-γ) (**Figure 3.8**).

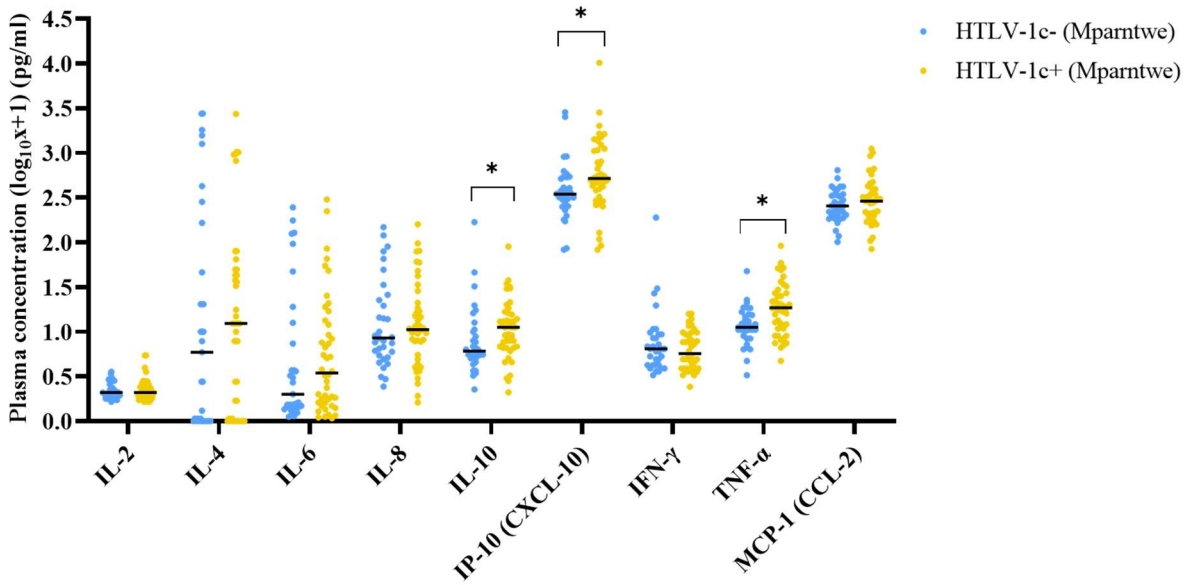


Figure 3.8 – Plasma levels of IL-10, IP-10 and TNF-α are elevated in primary HTLV-1c infection. Plasma cytokine levels were determined using a multiplex bead-based assay. Plasma cytokine concentrations (pg/ml) were log_{10}(x+1) transformed. HTLV-1c- donors from Mparntwe (n, 33) are shown in blue, HTLV-1c+ donors from Mparntwe (n, 44) are shown in yellow. Each dot represents an individual donor, with group median shown by black line. Statistical significance was assessed using multiple Mann-Whitney U test with two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli as an FDR control for multiple comparisons, adjusted p value (q value) is shown when *p<0.05.

3.4.2 Multivariate analyses expose dysregulated cytokine networks in HTLV-1c infection

Given the complexity of the host immune response and cytokines produced during a chronic viral infection, we then applied further statistical analyses and modelling of the host cytokine response to tease out risk factors and associations of HTLV-1c infection. Analyte levels were similarly transformed ($\log_{10}(x+1)$) so that variables approximated a more normal distribution, and so that analyses would be robust to outliers without needing to remove any samples from a small dataset. For variable selection (elastic net regression), decomposition (PCA) and visualisation (PLS-DA), data were further normalised by mean centring and variance scaling, as required by these methods.

We first examined the relationship between HTLV-1c infection and analyte levels on the entire study population. We began with a simple pairwise Pearson's correlation plot of analyte levels, which demonstrated moderate to high levels of correlation in two distinct clusters of cytokines; (1) TNF-α, IP-10, IL-10 and MCP-1; (2) IL-2, IL-4, IL-6, IL-8 and IFN-γ (**Figure 3.9A**).

While more common for very wide sets of data (100+ variables), we considered variable selection methods prior to further analysis given the correlation present in the dataset. We conducted 10-fold cross-validation (CV) using an elastic net logistic regression model. The optimal model ($C=10^{0.5}$, L1 ratio=0) produced non-zero coefficients for all variables (data not shown), so all analytes were kept for subsequent analysis.

We then performed a PCA, which aims to linearly transform the dataset into ordered principal components that explain the most variance in the dataset while being mutually orthonormal. This gives each sample new coordinates in an uncorrelated basis, and each input variable a loading for each principal component, but does not take into account the dependent variable in any way. Firstly, the PCA confirms that the dataset from all donors is highly correlated, as indicated by the fact that 82.9% of the variance of the initial 9 cytokine inputs can be explained with only the first 4 principal components (**Figure 3.9B**). The two distinct clusters of correlated analytes we identified in the pairwise correlation analysis ((1) TNF-α, IP-10, IL-10 and MCP-1; (2) IL-2, IL-4, IL-6, IL-8 and IFN-γ) were confirmed by their similar loadings on the first two principal

components (**Figure 3.9C**). When plotting the PC1/PC2 scores for each donor, those with high scores of both PC1 and PC2 are mostly HTLV-1c-infected (**Figure 3.9D**).

PLS-DA, another multivariate analysis method which linearly transforms independent variables into latent variables that maximise their covariance with a dependent variable, was applied to the cytokine dataset with the binary output of HTLV-1c infection. Latent variables were oriented such that their positive directions were associated with positive HTLV-1c status. There was a similar clustering of cytokine analytes for the two latent variables, when compared to the PCA (**Figure 3.9E**). TNF-α, IP-10, IL-10 and MCP-1 were assigned larger, positive loadings for LV1 and small loadings for LV2, while IL-2, IL-4, IL-6, IL-8 and IFN-γ were assigned larger, negative loadings for LV2. We observed reasonable separability between the 95% confidence regions of the uninfected and infected donor groups (**Figure 3.9F**).

Finally, a logistic regression was run for all donors with an output of HTLV-1c-infected or -uninfected, and controlling for age and sex AAB (**Table 3.3**). Of the controls, age was significantly positively associated with HTLV-1c infection (p, 0.03). Importantly, the concentrations of plasma TNF-α (p, 0.009) and IL-2 (p, 0.02) were significantly positively associated to HTLV-1c infection. Interestingly, IFN-γ (p, 0.005) and IL-8 (p, 0.007) plasma levels are significantly negatively associated with HTLV-1c infection. These results were largely consistent with the less rigorous findings from PLS-DA; TNF-α was significantly positively associated with HTLV-1c infection and was also from the cluster with a large positive loading on LV1 in PLS-DA, while IFN-γ and IL-8 were significantly negatively associated with HTLV-1c infection and from the cluster with a larger negative loading on LV2 in PLS-DA.

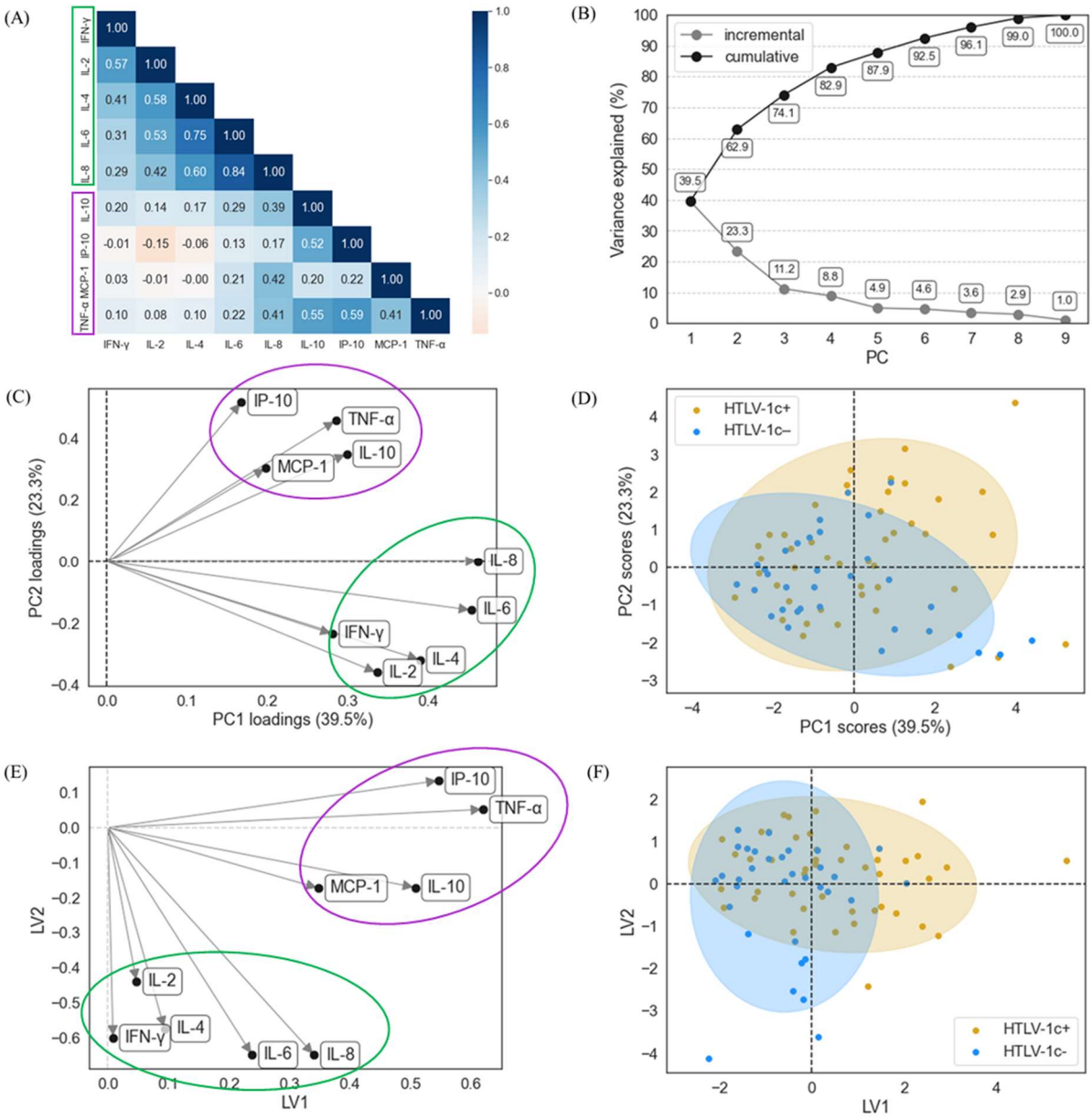


Figure 3.9 – Multivariate analyses for host cytokine profiles in HTLV-1c seropositive and seronegative individuals. Plasma cytokine levels were determined using a multiplex bead-based assay, and concentrations (pg/ml) were $\log_{10}(x+1)$ transformed. (A) Pearson's correlation matrix between each cytokine input. (B) PCA cumulative and incremental variance, with each PC denoted. (C) PCA input loadings for all cytokines. (D) PC1 and PC2 scores for each donor with 95% confidence regions of HTLV-1c-infected and uninfected donors. (E) PLS-DA LV input loadings for all cytokines. (F) PLS-DA LV1 and LV2 scores for each donor with 95% confidence regions of HTLV-1c-infected and uninfected donors. HTLV-1c- donors from Mparntwe (n, 33) are shown in blue, HTLV-1c+ donors from Mparntwe (n, 44) are shown in yellow.

Table 3.3 – Logistic regression analysis demonstrates that HTLV-1c infection disrupts the host cytokine network.

Logistic regression results
Dependent variable: HTLV-1c status
Number of observations: 77

Input	Coefficient	Std error	z	*p*	[0.025	0.975]
IL-2	14.62	6.05	2.42	**0.02** *	2.76	26.47
IL-4	0.44	0.44	0.99	0.32	-0.43	1.31
IL-6	1.60	1.06	1.50	0.13	-0.486	3.69
IL-8	-4.61	1.71	-2.70	**0.007****	-7.97	-1.26
IL-10	1.53	1.16	1.32	0.19	-0.74	3.80
IP-10 (CXCL-10)	0.73	1.17	0.62	0.53	-1.56	3.01
IFN-γ	-8.28	2.92	-2.83	**0.005****	-14.01	-2.55
TNF-α	4.77	1.82	2.62	**0.009****	1.20	8.34
MCP-1 (CCL-2)	2.59	1.93	1.34	0.18	-1.19	6.37
Age	0.05	0.03	2.16	**0.03***	0.01	0.10
Sex AAB (M)	0.01	0.70	0.02	0.99	-1.36	1.38

Logistic regression was performed for all HTLV-1c-infected and uninfected individuals with $\log_{10}(x+1)$ transformed plasma cytokine concentrations as independent variables and HTLV-1c infection status as dependent variable. Age and sex AAB were controlled for in the model. n, 77 (33 HTLV-1c- (Mparntwe), 44 HTLV-1c+ (Mparntwe)). *p<0.05; **p<0.01.

3.4.3 Univariate analyses of plasma cytokines does not reveal any trends for HTLV-1c-associated bronchiectasis

Next, we looked at the HTLV-1c infected cohort only, to see if there were distinct differences in the host plasma cytokine levels between HAPD- and HAPD+ test groups. The demographics and clinical characteristics of these two subgroups of HTLV-1c seropositive donors are summarised in **Table 3.4**. We included the PVL of these groups here, determined by ddPCR of matching PBMCs. Surprisingly, our data indicates there are no significant differences between the two groups across any of the cytokine analytes by univariate analysis (**Figure 3.10A**). Furthermore, none of the plasma cytokine concentrations measured correlated with PVL for the HTLV-1c infected donors (**Figure 3.10B**).

Table 3.4 – Demographics and clinical characteristics of HTLV-1c+ donors of host cytokine response study.

	HAPD- Mparntwe (n, 30)	HAPD+ Mparntwe (n, 14)	p-value
Demographics, n (%)			
Median Age	54.0	55.0	0.86
Female at birth	21/30 (70%)	8/14 (57%)	0.50
Male at birth	9/30 (30%)	6/14 (43%)	
Lifestyle, n (%)			
Smoker (current or previous)	11/30 (37%)	3/14 (21%)	0.49
Harmful alcohol consumption	10/30 (33%)	6/14 (43%)	0.74
Comorbidities, n (%)			
Chronic Liver Disease	8/30 (27%)	0/14 (0%)	**0.04***
Diabetes	21/30 (70%)	8/14 (57%)	0.50
Chronic Kidney Disease	23/30 (77%)	5/14 (36%)	**0.02***
Malignancy	1/30 (3%)	4/14 (29%)	**0.03***
HTLV-1 associated diseases, n (%)			
HAPD (Bex)	0/30 (0%)	14/14 (100%)	**0.0002*****
Strongyloidiasis, current	8/30 (27%)	4/14 (29%)	>0.99
BSI, current or previous	12/30 (40%)	6/14 (43%)	>0.99
HAM	0/30 (0%)	0/14 (0%)	>0.99
ATL	0/30 (0%)	0/14 (0%)	>0.99
HTLV-1 burden, log_{10}(copies per 10^6 cells)			
Median PVL	3.15	3.65	0.22

Participants consented to whole blood being taken and used for HTLV-1c pathogenesis research following written and verbal consent in primary language. Information was obtained through clinician and hospital records at Alice Springs Hospital, Mparntwe, NT, Australia. PVL was determined with gDNA extracted from PBMCs and quantified using ddPCR. n, 44 (30 HTLV-1c+ HAPD- (Mparntwe), 14 HTLV-1c+ HAPD+ (Mparntwe). Statistical significance was determined by using Mann-Whitney test for continuous variables, and Fisher's exact test for categorical variables. *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001.

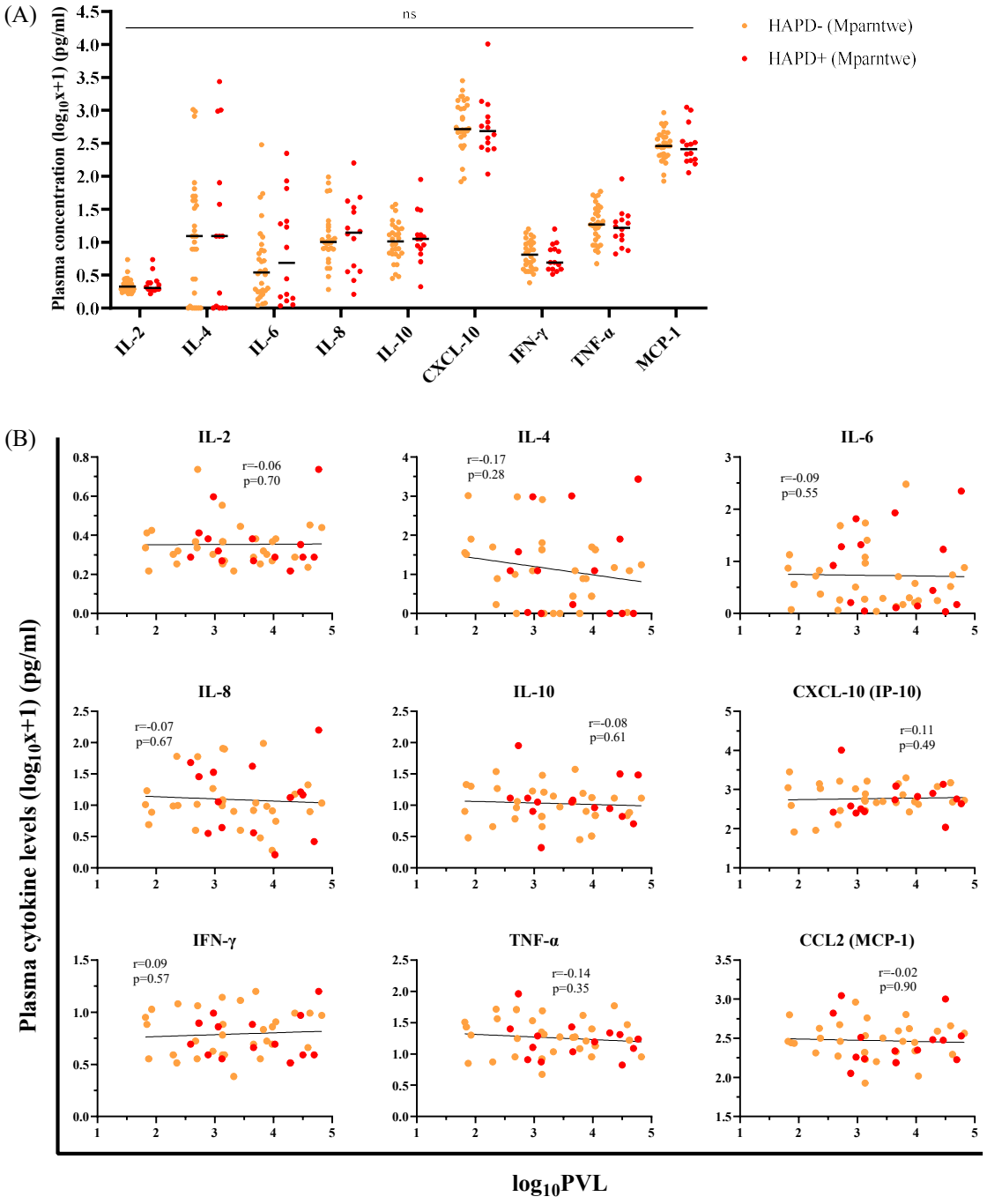


Figure 3.10 – Plasma cytokine levels in primary HTLV-1c infection. Plasma cytokine levels were determined using a multiplex bead-based assay and concentrations (pg/ml) were $\log_{10}(x+1)$ transformed. gDNA was extracted from matching PBMCs and PVL was determined by ddPCR. (A) Plasma cytokine levels in HTLV-1c infection. (B) Correlation of plasma cytokine concentrations and $\log_{10}$PVL for HTLV-1c seropositive donors. Each dot represents an individual donor, with group median shown by black line. HAPD- donors from Mparntwe (n, 30) are shown in orange, HAPD+ donors from Mparntwe (n, 14) are shown in red. Statistical significance was assessed using multiple Mann-Whitney U test with two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli as an FDR control for multiple comparisons. Black trend line represents nonlinear regression straight line of best fit. Correlation significance was assessed using Spearman's correlation test. ns, not significant, $p>0.05$.

3.4.4 Multivariate analyses reveal unique cytokine profile of HTLV-1c-associated bronchiectasis distinct from other HTLV-1a associated diseases

To better understand the complex nature of HTLV-1c infection and associated bronchiectasis, it was important to apply the same multivariate analyses strategy used previously in this chapter, particularly given the univariate analyses of the HTLV-1c seropositive cohort did not reveal any significant differences of plasma cytokine concentrations associated with the development of Bex. In this subgroup analyses, we included the similarly transformed ($\log_{10}(x+1)$) PVL of each donor to match the corresponding plasma cytokine analytes.

A pairwise Pearson's correlation plot demonstrated two distinct clusters of variables; (1) TNF-α, IP-10 and MCP-1; (2) IL-2, IL-4, IL-6 and IFN-γ. In this HTLV-1c+ group, IL-8 and IL-10 were moderately correlated to both clusters, while PVL was largely uncorrelated to any other analyte levels (**Figure 3.11A**).

We again considered variable selection with 10-fold CV using an elastic net logistic regression model. This time, there was minimal out-of-sample predictive power and therefore no optimal model was identified (data not shown); all variables were kept for subsequent analysis.

PCA confirmed the correlation structure of the data: 81.7% of the variance of all 10 inputs can be explained with the first 4 principal components (**Figure 3.11B**). Similar clustering as the correlation plot emerged; (1) TNF-α, CXCL-10 (IP-10) and CCL2 (MCP-1); (2) IL-2, IL-4, IL-6 and IFN-γ. Both clusters had similar PC1 loadings, while cluster (1) had a negative PC2 loading and cluster (2) had a positive PC2 loading, and IL-8 and IL-10 similarly lay between these two groups (**Figure 3.11C**). PVL had a very small loading on the first 2 PCs as it was uncorrelated to the two variable clusters that explained most of the variance of the dataset; its loading dominated the 3rd PC, but this could not be seen in a 2D visualisation of only the first 2 PCs. Overall, there was no obvious separability between HAPD- and HAPD+ donors when plotting their PC1/2 scores, indicated by the overlap of the 95% confidence regions (**Figure 3.11D**). This is expected in the absence of prior output-based variable selection.

During PLS-DA, the latent variables were oriented such that their positive directions were associated with HAPD. PVL notably had the most positive LV1 loading and the only positive LV2 loading. IL-8 and IL-10 were pushed towards cluster (2) involving IL-2, IL-4, IL-6 and IFN-γ, with positive LV1 loadings and negative LV2 loadings. Cluster (1) involving TNF-α, IP-10 and MCP-1 had small, negative loadings for both LV1 and LV2 (**Figure 3.11E**). When plotting each individual donors LV1/LV2 score, the PLS-DA shows some separability between the HAPD- and HAPD+ groups, as indicated by the 95% confidence regions (**Figure 3.11F**).

Finally, we ran a logistic regression with the output of HAPD for HTLV-1c-infected donors only, including all transformed cytokine analytes and PVL while controlling for age and sex AAB (**Table 3.5**). Interestingly, we determined that IL-8 (p, 0.04) and IFN-γ (p, 0.04) are significantly negatively associated with exhibiting HAPD in HTLV-1c infection. This mirrors the same negative associations that we found in the logistic regression of all infected and uninfected donors (**Table 3.3**). Moreover, our model indicated that IL-6 and IL-10 were significantly positively associated to HAPD in HTLV-1c infection (p, 0.03, 0.04, respectively). Additionally, our data shows that PVL was positively associated with HAPD among the HTLV-1c seropositive donors (p, 0.05). We determined the proportional risk was a 10-fold increase in PVL resulted in a 3.7-fold increase in the odds of exhibiting Bex. We re-ran the regression with $\log_{10}(x)$ for PVL and confirmed that both coefficients corresponded to the same odds ratio. Moreover, the results of the logistic regression were again reasonably consistent with PLS-DA despite the different purpose and underlying mathematics of the two methods. IL-6 and PVL had the two most positive loadings on LV1 and were positively associated with HAPD in the regression, while IFN-γ had the most negative loading on LV2 and was significantly negatively associated (**Figure 3.11E**).

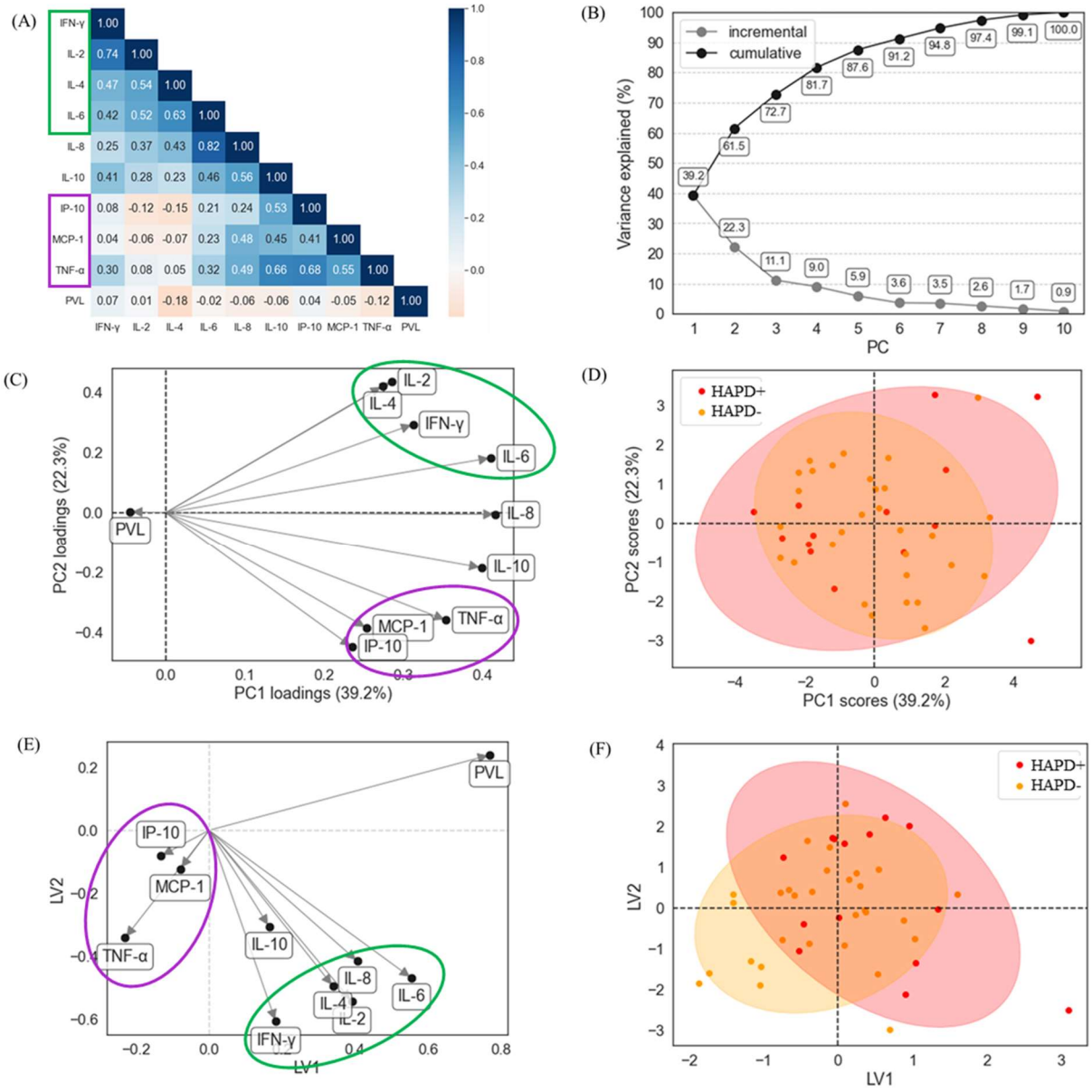


Figure 3.11 – Multivariate analyses for host cytokine profiles in HTLV-1c seropositive individuals and bronchiectasis. Plasma cytokine levels were determined using a multiplex bead-based assay, and concentrations (pg/ml) were $\log_{10}(x+1)$ transformed. gDNA was extracted from matching PBMCs and PVL was determined by ddPCR. (A) Pearson's correlation matrix between each cytokine input. (B) PCA cumulative and incremental variance, with each PC denoted. (C) PCA input loadings for all cytokines and PVL. (D) PC1 and PC2 scores for each donor with 95% confidence regions of HAPD- and HAPD+ patients. (E) PLS-DA LV input loadings for all cytokines and PVL. (F) PLS-DA LV1 and LV2 scores for each donor with 95% confidence regions of HAPD- and HAPD+ groups. HAPD- donors from Mparntwe (n, 30) are shown in orange, HAPD+ donors from Mparntwe (n, 14) are shown in red.

Table 3.5 – HTLV-1c-associated bronchiectasis shows a cytokine profile distinct from other HTLV-1a-associated diseases.

Logistic regression results
Dependent variable: HAPD
Number of observations: 44

Input	Coefficient	Std error	z	*p*	[0.025	0.975]
IL-2	11.12	7.49	1.48	0.14	-3.57	25.80
IL-4	-0.06	0.80	-0.07	0.95	-1.62	1.51
IL-6	4.80	2.22	2.16	**0.03***	0.45	9.15
IL-8	-6.01	2.88	-2.09	**0.04***	-11.66	-0.36
IL-10	5.25	2.55	2.06	**0.04***	0.25	10.25
IP-10 (CXCL-10)	-3.20	2.06	-1.56	0.12	-7.24	0.83
IFN-γ	-14.17	6.95	-2.04	**0.04***	-27.79	-0.54
TNF-α	0.21	2.81	0.07	0.94	-5.29	5.71
MCP-1 (CCL-2)	2.61	2.42	1.08	0.28	-2.12	7.35
PVL	1.30	0.67	1.93	**0.05***	-0.02	2.62
Age	0.04	0.03	1.28	0.20	-0.02	0.11
Sex AAB (M)	-0.11	0.94	-0.12	0.91	-1.95	1.73

Logistic regression was performed for all HTLV-1c-infected donors with $\log_{10}(x+1)$ transformed plasma cytokine concentrations and PVL as independent variables and HTLV-1c-associated Bex as the dependent variable. Age and sex AAB were controlled for in the model. n, 44 (30 HAPD- (Mparntwe), 14 HAPD+ (Mparntwe)). *p<0.05.

3.5 Chapter three discussion

Understanding the immune response to viral infections is a key element of the host-virus interplay. In the context of retroviruses and HTLV-1, where infection is lifelong, the immune response is a constantly evolving factor that determines potential disease outcomes and comorbidities versus the host efficiently managing infection. Host immune responses to HTLV-1a infection and associated diseases HAM and ATL have been well studied, but here we have uniquely characterised the host immune response to HTLV-1c infection and HAPD.

3.5.1 HTLV-1c causes high levels of immune activation and dysfunctional CD4+ T cell responses

HTLV-1a-infected cells can undergo clonal expansion, cause the proliferation of bystander cells, alter phenotype based on exposure to antigen, or result in the expansion of effector cells in response to infection. Therefore, we were interested in looking at CD4+ T cell frequencies in HTLV-1c infection of donors with and without HAPD. The increase in Tmem and cTFH frequencies in HTLV-1c-infected donors when compared to healthy donors, but no correlation to PVL, indicates that infection contributes to the expansion of these cell phenotypes but is not solely determined by the amount of provirus present, and may also be influenced by other viral and host factors. Expansion of cTFH cells during HTLV-1c infection may impact B cell differentiation and anti-HTLV-1c antibody production by the host. Expansion of the Tmem population could indicate that there is continual exposure to antigen during the chronic, lifelong HTLV-1c infection, and the provirus structure and viral genes that are expressed may play a role in determining this, something that will be explored in chapter four of this thesis. The total frequencies of the CD4+ T helper phenotypes (Th1, Th17, Th1Th17) were not significantly different across the groups of seropositive and seronegative donors, indicating that there was no HTLV-1c-associated expansion. On the other hand, the highly negative correlation between the total Th1 frequency and PVL suggests that HTLV-1c infection is interfering with Th1 differentiation or phenotype.

A profound finding of this study was the significant levels of background activation for every CD4+ T cell phenotype investigated for HTLV-1c infected donors. Chronic inflammation from this lifelong infection likely contributes to overall immune dysfunction resulting in increased susceptibility to comorbidities, tissue damage at sites of inflammation such as the lungs in Bex, and overall lower life expectancy [312]. Additionally, the positive correlation of both Tmem and Th1 phenotype background activation frequency with PVL indicates that HTLV-1c is causing overactivation and most likely dysfunction of these key CD4+ T cell phenotypes that otherwise would contribute to host control of infection, such as providing effective B cell and CD8+ T cell help, and triggering the production of, and the right balance of, effector cytokines.

Antigen-specific T cells are a major component of cell-mediated adaptive immunity. Traditionally, characterising antigen-specific T cells is determined by activation

following stimulation that is greater than three-fold above unstimulated background levels. However, we had to forgo this threshold due to the extensive levels of background activation of both CD4+ and CD8+ T cells present in the HTLV-1c seropositive donors in the absence of stimuli. This indicates just how striking the baseline levels of unstimulated activation are in HTLV-1c infected donors. It is important to note that not necessarily all HTLV-1-specific cells are infected with HTLV-1, although studies have shown that there is preferential infection of them [290]. Overall, we determined that there were elevated levels of HTLV-1c specific T cells in both HAPD- and HAPD+ groups when compared to negative controls. Interestingly, corresponding to the aforementioned background activation correlations, the frequency of both Env-specific Tmem and Th1 cells was significantly positively correlated to PVL for infected donors, further emphasizing the direct effect that HTLV-1c has on these cell phenotypes. Generally, of the HTLV-1c-specific CD4+ T cells, HAPD- donors showed more Tax-specific T cells, while HAPD+ donors had responses skewed towards Env, although more investigation into the T cell responses is required to confirm this. We hypothesize that differences in response to different peptides could be related to the provirus structure, and the viral mRNA and proteins that can be produced during chronic infection. Perhaps Bex donors have higher levels of intact provirus capable of producing infectious virions transmitted through cell contact and thus are exposed more to Env antigens, while HAPD- donors have higher levels of defective provirus with internal deletions, thus having higher responses to peptides of Tax which is encoded for mostly in the *pX* region and may be retained. This concept will be explored in the next chapter of this thesis which focuses on the proviral component of the host-virus relationship. Additionally, if Tax is expressed more than Env during in vivo infection, it might be suggested that HAPD- have more antigen-specific memory cells that can recognise and amount an immune response to the Tax-expressing cells, and thus keep the infection under control and limit extensive tissue damage and development of bronchiectasis. Alternatively, differences in pulmonary disease outcome may be due to the functionality of responding T cells related to other co-infections and co-morbidities – something which needs to be further investigated.

3.5.2 HTLV-1c elicits very strong CD8+ T cell responses

The strong CD8+ T cell response to HTLV-1c infection suggests that it is important in the host-virus relationship. Firstly, the increase in memory CD8+ T cells in HTLV-1c

infection may be due to the expansion of infected cells, similarly to subtype-A [313], and ongoing low levels of viral gene expression. We reported very high levels of baseline CD8+ T cell activation, and additionally this was correlated to the PVL, suggesting the presence of provirus, and potentially viral gene expression, is driving the high levels of CD8+ T cell activation. Tax was previously reported as the immunodominant peptide for CD8+ T cells in subtype-A [171], and this subtype-C study also reflects that. Interestingly, we see overall higher Tax-specific CD8+ T cells in HAPD- donors than HAPD+, which may allude to some level of effective CTL function in HAPD- donors that protects the host from developing Bex. Although, persistent lifelong infection of HTLV-1c suggests the CTL response is overall ineffective. On the other hand, HAPD- donors may have increased expression of Tax throughout infection, resulting in a more robust Tax-specific CD8+ T cell response than HAPD+ donors. Ultimately, there is still incredibly high levels of CD8+ T cell background activation of all HTLV-1c-infected donors, which likely contributes to overall chronic immune activation and inflammation.

3.5.3 DN T cells may be an untapped reservoir of HTLV-1c infection and inflammation

Finally, we examined the DN T cell response, for which their role and impact in HTLV-1 infection is relatively unknown. In HIV-1 infection, the frequency of DN T cells was negatively correlated with HIV-RNA plasma levels, suggesting that DN T cells may contribute to the control of infection [314]. While the DN T cell frequency did not correlate to PVL in our HTLV-1c study, we did not examine viral RNA levels, which are inherently low in HTLV-1a infection [44, 315]. In addition, DN T cells were demonstrated to be infected with HIV-1 in vivo [316], so these cells may be a source of persistent viral reservoir in both HIV-1 and HTLV-1c infection. Despite being such a rare T cell reservoir, DN T cells contribute to immune modulation and function during HIV-1 infection: they help regulate CD8+ T cell activation, and when activated will produce immunosuppressive cytokines IL-10 and TGF-β, in response to the chronic viral infection [314]. Our data indicated that circulating DN T cells were highly activated in HTLV-1c infection, and we also showed the presence of HTLV-1c-specific DN T cells. Further investigations are required to determine if these activated DN T cells regulate immune activation, or if being excessively activated themselves leads to dysfunctional

responses in HTLV-1c infection. The DN T cell reservoir may be an untapped contributor to the chronic inflammation that is seen across all HTLV-1c infected donors.

3.5.4 Future directions of the HTLV-1c-specific T cell response studies

Going forward, we suggest that additional uninfected control donors from the same region as HTLV-1c infected donors be included to further validate the findings of the HTLV-1c-specific T cell responses. Unfortunately, as the AIM assay requires a large number of cells ($1x10^6$ cells per condition) to test each donor sample across four conditions, our seronegative samples from Mparntwe did not have enough viable cells, so we had to source other negative controls from Naarm, Melbourne. While the HTLV-1c seropositive patients recruited in Mparntwe were First Nations people from remote community settings, the background of donors from Naarm were unknown. It is possible that these samples came from donors with different genetics and social determinants of health, meaning there could be differences in the overall efficiency and robustness of the effector functions of their immune cells. However it must be noted that these control samples from Naarm were indeed seronegative as per blood donation screening, confirmed by the absence of any HTLV-1c-specific cells in our study.

This chapter mainly focused on CD4+ T cell phenotypes given they are the predominant HTLV-1 reservoir in vivo, however the insights into the CD8+ T cell response to HTLV-1c infection uncovered by this study prompts further investigations. Future experiments should examine the activation levels of and the HTLV-1c-specific CD8+ T cell effector and memory subsets.

Furthermore, we should also assess the T cell responses to a wider variety of HTLV-1c peptide libraries, including those that may produce suboptimal responses, such as HBZ, as well as determining the immunodominant peptides. Additional studies should consider the expression of exhaustion markers such as programmed cell death protein 1 (PD-1), to help elucidate the effectiveness of the HTLV-1c-specific T cell response. Understanding these responses will guide the design and development of HTLV-1c (and perhaps even cross-clade) vaccines, to elicit effective T cell immunity. It is a fine balance between eliciting an immune response that is effective but not harmful: one that should be treated with caution and consideration.

3.5.5 Distinct and dysfunctional host cytokine response in HTLV-1c infection

As T cell dysregulation was apparent in HTLV-1c infected individuals earlier in this chapter, we hypothesized that the host cytokine response was also dysfunctional, as they are very tightly intertwined. Ideally, intracellular cytokine levels would be measured by flow cytometry from matching PBMC samples to the T cell response study, however we could not procure the appropriate volumes of samples required for this. Measuring the plasma cytokine levels provided an indirect measurement of what is being produced by these immune cells during HTLV-1c infection.

Prior to in-depth analysis, we examined the raw data of cytokine and PVL levels and determined it appropriate to transform the whole data set. A log transform was chosen over other outlier detection and removal methods given the small sample size of data, and the non-normality of the untransformed data even if outliers were removed. In order to perform a log transform, given that some cytokine analytes in various donors were at undetectable levels and designated a value of 0, we first right shifted all variable levels by 1 unit for consistency.

We determined via simple univariate analysis that HTLV-1c infected donors have elevated plasma levels of three cytokines; IL-10, IP-10 and TNF-α. TNF-α is a characteristic pro-inflammatory cytokine, involved in many physiological processes such as cell activation and induction of other inflammatory cytokines, and can indeed exacerbate pathogenesis [317]. IP-10 is known to activate and attract various immune cells by binding to CXCR3 [318]. On the other hand, IL-10 is known as an anti-inflammatory cytokine which helps regulate the host immune response and limit tissue damage [319]. Given the interconnectedness of the host cytokine response, we wanted to further explore the relationship between cytokine levels within donors by applying multivariate analyses.

Firstly, we determined our dataset was highly correlated, shown by the correlation matrix, along with the clusters of cytokine loadings and the substantial variance explained by the first few principal components. This was expected; in a completely uncorrelated dataset, each principal component would explain an equal proportion of the cumulative variance. Cytokines, meanwhile, are components of a complex immune network of cause and effect. Furthermore, given that PCA does not consider the

dependent variable, infection status, and we did not start with a wide dataset from which to select independent variables strongly associated with the dependent variable using variable selection methods prior to PCA, we did not expect substantially divergent PC scores between uninfected and HTLV-1c infected donors. Nonetheless, we did observe some natural separability; despite the arbitrary directions of the PCs relative to disease status, patients with both positive PC1 and PC2 scores tended to be HTLV-1c+.

While the elastic-net regression used for variable selection indicated that all independent variables should be kept for subsequent analysis, this did retain the multicollinearity in the dataset, which can impact the stability of results. Further analysis would benefit from a much wider set of independent variables from which to select, along with a larger set of donors. This may increase the likelihood that correlated variables are dropped during variable selection and therefore reduce the multicollinearity of the retained variables, while facilitating the exploration of other variable selection methods such as step-forward and step-backward regressions.

The PLS-DA for all donors, which does consider the dependent variable by maximising the covariance of independent and dependent variables when projecting onto a latent space, gave more separability between HTLV-1c uninfected and infected donors, with less overlap of the 95% confidence regions. Notably, the three cytokines given the highest loading for LV1 were TNF-α, IL-10 and IP-10, which coincides with the significant cytokines found in the previous univariate analyses from this cohort. Furthermore, TNF-α, and IL-10 were the only inputs given a positive loading for LV2.

We then applied a logistic regression to the dataset, a more finely tuned method which explicitly maximises the likelihood of the data given a linear association between a multivariate set of inputs and the log-odds of a binary dependent variable. This stands in contrast to the more visualisation-oriented PLS-DA, where maximising covariance is less well-suited to a binary outcome. The logistic regression determined the cytokine risk factors associated with HTLV-1c infection, taking into account all cytokines and also controlling for age and sex AAB. While some patterns are consistent with the rougher PLS-DA analysis, differences with the logistic regression may be explained by the known multicollinearity in the dataset, along with the very different objectives and optimisation functions of the two methods; consistency is by no means guaranteed between them.

Firstly, the positive association of HTLV-1c infection with age is likely due to the increasing chances of facing exposure events with age: breastfed as a baby, participation in various cultural practices, and cumulative sexual encounters throughout young adulthood and into adult life. Furthermore, with increasing age there could potentially be an increased chance of having comorbidities, either associated with HTLV-1c infection or unrelated. This could increase the number of health care visits for an individual, which may increase the likelihood of HTLV-1 screening and detection. TNF-α, which is positively associated with HTLV-1c infection, is a pro-inflammatory cytokine that helps regulate the immune response, however, can also be associated with pathogenesis in autoimmune diseases and infections if secreted excessively [319]. IL-2, another cytokine that is positively associated with HTLV-1c infection, supports the proliferation and survival of infected cells [320], and thus likely contributes to viral persistence. IFN-γ, the characteristic pro-inflammatory cytokine produced in HAM donors, is in fact negatively associated with HTLV-1c infection, much like ATL donors, and this may contribute to the dysfunctional CTL response [239]. IL-8, known to activate neutrophils to release chemoattractant that assist in the migration of T cells and monocytes to sites of inflammation [321], is negatively associated with HTLV-1c infection. With reduced circulating levels of this chemokine during infection, the host cannot direct an effective immune response against the virus to the sites it needs to. Taken together this shows that overall, in HTLV-1c infection regardless of the presence or absence of a known associated disease, there is dysregulation of the cytokine networks, and this may play a role in chronic inflammation, dysfunction in the host immune response and overall reduced life expectancy for all those infected with HTLV-1c.

3.5.6 Unique cytokine signature of HTLV-1c-associated bronchiectasis

For the HTLV-1c infected donors only, we did not detect any significant cytokine nor PVL differences between HAPD- and HAPD+ donor groups using simple univariate analysis. This indicates there may be a more complicated and subtle effect on the host immune response in the development of bronchiectasis, which is disrupting the finely tuned balance of cytokine levels in each individual donor. This could be especially relevant in resource poor settings such as remote communities in Central Australia, where overall health inequity may give rise to a plethora of other concurrent health challenges, making the distinct effect of HTLV-1c infection more complex to tease apart.

The lack of out-of-sample predictive power during elastic-net variable selection for HTLV-1c infected donors suggests that this study could be expanded to improve the generalisability of our modelling. Firstly, we had a relatively lower number of donors (n, 14) that we had matching plasma and PBMCs for who were infected with HTLV-1c and radiologically diagnosed with bronchiectasis. Ideally, we would also like to have a cohort larger than 77 participants, however we proceeded with the study at this size given the difficulties and limitations with procuring donor samples with HTLV-1c in Central Australia. Parallel to the all-donor analysis, our HTLV-1c-infected donor analyses may benefit from a much wider set of input biomarkers, along with more complex modelling methods such as introducing polynomial terms or tree-based regressions, to better generalise how the complex immune response in the development of Bex may affect cytokine levels in a wider population.

Consistent with the total cohort, correlation plots and PCA again showed that this HTLV-1c+ subgroup dataset is moderate-to-highly correlated. However, unlike the all-donor PCA which incidentally showed some separation between disease status, the full overlap of 95% confidence regions for HAPD- and HAPD+ PC1/PC2 scores is more typical of what we would expect when we did not start with hundreds of variables and preselect only those which give the best separation between outcomes.

Consistent with the total cohort, the PLS-DA for the HTLV-1c+ only population showed some visual differences between the distributions of HAPD- and HAPD+ donors on the latent variables, which prompted us to explore further with the logistic regression. We determined several risk factors associated with Bex development in HTLV-1c infection. Previously it has been shown that PVL is higher in Bex donors than ACs [9], and this study further concurs that PVL is positively associated with HAPD. However, we also determined the proportional risk of PVL to HAPD, which may assist in monitoring HTLV-1c infected individuals for disease progression. While right-shifting by 1 unit prior to a $\log_{10}$ transform during data normalisation can distort the interpretability of odds ratios from logistic regression, especially for variables with minima close to 0, the range of PVL (copies per 10^6 cells) was [64.8, 67070.2], giving us confidence that odds ratio (OR) could be interpreted directly. This was further confirmed by removing the right-shift only for PVL before re-running the logistic regression; the OR was effectively identical. IL-6, associated with a multitude of cell types, has both pro- and anti-inflammatory effects depending on the context of the host state [322], and was

positively associated with exhibiting Bex in HTLV-1c infection. IL-6 is produced due to tissue damage during viral infections [323], so is likely upregulated during inflammation and damage to the bronchial tubes in the lungs during HTLV-1c-associated Bex. IL-6 decreases AICD [324], thus is promoting survival of chronically activated HTLV-1c-infected and bystander cells. Furthermore, IL-6 has been shown to impair the effector CD8+ T cell response post-activation [325], which could indicate that despite high levels of CD8+ T cell activation and HTLV-1c-specific CD8+ T cells, they are ineffective in mounting an attack on infected cells. IL-6 upregulates migration of activated and antigen-specific T cells in vitro, thus in Bex development there may be activated and infected T cells, of which we previously showed there is a plethora of, that migrate to the lungs and contribute to pathogenesis [326]. We also determined that IL-10, a pleiotropic chemokine broadly involved in downregulating the host immune response and contributing to the establishment and maintenance of chronic viral infections [327], was positively correlated to HAPD. HIV-1 studies show that IL-10 inhibits the function of virus-specific T cells [328], so even though HTLV-1c-infected donors display Env- and Tax-specific T cells, they may not be effective in controlling the infection. Furthermore, IL-10 impairs the functions of monocytes and macrophages through inhibiting expression of MHC class II and costimulatory molecules, and thus downregulates the production of pro-inflammatory cytokines and chemokines [329]. Interestingly, both IL-8 and IFN-γ were significantly negatively associated with HAPD in HTLV-1c infection, similarly to the total cohort analyses that determined the correlates with seropositivity as previously mentioned. Taken together, it seems that HTLV-1c-associated Bex is in fact associated with some immunosuppressive cytokines, more akin to the ATL signature than HAM, summarised in **Table 3.6**. This suggests that the chronic T cell activation characterised in this chapter leads to immune exhaustion, facilitating the progression of HAPD. Nonetheless, correlation does not necessarily equate to causality, hence further investigation will be required to confirm the link between cytokine signatures and HTLV-1c-associated Bex.

Table 3.6 – Summary of plasma cytokine networks in HTLV-1-associated diseases.

HTLV-1-associated disease	Plasma cytokines	
	Upregulated	*Downregulated*
ATL (subtype-A) [239]	IL-10; TGF-β; IL-6	IFN-γ
HAM (subtype-A)[239]	IFN-γ; TNF-α; IL-4; IL-6; IL-8; IP-10; CXCL9	MCP-1
HAPD (subtype-C)	IL-6; IL-10	IL-8; IFN-γ

3.5.7 Dysfunctional Th1 response in HTLV-1c infection

Arguably one of the most significant findings of this study which combines both the HTLV-1c-specific T cells and the cytokine network, was the dysregulation of the Th1 response in HTLV-1c infection. Th1 cells are involved in triggering the pathway to signal virus-specific CD8+ T cells to differentiate into effector CTLs which can result in the killing of virus-infected cells. With lower frequency of Th1 cells in HTLV-1c infection which are highly activated and may have reduced function, the HTLV-1c-specific CD8+ T cells may be unable to differentiate into effector CTLs efficient at killing the persistent viral infection. The negative correlation of PVL to frequency of Th1 cells suggests that HTLV-1c infection may be interfering with Th1 differentiation. Furthermore, the strong correlation of PVL to the baseline level of Th1 activation, as well as the correlation of PVL to the frequency of Env-specific Th1 cells, shows that HTLV-1c is driving this dysregulation. Additionally, IL-6 is known to inhibit Th1 polarization by inhibiting the secretion of IFN-γ by CD4+ T cells [330], just as IL-10 can downregulate the expression of Th1 cytokines [331], both of which are significantly associated with HAPD in HTLV-1c-infected individuals. Indeed, we see this play out: IFN-γ levels are significantly negatively associated with HTLV-1c infection, and also with Bex disease progression.

3.5.8 Chapter three conclusion

In summary, our results indicate that HTLV-1c infection instigates a unique host immune response in terms of both the T cell response and the cytokine networks and may be due to a combination of both viral and host factors. We showed that subtype-C infection and bronchiectasis are associated with several immunosuppressive cytokines, present in conjunction with inflammatory cytokines, extensive T cell activation, and a dysregulated Th1 response. When considered together, HTLV-1c infection triggers a

complicated combination of immune activation and inflammation, alongside dysfunctional and exhaustive immune responses, that ultimately contribute to the higher all-cause mortality in HTLV-1c infection, and the development of HAPD and bronchiectasis.

4. Characterising the phenotype of HTLV-1c infected cells and the proviral structure within these cells

4.1 Chapter four background

The last chapter examined the host immune response to HTLV-1c infection, including immune factors, such as the T cell response and the cytokine networks, that influence disease progression. The chronic immune activation, together with immune dysregulation and dysfunction were significantly linked to HTLV-1c and bronchiectasis. This chapter focuses on which immune cells carry HTLV-1c, what is the nature of the proviral genome within these cells, and how this may impact pathogenesis and disease.

4.1.1 HTLV-1 exhibits CD4+ T cell tropism

In vivo, HTLV-1 displays tropism for CD4+ T lymphocytes [29], although it can also infect CD8+ T cells [34, 35], DCs [36], macrophages and monocytes [37]. The ability of HTLV-1 to infect a plethora of immune cells stems from the ubiquitous nature of its cellular receptors: GLUT-1 [33], NRP-1 [32] and HSPs [30]. Nevertheless, characterisation of the immunophenotype of HTLV-1c infected cells and the molecular virology of the provirus will inform understanding of disease pathogenesis, and must start with the expected main reservoir: the CD4+ T cells.

Soon after the discovery of the virus, the cellular tropism of HTLV-1 and surface protein expression of HTLV-1-infected cells has been deeply investigated for the subtype-A infection. CD25 is highly expressed on the surface of HTLV-1 infected CD4+ T cells [332], such that administration of a humanised mAb anti-CD25 was among the first targeted treatments trialled with ATL patients [112]. ICAM-1 and lymphocyte function-associated antigen 3 (LFA-3), both adhesion molecules, are upregulated on the surface of HTLV-1a infected T cells, which assists in cell-to-cell transmission of the virus [333, 334]. Furthermore, cell adhesion molecule 1 (CADM1) is similarly highly expressed on CD4+ T cells harbouring provirus [335-337]. In addition, the frequency of CADM1+CD4+ T cells is correlated to PVL and may influence the susceptibility of infected cells to CTL-mediated lysis, as well as predict the severity of ATL progression [336-339]. Most recently, CC chemokine receptor 4 expressing (CCR4+) CD4+ T cells were identified as a reservoir highly infected with HTLV-1a in both ATL and HAM donors [340-343]. Furthermore, the virus will preferentially continue to transmit to, and functionally alter, these cells [340-343]. Indeed, phase I and II clinical trials have been undertaken with ATL and HAM cohorts where participants were administered an anti-

CCR4 mAb, mogamulizumab [110, 111, 344, 345]. However, the outcomes for participants were varied: some patients demonstrated a PVL reduction, some experienced temporary remission, while for others the proviral burden wasn't effectively minimised.

It is important to identify consistent cellular and serum biomarkers of HTLV-1 infection, to efficiently monitor disease progression and design therapeutic treatments that effectively target infected cells, while limiting toxicities for other non-infected immune cells and negative side effects on the patient. This makes understanding the immunophenotype of HTLV-1c infected cells and associated pulmonary disease a priority that is yet to be resolved.

4.1.2 HTLV-1-infected cells are found in tissue reservoirs

During viral infections, specific cytokine and chemokine signalling directs the migration of host immune cells to sites of inflammation. While this is usually a beneficial host mechanism to combat infection and repair damage, for an immune targeting virus like HTLV-1c, it can also heighten pathogenesis and tissue damage through viral gene expression by infected cells and a positive feedback loop of chronic inflammation [346]. As such, infiltration of HTLV-1a infected cells into a range of tissues likely contributes to the development and progression of HTLV-1-associated diseases. HTLV-1a infected CD4+ T cells have been detected in the CNS of HAM patients [347], with transmigration facilitated through disruption of the blood brain barrier [348]. Furthermore, the development of HTLV-1-associated Sjogren's autoimmune syndrome in HAM donors was associated with CD4+ T cell infiltration into tissues implicated in disease such as the salivary glands, and HTLV-1 DNA was detected in these tissue reservoirs [349, 350]. Similarly, the infiltration of T lymphocytes into the lung tissues and BALF, alongside the presence of HTLV-1 DNA in the respiratory exudates, has been reported during infection [131, 218, 351, 352] and may contribute to the various lung tissue abnormalities and lung diseases associated with HTLV-1. In HTLV-1-associated uveitis, non-malignant, HTLV-1-infected T cells infiltrate into the eye [141, 142]. Together, these studies show the implications of HTLV-1 infected T cell migration on pathogenesis and disease development at specific tissue homing sites.

4.1.3 Defective provirus is present in HTLV-1a infection

An integral component of all host-virus interplay is viral gene expression. Viral products, *mRNA* or proteins, can exert pathogenic effects on the host as directly cytotoxic themselves, or through modulating host cell function. Gene expression of a retrovirus, such as HTLV-1, is dependent upon a wide variety of factors: including the sequence of the provirus and the presence of mutations, insertions or deletions which may affect the gene coding regions; proviral integration site and interactions with surrounding host genes, various epigenetic modifications such as 5'-LTR CpG methylation [173] and pre- and post- transcriptional regulation by viral and host proteins and functional mRNA. The studies in this thesis chapter focus on defining which immune cells are directly infected and how the proviral sequences play their role in promoting the persistence and pathogenesis of HTLV-1.

Shortly after the discovery of HTLV-1, defective HTLV-1a provirus was frequently documented using a variety of molecular and genetic sequencing techniques. Early studies using ATL derived T cell lines and primary ATL cells detected the presence of defective clones that were missing the 5'LTR region, or internal *gag* and *pol* regions [153, 272, 273]. In fact, it was reported that 56% of ATL donors had defective proviruses, and of these, 43% of the proviruses retained the LTRs but had internal deletions [153]. Furthermore, some of these defective clones were shown to still retain the negative strand *hbz* gene with capacity to express HBZ protein [154]. Moreover, defective proviruses detected in HTLV-1 seronegative HAM patients demonstrated the ability of the virus to avoid immune system detection, while still exerting pathogenic effects and associated disease development [353]. Southern blot hybridisation and restriction digestion determined that over 50% of individuals in an ATL donor cohort had predominantly intact provirus [354]. Characterisation of proviral sequences with a short-read next generation sequencing (NGS) capture-seq approach suggested that for all HTLV-1a infected persons, including ATL, HAM and AC donors, over 75% of proviruses were intact [152]. The variability in the overall HTLV-1a proviral landscape reported in the literature is likely attributable to the different methodologies applied.

The numerous proviral sequencing studies of HTLV-1a, using a variety of methods each with their own strengths and limitations, highlighted various types of defective provirus which contribute to viral pathogenesis and persistence. However, the nature of

individual provirus sequences, deletions and breakpoints are not well characterised in these studies. Furthermore, minimal investigation has been undertaken to understand the proviral landscape of HTLV-1c infection and how it may contribute to the development of HTLV-1c-associated bronchiectasis and overall lower life expectancy documented for all infected persons [14].

4.2 Chapter four aims

This study aims to characterise the phenotype and provirus structure of HTLV-1c infected immune cells. We hypothesize that HTLV-1c infected individuals have a higher frequency than uninfected individuals of CD4+ T cells in peripheral circulation bearing markers associated with the capacity to traffic to the respiratory tract, and that donors exhibiting Bex show further expansion of this phenotype. Additionally, we hypothesize that these phenotypic lung homing CD4+ T cells are highly infected with HTLV-1c and may contribute to pathogenesis and respiratory tract damage. Given the nature of extensive T cell activation elucidated in chapter three of this thesis, we hypothesize that CD4+ T cells trafficking to the respiratory tract are similarly activated, and contribute to the dysfunctional host immune response, chronic inflammation, and tissue damage. As defective proviruses are found in both in HIV-1 and HTLV-1a infection, we hypothesize that HTLV-1c infection accumulates defective provirus. Furthermore, we hypothesize that among these defective proviral clones, there will be preferential retention of regulatory and accessory genes, such as *tax* and *hbz,* which may facilitate persistence, proliferation, and pathogenesis, as opposed to the genes for the viral structural proteins *gag, pol* and *env*. Viral gene expression can be modulated through the provirus structure and the retention or absence of genomic regions essential for transcription and translation of genes, which may have implications on the host-virus relationship and disease progression. Accordingly, we hypothesize that there is a difference in the overall provirus structure of HAPD- and HAPD+ donors.

4.3 Characterising the phenotype of HTLV-1c infected cells

4.3.1 Strategy for infected cell phenotype characterisation was structured around CCR4 marker

To identify the phenotype of HTLV-1 infected immune cells, hereafter called the immunophenotype of HTLV-1c positive cells, we designed a fluorescence panel and gating strategy (**Figure 4.1**) structured around the expression of surface marker CCR4+, as these cells are highly infected with HTLV-1a [340-343]. Moreover, given HTLV-1a infection is characterised by an expansion of FOXP3+CD4+ T cells in ATL, HAM and high PVL ACs [355-357], we were interested in investigating this subset of T cells. FOXP3 is a transcription factor that is often expressed by Tregs [358, 359], although further functional studies to determine their suppressive nature are required for true classification. Staining of internal FOXP3 for FACS requires permeabilization of cells, and the fixing process often results in the loss of a proportion of cells. With the precious nature and small volume of PBMC samples from Central Australia, and downstream analysis of high throughput gDNA sequencing to examine the provirus structure, we could not afford to lose any cells during the fixing process. We therefore targeted this CD4+ phenotype using the previously characterised FOXP3+ proxy phenotype surface markers, CD49d-CD127- [360]. Furthermore, given Bex is a particularly prevalent associated disease of HTLV-1c infection in Australia, we were interested to explore the trafficking of infected cells to the respiratory tract, and targeted this phenotype with CCR4+CD49d+Integrin-β7-. The rationale behind this was that CCR4+CD49d- T cells preferentially traffic to the skin, and CD49d+Integrin-β7+ cells preferentially traffic to the gut [361, 362], so the CD49d+Integrin-β7- phenotype likely contains cells that will traffic to the respiratory tract and lungs, especially if this phenotype also expresses the CCR4+ marker [363-365] that was previously documented as high expressed on many HTLV-1 infected cells [340-343]. Finally, we included activation markers HLA-DR+CD38+ into the phenotyping panel, to further validate the extensive activation and chronic inflammation demonstrated in chapter three, but with a different approach and cohort of donor samples.

Cryopreserved PBMCs from HTLV-1c infected and uninfected First Nations donors from Mparntwe (Alice Springs), Central Australia, as well as uninfected controls from

Naarm (Melbourne), were thawed and stained using our 10-colour panel and processed via FACS (**Figure 4.1**). Like the cytokine profiling in chapter three of this thesis, we included uninfected donor controls from the same region and genetic background as our infected donors, to confidently ascertain the differences of immune cell frequencies between those with HTLV-1c and those testing seronegative. These community matched controls should account for the higher levels of other acute and chronic comorbidity conditions in First Nations people in Central Australia and reduced life expectancy, when compared to non-First Nations individuals. Demographic and clinical details of community control and HTLV-1c infected groups used for immunophenotyping study are detailed in **Table 4.1**. Additional negative control blood samples from Naarm, Melbourne, were obtained through Red Cross Lifeblood, but ethics prevented the distribution of clinical or demographic information of these donors for our research purposes.

Given the under resourced staff and facilities in Alice Springs Hospital, where the collection of whole blood and isolation of PBMCs occurred, the quality of cryopreserved samples was highly variable and sometimes quite poor. For the analysis of samples, when gating from single cell lymphocytes, we had to set a minimum threshold of 2% live CD3+ T cells, and we disregarded any samples that were below this quality. Furthermore, any samples with gates that contained less than 50 cells were excluded from analysis, as they may not accurately represent the phenotype frequencies in peripheral blood of the donor.

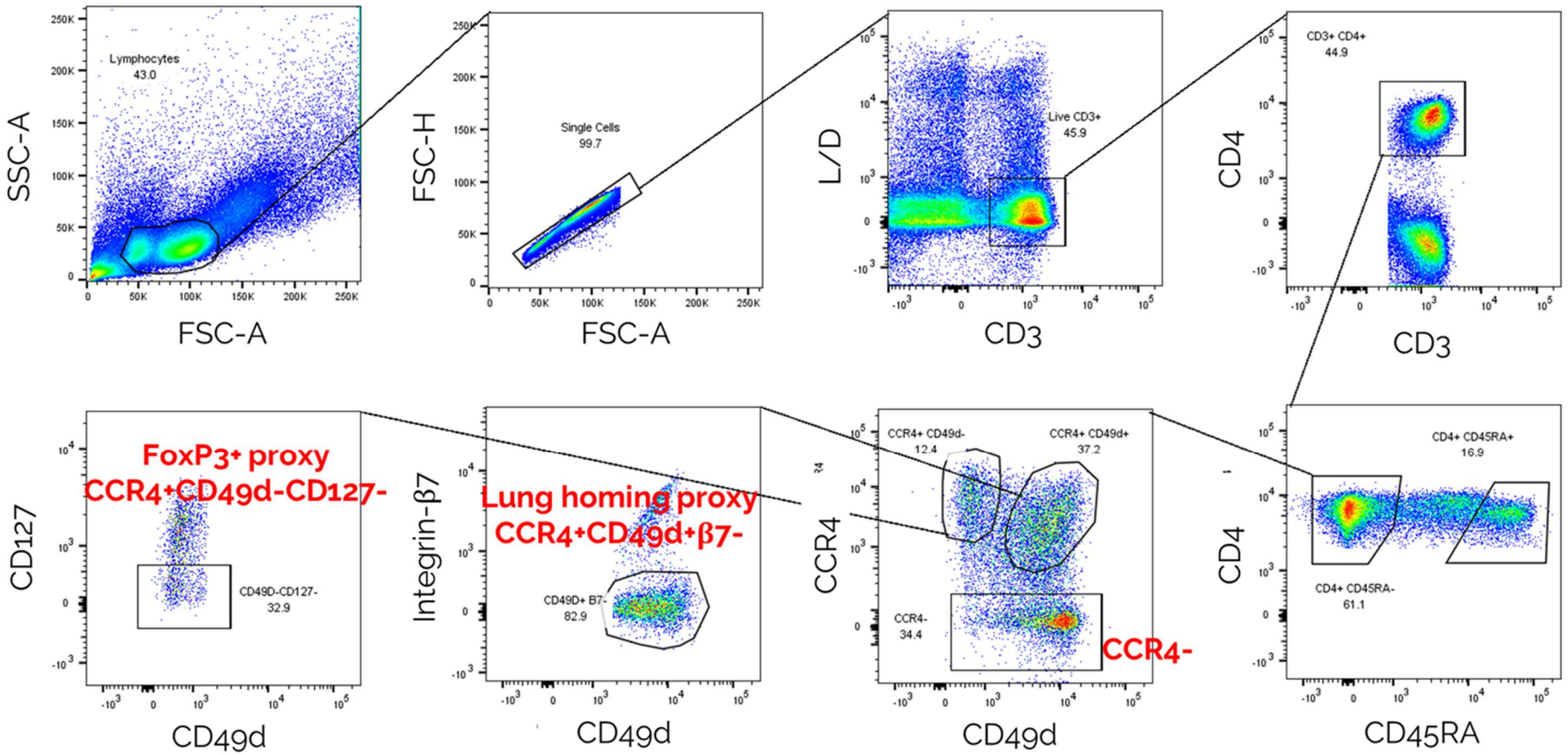


Figure 4.1 – Fluorescent activated cell sorting (FACS) gating strategy for HTLV-1c proviral reservoir phenotyping study. Cryopreserved PBMCs were thawed and stained with a cocktail of antibodies prior to sorting. Cells were first gated by lymphocytes>single cells>live CD3+T cells>CD4+ T cells>CD45RA- (effector/memory). The three phenotype populations of focus were FOXP3+ proxy (CCR4+CD49d-CD127-), lung homing proxy (CCR4+CD49d+Integrin-β7-) and CCR4-.

Table 4.1 - Demographics and clinical characteristics of donors in the HTLV-1c proviral reservoir phenotyping cohort.

	HTLV-1c- Mparntwe (n, 5)	HTLV-1c+ HAPD- Mparntwe (n, 11)	HTLV-1c+ HAPD+ Mparntwe (n, 8)
Demographics, n (%)			
Median Age	54	55	49
		p, 0.56	
Female at birth	4/5 (80%)	7/11 (87.5%)	5/8 (62.5%)
Male at birth	1/5 (20%)	4/11 (12.5%)	3/8 (37.5%)
Lifestyle, n (%)			
Smoker (current or previous)	2/5 (40%)	4/11 (12.5%)	2/8 (25%)
Harmful alcohol consumption	2/5 (40%)	4/11 (37.5%)	3/8 (37.5%)
Comorbidities, n (%)			
Chronic Liver Disease	0/5 (0%)	1/11 (25%)	0/8 (0%)
Diabetes	4/5 (80%)	7/11 (87.5%)	5/8 (62.5%)
Chronic Kidney Disease	3/5 (60%)	9/11 (87.5%)	4/8 (50%)
Malignancy	0/5 (0%)	1/11 (0%)	2/8 (25%)
HTLV-1 associated diseases, n (%)			
HAPD (Bex)	0/5 (0%)	0/11 (0%)	8/8 (100%)
Strongyloidiasis, current	0/5 (0%)	2/11 (0%)	4/8 (50%)
BSI, current or previous	0/5 (0%)	6/11 (37.5%)	4/8 (50%)
HAM	0/5 (0%)	0/11 (0%)	0/8 (0%)
ATL	0/5 (0%)	0/11 (0%)	0/8 (0%)
HTLV-1 burden, log_{10}(copies per 10^6 cells)			
Median PVL	N/A	3.53	3.38
		p, 0.97	

Participants consented to whole blood being taken and used for HTLV-1c pathogenesis research following written and verbal consent in primary language. Information was obtained through clinician and hospital records at Alice Springs Hospital, Mparntwe, NT, Australia. Seronegative community controls (n, 5), HTLV-1c+ HAPD- (n, 11), and HTLV-1c+ HAPD+ patients (n, 8). Other negative control samples from Naarm, Melbourne, Australia were collected through Lifeblood Red Cross, and no demographics are supplied through research agreements, however these donor samples are screened for infectious diseases upon acquisition at Lifeblood. No p values are reported for categorical variables, as Chi-square calculations are only valid when at least 20% of the expected values are greater than 5. For age, statistical significance was assessed using Kruskal-Wallis test. For PVL, statistical significance assessed using Mann-Whitney test.

4.3.2 Lung homing proxy CD4+ T cell phenotype is expanded in HTLV-1c infection

Firstly, we looked at the naïve (CD45RA+) and effector/memory (E/M) (CD45RA-) CD4+ T cell phenotypes across all four test groups. HTLV-1c-infected HAPD- (median, 28.3%; IQR, 17.7-37.1; p, 0.03) and HADP+ donors (median, 27.5%; IQR, 17.4-31.7; p, 0.02), had significantly less proportions of naïve CD4+ T cells when compared to the Naarm negative control group (median, 52.9%; IQR, 38.8-59.3). However, there were no significant differences in proportion of naïve CD4+ T cells between the Mparntwe community negative control group (median, 29.6%; IQR, 21.2-41.5) and infected donor groups (p, 0.41, 0.32, respectively) (**Figure 4.2A**). On the other hand, the frequencies of E/M CD4+ T cells were significantly higher in the HTLV-1c seropositive groups: HAPD- (median, 64.6%; IQR, 49.5-72.1; p, 0.05) and HAPD+ donors (median, 63.5%; IQR, 60.8-71.93; p, 0.04), than uninfected controls from Naarm (median, 41.1%; IQR, 32.3-54.8), and trended higher compared to the community matched negative controls (median, 53.8%; IQR, 46.9-69.6; p, 0.46, 0.32, respectively). (**Figure 4.2A**). Between the HTLV-1c infected groups themselves, there was no discernible difference in the proportion of naïve or E/M phenotypes (p, 0.39, 0.50) (**Figure 4.2A**). Furthermore, the frequency of both naïve (r, -0.04; p,0.88) and E/M (r, -0.04; p, 0.86) CD4+ T cell phenotypes of HTLV-1c infected donors did not correlate to PVL (**Figure 4.2B**). Notably, the frequency of activated (HLA-DR+CD38+) subpopulation of E/M CD4+ T cells in HAPD- (median, 8.91%; IQR, 3.84-13.60; p, 0.002, 0.005, respectively) and HAPD+ (median, 6.66%; IQR, 4.31-8.63; p, 0.007, 0.02, respectively) donors was significantly greater when compared to both Naarm (median, 1.33%; IQR, 1.22-1.99) and Mparntwe (median, 2.51%; IQR, 0.96-4.22) control groups (**Figure 4.2C**). However, the PVL of HTLV-1c+ donors did not correlate to the proportion of activated subpopulation of E/M CD4+ T cells (r, 0.28; p, 0.24) (**Figure 4.2D**).

Next, we examined various CD4+ T cell phenotypes in HTLV-1c infection with CCR4 surface expression, where HTLV-1a provirus was reported to accumulate [340-343]. Firstly, there was no expansion of the FOXP3+ proxy phenotype (CCR4+CD49d-C127-) observed in HTLV-1c infection (**Figure 4.3A**): the frequencies in HAPD- (median, 9.24%; IQR, 7.16-14.80) and HAPD+ (median, 12.00%; IQR, 6.61-18.45) donors were very similar to HTLV-1c uninfected control groups from Naarm (median, 11.30%; IQR,

8.77-12.80; p, 0.96, 0.87, respectively) and Mparntwe (median, 12.3%; IQR, 9.84-16.90; p, 0.87, 0.87, respectively). Furthermore, the frequency of FOXP3+ proxy cells in HTLV-1c infected individuals did not correlate to PVL (r, -0.02; p, 0.95) (**Figure 4.3B**). Nevertheless, the activated subset, HLA-DR+CD38+, of the FOXP3+ proxy phenotype was significantly higher in both HAPD- (median, 8.33%; IQR, 3.90-14.50; p, 0.04, 0.006, respectively) and HAPD+ (median, 10.77%; IQR, 7.05-15.68; p, 0.02, 0.004, respectively) donors than control groups from Naarm (median, 3.53%; IQR, 2.62-4.79) and Mparntwe (median, 1.70%; IQR, 0.75-4.78) (**Figure 4.C**). However, the proportion of FOXP3+ proxy phenotype cells with positive activation surface markers was not correlated to the PVL of HTLV-1c seropositive individuals (r, -0.13; p, 0.60) (**Figure 4.3D**).

With a high association of Bex to HTLV-1c infection in Central Australia, we next explored the potential for trafficking of HTLV-1c infected CD4+ T cells to the respiratory tract, as lymphocyte infiltration of lung tissues occurs in HTLV-1c seropositive individuals, and HTLV-1 DNA has been previously detected in BALF (reviewed in [128]). Importantly, our results indicate there is a large expansion of CD4+ T cells expressing surface markers for the lung homing proxy phenotype (CCR4+CD49d+Integrin-β7-) in HTLV-1c seropositive individuals (**Figure 4.4A**): HAPD- (median, 24.40%; IQR, 17.40-28.90; p, 0.002, 0.009, respectively) and HAPD+ (median, 26.30%; IQR, 22.23-36.28; p, 0.0004, 0.002, respectively) donors showed significant expansion of this phenotype above the negative control groups from Naarm (median, 8.36%; IQR, 3.38-11.47) and Mparntwe (median, 16.2%; IQR, 9.39-19.70) (**Figure 4.4A**). Furthermore, HAPD+ donors trended higher than HAPD- (p, 0.06) (**Figure 4.4A**). Interestingly, we also found the frequency of lung homing proxy phenotype T cells is correlated to the PVL in HTLV-1c infected individuals, (r, 0.44; p, 0.06) however not statistically significant (**Figure 4.4B**). Furthermore, we found a higher frequency of respiratory trafficking proxy CD4+ T cells expressing activation phenotype markers was associated with HTLV-1c infection (**Figure 4.4C**): significantly higher proportions in HAPD- (median, 13.30%; IQR, 6.49-14.50; p, 0.007, 0.005, respectively) and trending higher in HAPD+ (median, 6.42%; IQR, 4.00-11.83; p, 0.07, 0.06) donors when compared to uninfected Naarm-based (median, 3.2%; IQR, 2.21-4.78) and Mparntwe-based (median, 3.24%; IQR, 0.61-4.99) controls. However, for HTLV-1c

infected donors, the proportion lung homing proxy phenotype cells with activation marker expression did not correlate to PVL (r, 0.16; p, 0.51) (**Figure 4.4D**).

As a control phenotype, we also investigated the circulating frequency and activation of CCR4-CD45RA-CD4+ T cells, where HTLV-1a subtype provirus was reported to not preferentially infect [340-343]. While the frequency of the CCR4- phenotype was significantly lower across HAPD- (median, 26.70%; IQR, 16.10-41.50) and HAPD+ (median, 21.05%; IQR, 15.80-30.08) donors when compared to uninfected controls from Naarm (median, 47.50%, IQR, 36.60-54.00; p, 0.03, 0.03, respectively), there was no discernible difference between the two HTLV-1c+ groups and control HTLV-1 seronegatives from Mparntwe (median, 30.20%; IQR, 18.30-50.90; p, 0.31, 0.22, respectively) (**Figure 4.5A**). Moreover, the frequency of CCR4- phenotype in E/M CD4+ T cells did not correlate to PVL in HTLV-1c infected individuals (r, -0.2282; p, 0.35) (**Figure 4.5B**). Our results do indicate, however, that in the two HTLV-1c seropositive groups from Mparntwe, a significantly higher proportion of CCR4- cells expressed the HLA-DR+CD38+ activation markers in HAPD- (median, 6.50%; IQR, 3.43-13.10; p, 0.007, 0.0009, respectively) and HAPD+ (median, 6.67%; IQR, 2.43-8.69; p, 0.02, 0.002, respectively) donors compared to the local community matched uninfected Mparntwe controls (median, 1.65%; IQR, 1.06-2.96) and even more so than those from Naarm (median, 0.58%; IQR, 0.54-0.87) (**Figure 4.5C**). Like other CD4+ T cell phenotypes investigated in this study, the activation frequency of CCR4- cells did not correlate with PVL in HTLV-1c infection (r, 0.35; p, 0.14) (**Figure 4.5D**).

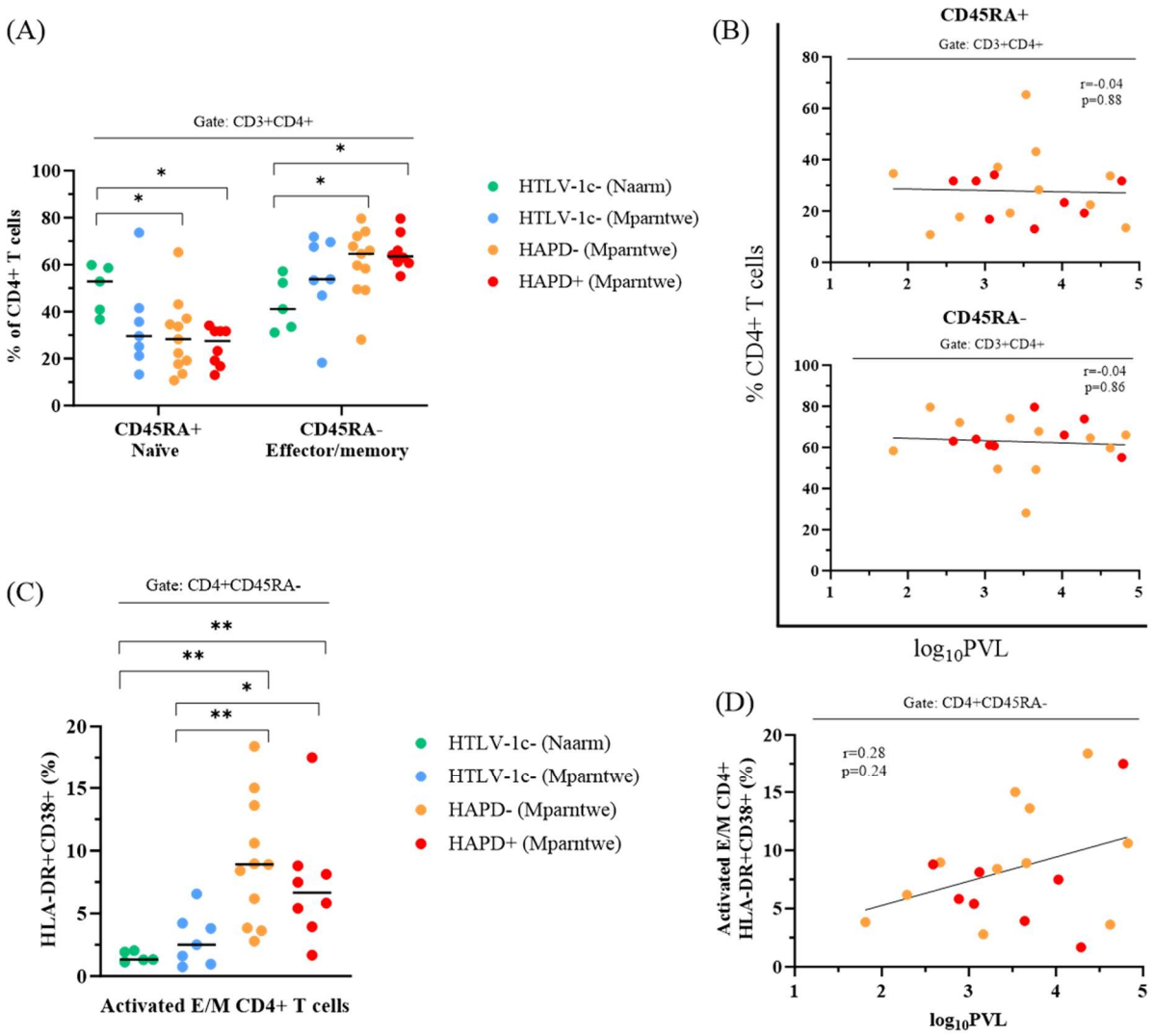


Figure 4.2 – Effector/memory CD4+ T cells in patient PBMCs are activated in HTLV-1c infection. Cryopreserved patient PBMCs were stained for surface marker expression with a cocktail of mAb and analysed by FACS. gDNA was extracted from PBMCs and HTLV-1c PVL was determined by ddPCR. (A) Frequency of naïve (CD45RA+) and E/M (CD45RA-) CD4+ T cells. (B) Correlation of HTLV-1c+ patient PVL with frequency of naïve and E/M CD4+ T cells. (C) Subset of E/M CD4+ T cells with activation marker (HLA-DR+CD38+) expression. (D) Correlation of HTLV-1c+ patient PVL and proportion of activated E/M CD4+ T cells. Negative controls from Naarm (n, 4) are shown in green, negative community controls from Mparntwe (n, 7) are shown in blue, HTLV-1c+ HAPD- (n, 11) from Mparntwe are shown in orange, and HTLV-1c+ HAPD+ donors (n, 8) from Mparntwe are shown in red. Median frequency of each test group is shown by black line. Statistical significance between groups was assessed using Kruskal-Wallis test with two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli as an FDR control for multiple comparisons. Black trend line represents nonlinear regression straight line of best fit. Correlation significance was assessed using Spearman's correlation test. * $p<0.05$; ** $p<0.01$.

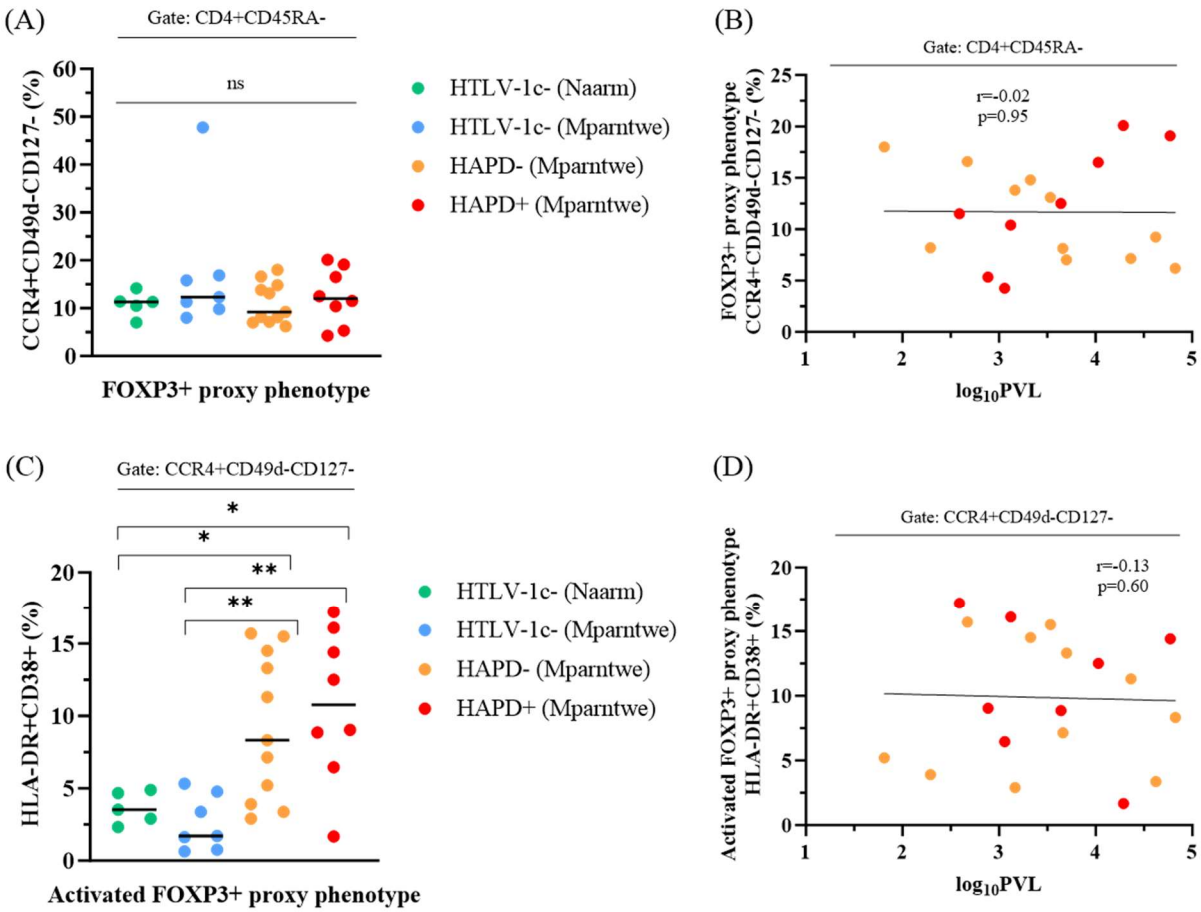


Figure 4.3 – FOXP3+ proxy phenotype CD4+ T cells in patient PBMCs do not expand during HTLV-1c infection but are highly activated. Cryopreserved patient PBMCs were stained for surface marker expression with a cocktail of mAb and analysed by FACS. gDNA was extracted from PBMCs and HTLV-1c PVL was determined by ddPCR. (A) Frequency of FOXP3+ proxy (CCR4+CD49d-CD127-) phenotype CD4+ T cells. (B) Correlation of HTLV-1+ PVL with frequency of FOXP3+ proxy phenotype CD4+ T cells. (C) Subset of FOXP3+ proxy phenotype CD4+ T cells with activation marker (HLA-DR+CD38+) expression. (D) Correlation of HTLV-1c+ patient PVL with proportion of activated FOXP3+ proxy cells. Negative controls from Naarm (n, 4) are shown in green, negative community controls from Mparntwe (n, 7) are shown in blue, HTLV-1c+ HAPD- (n, 11) from Mparntwe are shown in orange, and HTLV-1c+ HAPD+ donors (n, 8) from Mparntwe are shown in red. Median frequency of each test group is shown by black line. Statistical significance between groups was assessed using Kruskal-Wallis test with two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli as an FDR control for multiple comparisons. Black trend line represents nonlinear regression straight line of best fit. Correlation significance was assessed using Spearman's correlation test. ns, not significant, $p>0.05$; * $p<0.05$; ** $p<0.01$.

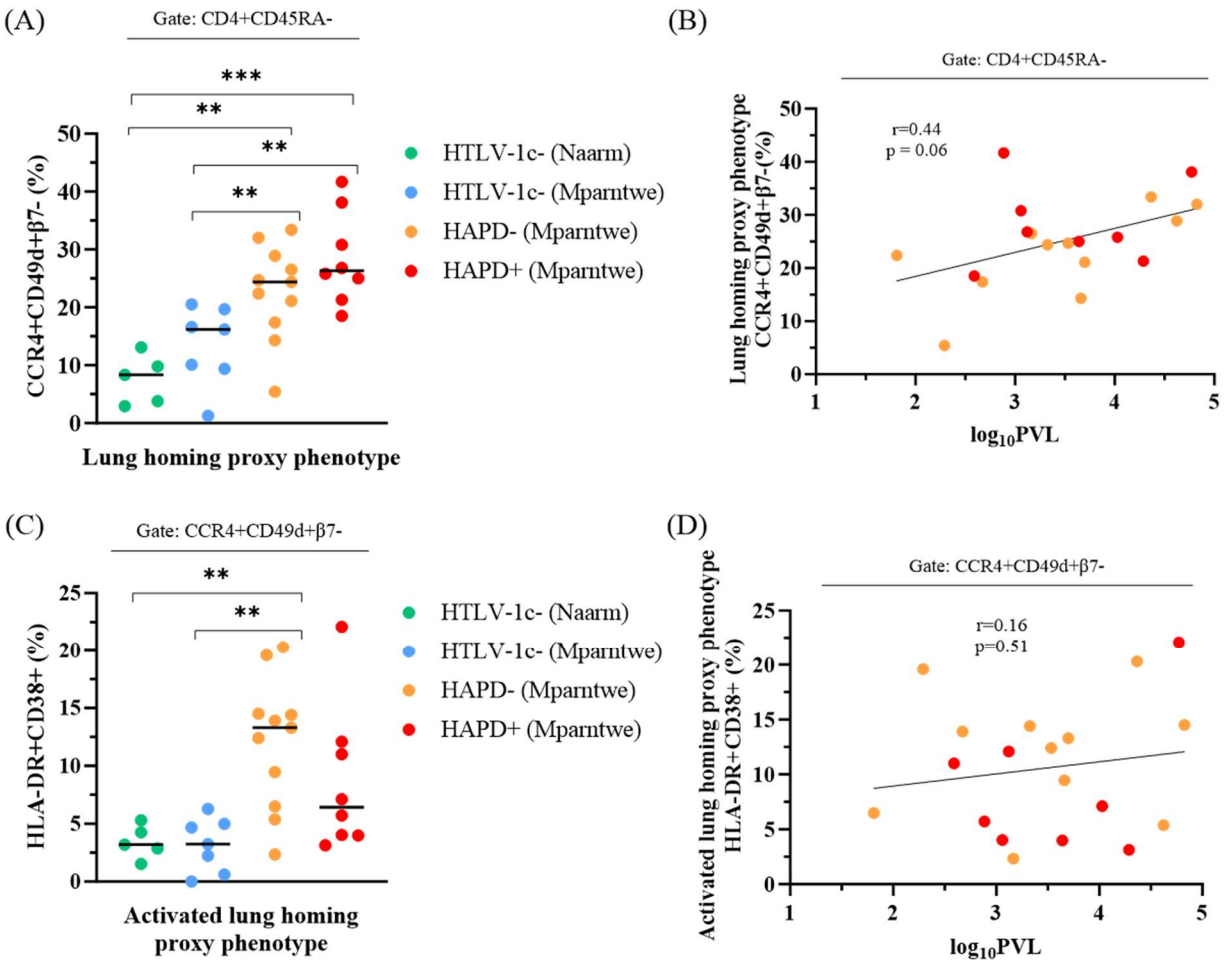


Figure 4.4 – CD4+ T cells with potential lung trafficking phenotype are significantly expanded and activated in HTLV-1c infection. Cryopreserved patient PBMCs were stained for surface marker expression with a cocktail of mAb and analysed by FACS. gDNA was extracted from PBMCs and HTLV-1c PVL was determined by ddPCR. (A) Frequency of lung homing proxy phenotype (CCR4+CD49d+IntegrinB7-) CD4+ T cells. (B) Correlation of HTLV-1c+ patient PVL with frequency of lung homing proxy phenotype cells. (C) Subset of lung homing proxy phenotype CD4+ T cells with activation marker (HLA-DR+CD38+) expression. (D) Correlation of HTLV-1c+ patient PVL and proportion of activated lung homing proxy CD4+ T cells. Negative controls from Naarm (n, 4) are shown in green, negative community controls from Mparntwe (n, 7) are shown in blue, HTLV-1c+ HAPD- (n, 11) from Mparntwe are shown in orange, and HTLV-1c+ HAPD+ donors (n, 8) from Mparntwe are shown in red. Median frequency of each test group is shown by black line. Statistical significance between groups was assessed using Kruskal-Wallis test with two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli as an FDR control for multiple comparisons. Black trend line represents nonlinear regression straight line of best fit. Correlation significance was assessed using Spearman's correlation test. * $p<0.05$; ** $p<0.01$; *** $p<0.001$.

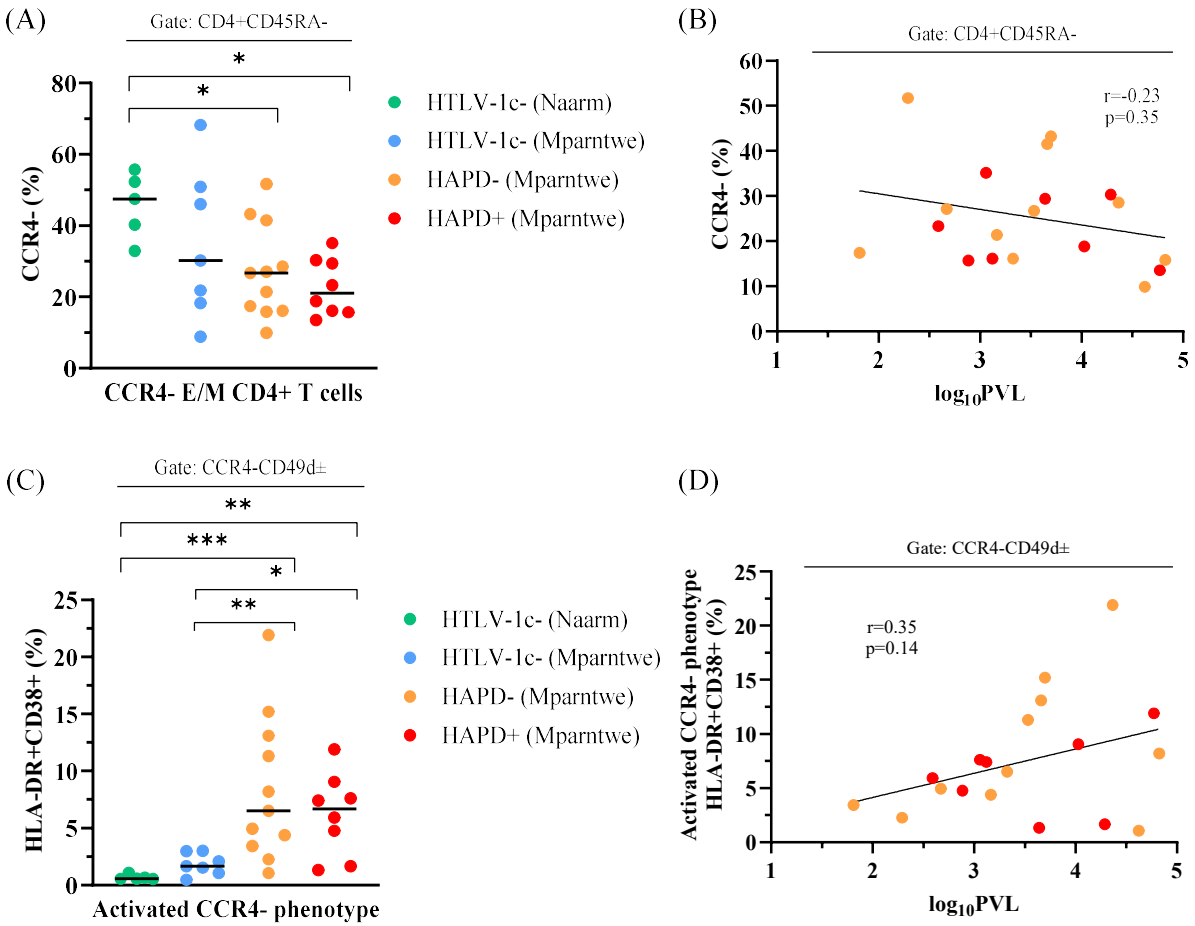


Figure 4.5 – CCR4- effector/memory CD4+ T cells are highly activated in HTLV-1c infection. Cryopreserved patient PBMCs were stained for surface marker expression with a cocktail of monoclonal antibodies and analysed by FACS. gDNA was extracted from PBMCs and HTLV-1c PVL was determined by ddPCR. (A) Frequency of CCR4- effector/memory (E/M) CD4+ T cells in patient PBMCs. (B) Correlation of HTLV-1c+ patient PVL with frequency of CCR4- phenotype E/M CD4+ T cells. (C) Subset of CCR4- phenotype E/M CD4+ T cells with activation marker (HLA-DR+CD38+) expression. (D) Correlation of HTLV-1c+ patient PVL and proportion of activated CCR4- phenotype E/M CD4+ T cells. Negative controls from Naarm (n, 4) are shown in green, negative community controls from Mparntwe (n, 7) are shown in blue, HTLV-1c+ HAPD- (n, 11) from Mparntwe are shown in orange, and HTLV-1c+ HAPD+ donors (n, 8) from Mparntwe are shown in red. Median frequency of each test group is shown by black line. Statistical significance between groups was assessed using Kruskal-Wallis test with two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli as an FDR control for multiple comparisons. Black trend line represents nonlinear regression straight line of best fit. Correlation significance was assessed using Spearman's correlation test. ns, not significant, $p>0.05$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$.

4.3.3 HTLV-1c provirus is enriched in both CCR4+ and CCR4- CD4+ T cell reservoirs

Given that HTLV-1 studies with subtype-A identified a concentration of provirus in CCR4+ cells and use this as a targeting method for antiviral therapy, we next assessed which CD4+ T cell phenotypes of interest were indeed infected with subtype-C provirus. We employed a previously established ddPCR assay to quantify the PVL, in each T cell subpopulation phenotype purified from 6 seropositive donors. We used primers and probe targeting the *tax* gene and housekeeping gene *RPP30* [218] (**Figure 4.6A**) and compared against the PVL measured in corresponding bulk PBMC samples. We $\log_{10}(x)$ transformed PVL values (copies provirus per 10^6 cells) for better visualisation of the widely distributed dataset. Our results indicate that HTLV-1c provirus was significantly enriched in the FOXP3+ proxy phenotype (median, 5.031; p, 0.001) when compared to the matching bulk PBMC PVL (median, 3.677) (**Figure 4.6B**). So, while the circulating frequency of FOXP3+ proxy phenotype does not expand during HTLV-1c infection (**Figure 4.3A**), it still contains a significant reservoir of integrated provirus (**Figure 4.6B**). Furthermore, the lung homing proxy phenotype is significantly enriched with provirus (median, 5.036; p, 0.02) when compared to the corresponding PVL of bulk PBMCs (**Figure 4.6B**), indicating these cells are both expanded (**Figure 4.4A**) and highly infected with HTLV-1c (**Figure 4.6B**). Unexpectedly, the CCR4- phenotype of E/M CD4+ T cells, which was initially designed as a control reservoir, also displayed a significant enrichment of HTLV-1c provirus (median, 4.755; p, 0.05) relative to bulk PBMCs (**Figure 4.6B**).

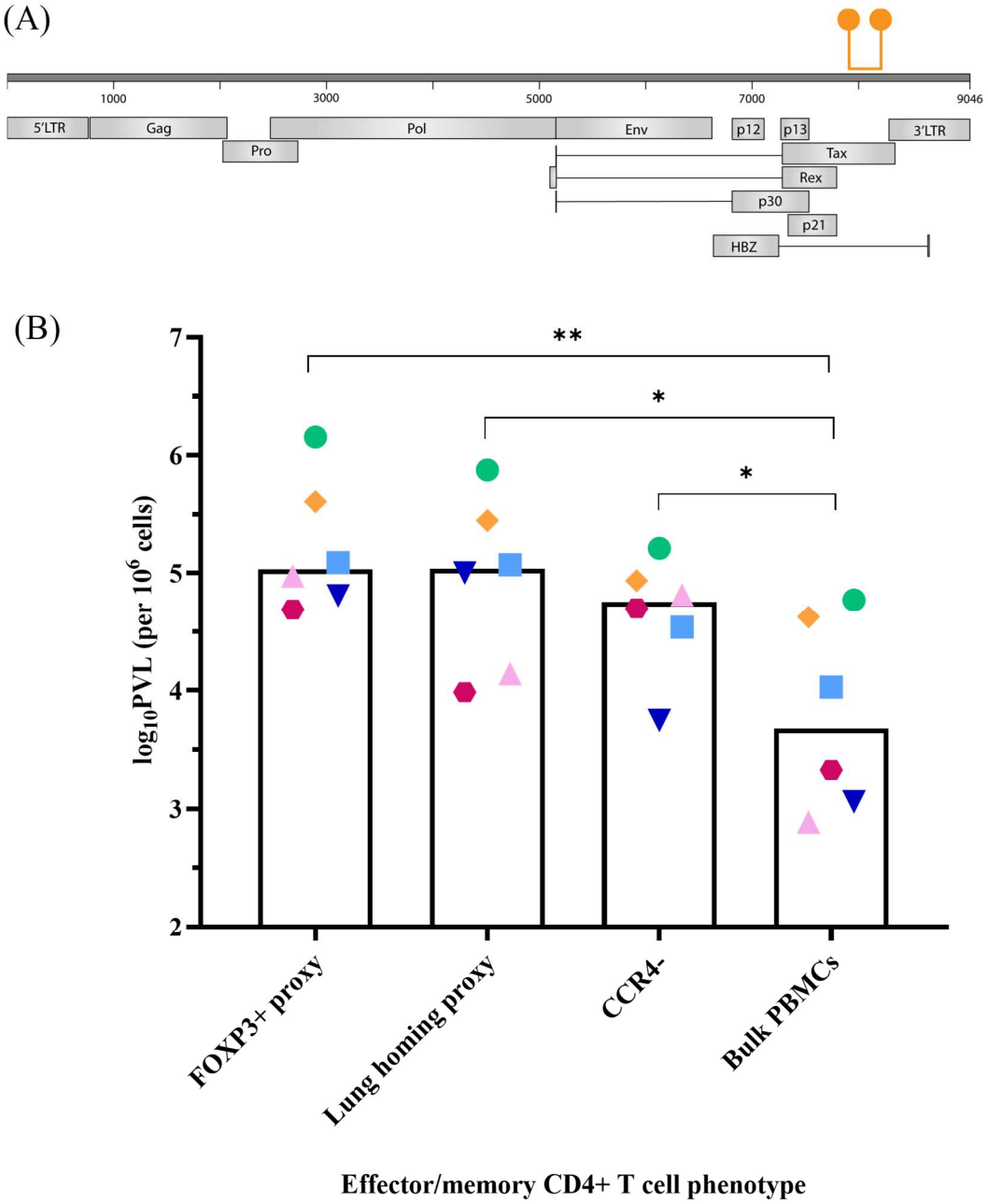


Figure 4.6 – HTLV-1c provirus is enriched in CCR4+ and CCR4- phenotypes of effector/memory CD4+ T cells in patient PBMCs. Cryopreserved patient PBMCs were stained for surface marker expression with a cocktail of monoclonal antibodies and analysed by FACS. gDNA was extracted from sorted E/M CD4+ T cell phenotypes and bulk PBMCs, and HTLV-1c PVL was determined by ddPCR. (A) Schematic of the HTLV-1c genome with ddPCR primers and probe in the *tax* gene region. (B) Enrichment of HTLV-1c provirus in E/M CD4+ T cell reservoirs in human PBMCs (n, 6). Each symbol/colour represents a matched donor samples and is the mean of duplicate measurements. Bar graphs show the median PVL for each reservoir. Statistical significance was assessed using Friedman test with two stage linear step-up procedure of Benjamini, Krieger and Yekutieli as an FDR control for multiple comparisons. * $p<0.05$; ** $p<0.01$.

4.4 Characterising the proviral landscape in HTLV-1c infection in vivo

4.4.1 *tax* and *hbz* gene regions are detected at higher levels than *gag* or *env* in HTLV-1c proviral reservoirs by ddPCR

Preliminary HTLV-1c genomic sequencing revealed there may be significant deletions in proviral sequences (Hirons et al., unpublished). Absolute proviral DNA levels measured relative to a reference cellular gene, *RPP30*, using specific ddPCR probes showed the *tax* gene region was often measured at higher levels than the *gag* region [155]. This aligned with bulk PCR sequencing data that mapped numerous missing genome segments (Hirons et al., unpublished). Therefore, we designed additional ddPCR primer/probe targets along the HTLV-1c genome to quantify and compare the measured amount of *gag*, *env*, *hbz* and *tax* regions in proviral DNA relative to the number of cellar DNA copies (**Figure 4.7A**). gDNA was extracted from FACS sorted CD4+ phenotypes (FOXP3+ proxy, lung homing proxy, CCR4-) and bulk PBMCs, and used as template for ddPCR. The PVL of each target region was performed in duplicate, for each reservoir, of 8 donors (HAPD- n, 3; HAPD+ n, 5). The retention frequency of a given gene was normalised to the highest detected gene for each phenotype, for each donor.

Overall, high levels of defective provirus were present in all donors, as indicated by the varied levels of detection of the four genomic regions of HTLV-1c (**Figure 4.7B**). In the bulk PBMC reservoir of all donors, both *hbz* and *tax* genomic regions were present in most provirus sequences (mean, 82.12%, 86.04%, respectively). In contrast, *gag* was retained in less than half of proviral sequences (mean, 45.10%), while *env* was mostly deleted (mean, 21.02%). *hbz* and *tax* gene regions were similarly the highest retained in the FOXP3+ proxy proviral reservoir (mean, 84.06%, 80.28%, respectively), while *gag* and *env* regions were less frequently detected (mean, 68.50%, 64.67%, respectively). On the other hand, the lung homing proxy proviral reservoir had approximately the same level of retention across *gag*, *hbz* and *tax* (mean, 71.32%, 76.79%, 70.56%, respectively), while *env* remained the most deleted genomic region (mean, 56.15%). In the CCR4- proviral reservoir of all donors, *gag* (mean, 79.16%) and *tax* (mean, 82.40%) coding regions were mostly retained, while *env* (mean, 69.20%) and *hbz* (mean, 69.80%) were detected slightly less.

We then compared the normalised gene retention between HAPD- and HAPD+ groups (**Figure 4.7C, 4.7D**). Overall, HAPD+ donors had higher retention of all genes across all HTLV-1c proviral reservoirs. In the bulk PBMC reservoir, HAPD- donors had highest retention of *hbz* coding region (mean, 100%), followed by *tax* (mean, 75.84%), while *gag* was less retained (mean, 36.61%). Notably, *env* was mostly deleted from the proviral reservoir for HAPD- donors: it was detected at a much lower rate than the other genomic regions (mean, 10.98%). In the bulk PBMC reservoir of HAPD+ donors, *tax* and *hbz* genes had the highest retention in proviral sequences (mean, 92.16%, 71.39%, respectively). In contrast, *gag* and *env* regions were more commonly deleted in the proviral sequences (mean, 50.20%, 27.04%), however, were retained at higher levels than the corresponding PBMC reservoir of HAPD- donors.

The FOXP3+ proxy phenotype showed similar proviral gene retention patterns when comparing both HTLV-1c disease groups (**Figure 4.7C, 4.7D**). *hbz* and *tax* regions were most frequently retained for both HAPD- (mean, 88.80%, 82.31%, respectively) and HAPD+ donors (mean, 80.50%, 79.06%, respectively). *gag* and *env* regions were detected at approximately 10-20% lower frequencies but were very similar for both disease groups: HAPD- (mean, 67.4%, 65.9%, respectively) and HAPD+ (mean, 69.0%, 63.9%, respectively) donors.

Interestingly, it appears there are some differences in the proviral characteristics of the lung homing proxy phenotype between HAPD- and HAPD+ donors. The frequency of gene retention in HAPD- donors was approximately 60% for each of the genomic regions: *gag* (mean, 57.43%), *env* (mean, 66.28%), *hbz* (mean, 66.60%), *tax* (mean, 64.59%) (**Figure 4.7C**). Contrastingly, in the respiratory trafficking proxy phenotype of HAPD+ donors, the provirus pool retained *gag*, *hbz* and *tax* regions at higher frequencies than HAPD- donors (mean, 79.65%, 82.9%, 74.14%, respectively) with less retention of *env* gene region (mean, 50.06%) (**Figure 4.7D**).

Finally, in the proviral reservoir of CCR4- phenotype of E/M CD4+ T cells, HAPD- donors showed noticeably low retention of the *hbz* genomic region (mean, 42.47%) (**Figure 4.7C**), while HAPD+ donors displayed extraordinarily high levels of *hbz* retention (mean, 97.13%) (**Figure 4.7D**). HAPD- donors showed similar levels of *gag*, *env* and *tax* detection in the CCR4- phenotype proviruses (mean, 89.69%, 82.47%, 78.68%, respectively) (Figure 4.7C). On the other hand, while HAPD+ donors showed

high levels of *tax* retention (mean, 85.19%) in the CCR4- proviral reservoir, *gag* (mean, 68.63%) and *env* (mean, 59.24%) regions were less frequently detected (**Figure 4.7D**).

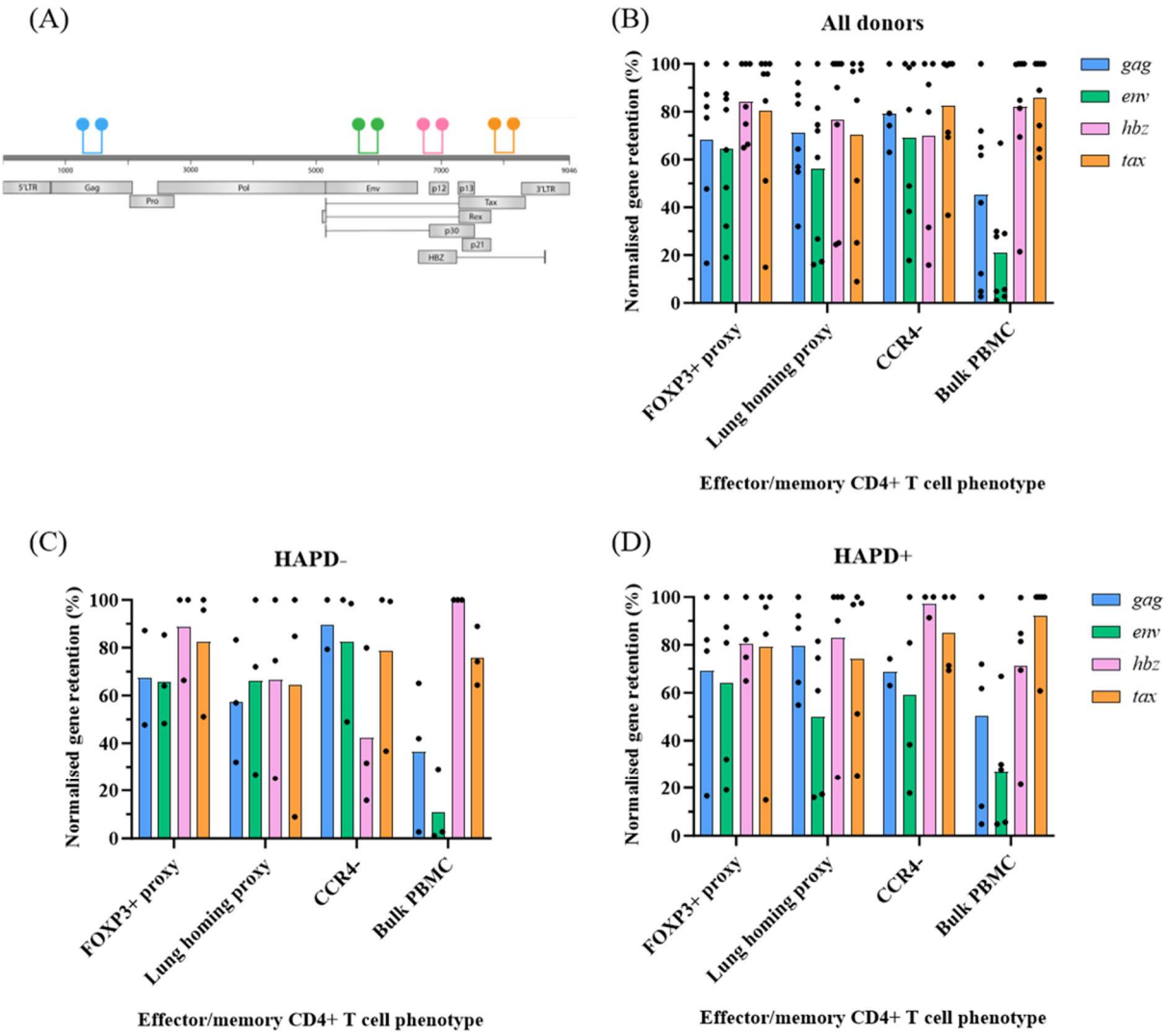


Figure 4.7 – Normalised gene retention of the HTLV-1c provirus pool in effector/memory CD4+ T cell reservoirs in human PBMCs. PBMCs were stained and processed via FACS into various E/M CD4+ T cell phenotypes. gDNA was extracted and ddPCR was performed to determine PVL. The PVL of each genomic region is measured in duplicate, and then normalised to the highest mean PVL of a given gene region in that sample. (A) HTLV-1c genome with primers and probes for ddPCR targeting the gene regions of *gag* (blue), *env* (green), *hbz* (pink) and *tax* (orange). (B) Normalised gene retention frequency of HTLV-1c provirus in E/M CD4+ T cell phenotypes for all seropositive donors (n, 8). (C) Normalised gene retention frequency of HTLV-1c provirus in E/M CD4+ T cell phenotypes reservoirs for HAPD- donors (n, 3). (D) Normalised gene retention frequency of HTLV-1c provirus in E/M CD4+ T cell phenotypes for HAPD+ donors (n, 5). The mean normalised gene retention is shown by bar graphs, with individual points representing different donors.

4.4.2 High levels of defective provirus are present in all HTLV-1c infected donors

With differences in the detection of structural genes compared to regulatory genes in the bulk provirus pool of HTLV-1c infected cells presented earlier in this chapter, we wanted to further investigate the proviral landscape at higher resolution. To achieve this, we designed the single provirus amplification (SPA) assay to purify individual clones of integrated provirus. Based on HIV-1 provirus sequencing [270], the SPA assay uses a limiting dilution, highly processive, touchdown, nested PCR to amplify single integrations of provirus. The limiting dilution is such that there is less than one copy of provirus per well (30%), hence eliminating any PCR biases that would occur for preferential amplification of shorter DNA sequences. The nested primers are staggered in the LTR regions to avoid multiple binding sites and amplifications of the LTRs only, meanwhile avoiding amplification of integration sites which would include First Nations people's DNA (**Figure 4.8A**). Sovereignty of Indigenous genomic information must always be considered when undertaking genomic sequencing studies and conducting ethical scientific research. In addition, we checked in silico that the primers shouldn't amplify endogenous retrovirus LTRs, and confirmed this by running control PCRs with HTLV-1c negative gDNA. Using the SPA assay, we characterised the proviral landscape in bulk PBMCs of six HTLV-1c infected donors (HAPD- n, 3; HAPD+ n, 3), and we further amplified individual proviruses from the FACS sorted CD4+ T cell phenotypes of two donors (HAPD- n, 1; HAPD+ n, 1). We sequenced all proviral amplicons on the Oxford Nanopore Technologies (ONT) platform, summarised in **Figure 4.8B**. In brief, we generated libraries using an ONT ligation kit (**Figure 4.8B(1)**) and sequenced the barcoded libraries using a minion Mk1c flow cell (**Figure 4.8B(2)**). Base calling of reads was performed with Guppy (**Figure 4.8B(3)**). Upon demultiplexing, reads were filtered to select only those flanked by the nested primer sequences (**Figure 4.8B(4)**). Stringent quality control was applied, and further size selection of reads was performed by comparison to original agarose gel of positive SPA wells (**Figure 4.8B(5)**). A consensus sequence was generated for each individual provirus, using 1000 reads, on the Lamassemble platform (**Figure 4.8B(6)**). Finally, each provirus consensus sequence was mapped to the previously characterised HTLV-1c consensus genome (Hirons et al., unpublished) (**Figure 4.8B(7)**). Additionally, we determined the PVL of each SPA donor with gDNA isolated from bulk PBMCs using ddPCR targeting the *tax* gene region.

Overall, 267 proviral consensus sequences were generated from the six HTLV-1c seropositive donors after stringent QC and filtering steps on the ONT platform. 73.6% of all proviral sequences had large internal deletions, while only 2.8% were full length (**Figure 4.9A**). Meanwhile, 2% of proviruses had an internal inverted sequence (**Figure 4.9A**). Most interestingly, 21.5% of proviral sequences were chimeric, such that they had an internal sequence that mapped to the human genome (hg) (**Figure 4.9A**).

Every defective provirus sequence amplified from bulk PBMCs from each donor was mapped to the HTLV-1c genomic consensus sequence, and a normalised read count was generated (to the arbitrary value of 200) for each nucleotide in the genome (**Figure 4.9B**). The LTRs were the highest genomic regions retained, with a normalised read count of approximately 180-200, explained by one of the filtering steps which selected reads with primer binding sites. The first half of the *gag* coding region was moderately retained with a normalised read count of approximately 130, however there was a steep drop off in the second half of the *gag* coding region to a normalised read count of approximately 60. The *pX* region, towards the 3' end of the genome, was retained quite consistently, with normalised read counts between 80-180 across the genomic region. Meanwhile, structural and enzymatic coding regions *pro*, *pol* and *env* were rarely retained in the defective proviruses, with a normalised read count of less than 20.

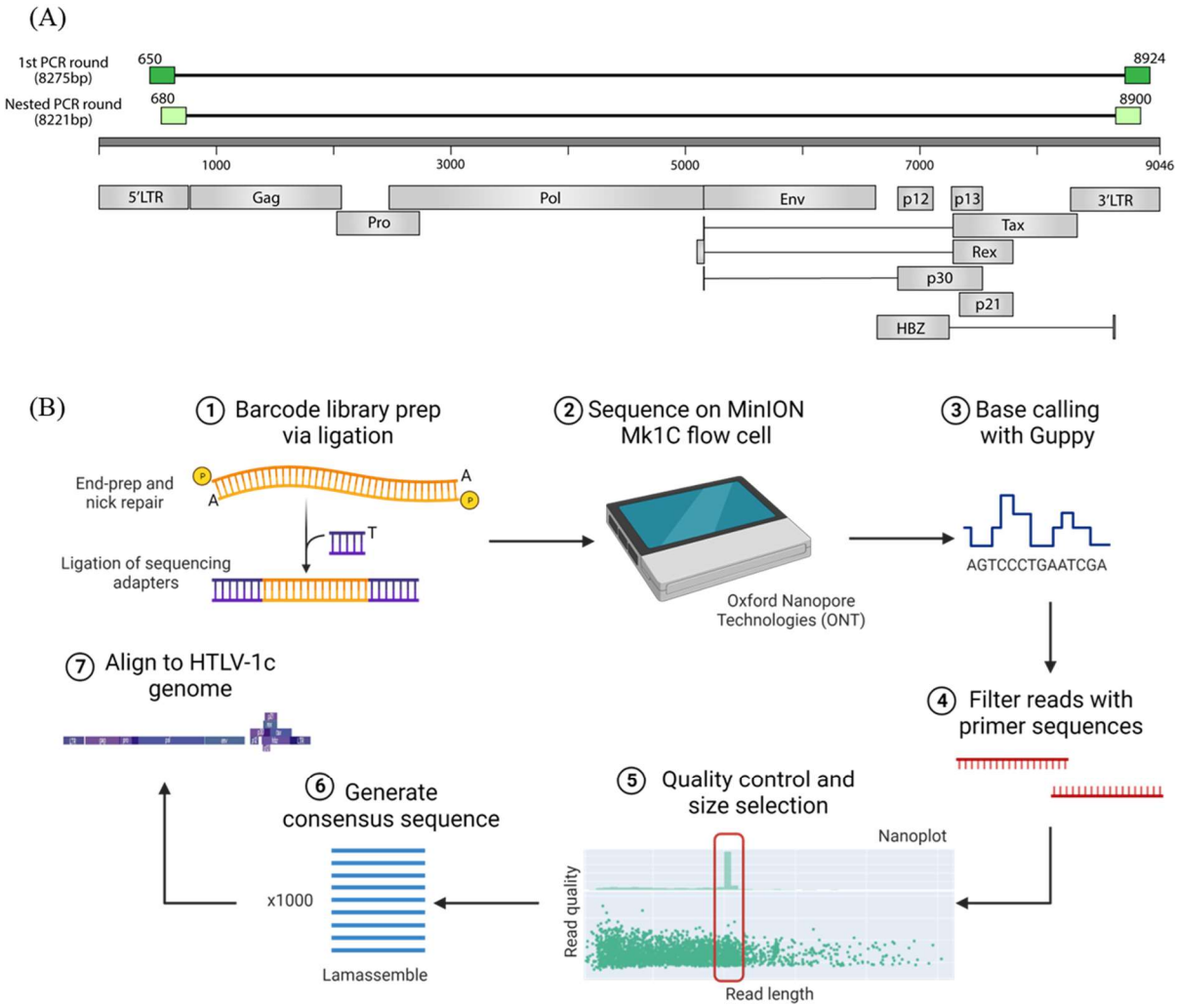


Figure 4.8 – Single provirus amplification assay with Oxford Nanopore Technologies sequencing strategy. (A) Individual proviruses were amplified with SPA assay by limiting dilution touchdown nested PCR and purified. (B) ONT sequencing workflow; (1) ONT barcode library prep was performed with ligation protocol; (2) Library was sequenced on the MinION Mk1C flow cell; (3) Base calling was performed with Guppy; (4) All reads were filtered to select only those that contain both nested primer sequences; (5) QC of reads was performed and read-length distribution determined with Nanoplot, to size select reads based on comparison to original agarose gel of SPA assay; (6). A consensus sequence using 1000 reads was generated for each individual provirus sequence on the Lamassemble platform; (7) The proviral sequences were aligned to the HTLV-1c consensus genome sequence. Section (B) was created with BioRender.com

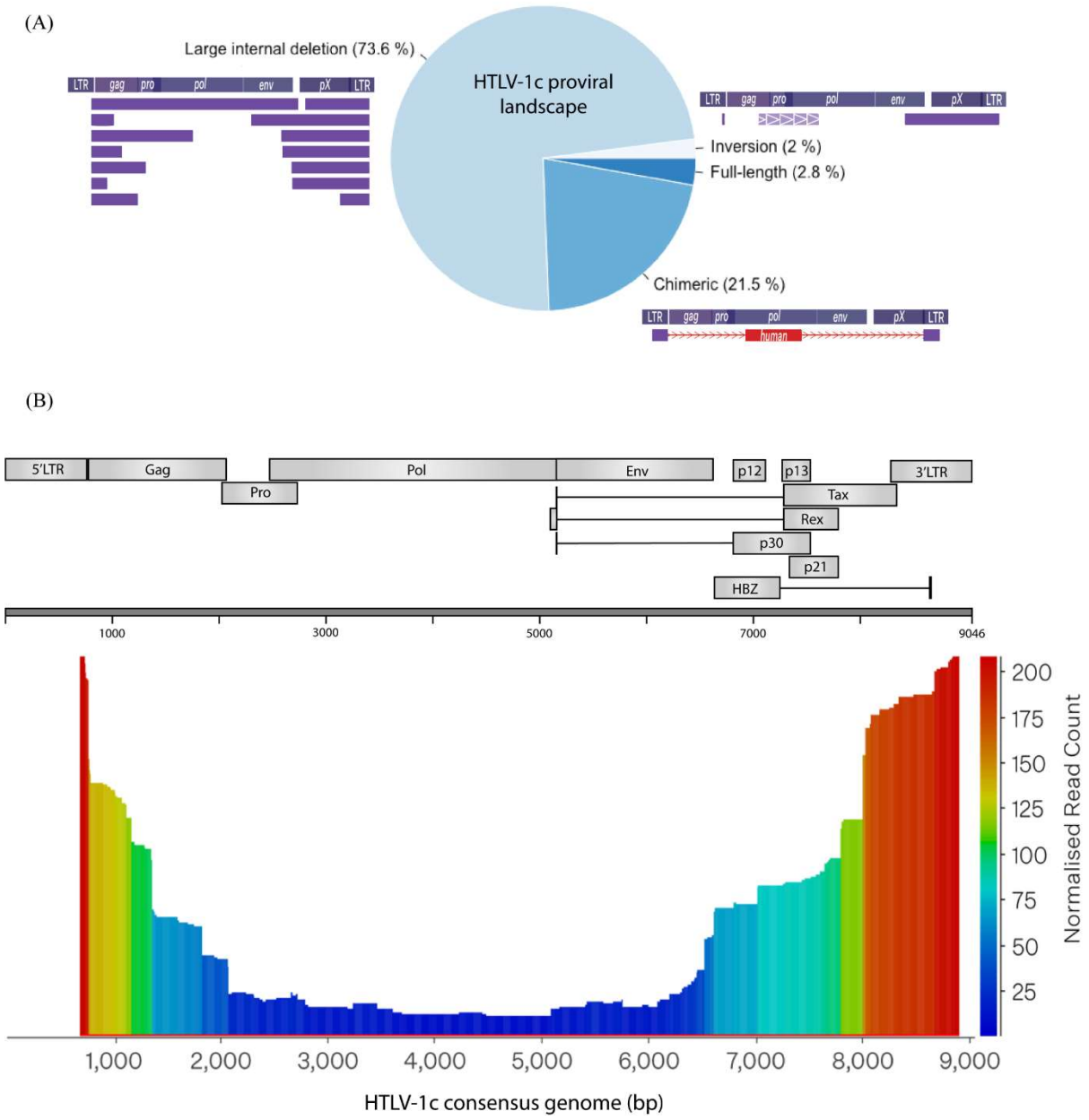


Figure 4.9 – The overall HTLV-1c proviral landscape in primary human PBMCs is highly defective. gDNA was extracted from primary patient PBMCs and sorted CD4+ T cell phenotypes, and SPA assay was performed by limiting dilution touchdown nested PCR. Individual proviral amplicons were sequenced with the ONT platform and aligned to the HTLV-1c consensus genome. (A) The overall frequency of provirus sequence categories for all amplicons: Large internal deletions, inversions, hg-HTLV-1c chimeras and full-length provirus. (B) The normalised read count of each nucleotide across the HTLV-1c consensus genome for the overall proviral landscape for primary human PBMCs, excluding hg-HTLV-1c chimeras.

4.4.3 HAPD+ donors show greater retention of HTLV-1c *pX* region in defective provirus than HAPD-

Given the overall proviral landscape of HTLV-1c infection showed significant deletions, we then sought to determine if any features of the proviral sequences may be driving pathogenesis, particularly HAPD. We first separated the proviral sequences amplified from bulk PBMCs (excluding chimeric or inversion proviruses) by donor (**Figure 4.10A**). While multiple provirus sequences within individual donors were identical, none were mirrored between different donors. Nevertheless, there remained some defining features and differences between the HTLV-1c proviral landscapes of HAPD- and HAPD+ patients. Firstly, the only full-length provirus amplified from PBMCs originated from a Bex patient: none were detected in HAPD- donors (**Figure 4.10A**). We then determined the mean percentage coverage of the genome for each donor by averaging the number of nucleotides of each individual provirus consensus sequence out of the entire possible amplification region (**Figure 4.10B**). In summary, the proviral landscape in PBMCs retained higher coverage of the HTLV-1c genome in HAPD+ patients (mean, 33.34%) than HAPD- donors (mean, 23.00%) (**Figure 4.10C**). However, the PVL of infected individuals did not correlate to the genome coverage frequency of the provirus pool in PBMCs (r, -0.14; p, 0.80) (**Figure 4.10D**).

Next, we compared the provirus sequences between pulmonary disease status by pooling the genome coverage of the donors to create a normalised read count of each nucleotide across the HTLV-1c genome for the two seropositive groups (**Figure 4.11A, 4.11B**). Next, the normalised read count of proviral landscape of HAPD- was subtracted from HAPD+ patients (**Figure 4.11C**). Strikingly, HAPD+ donors retain the entire *pX* region at a markedly higher normalised rate than HAPD- donors (**Figure 4.11C**). Taken together, Bex patients have higher HTLV-1c genomic coverage in their total provirus pool, and this difference is mostly concentrated in the entire *pX* region, when compared to HAPD-.

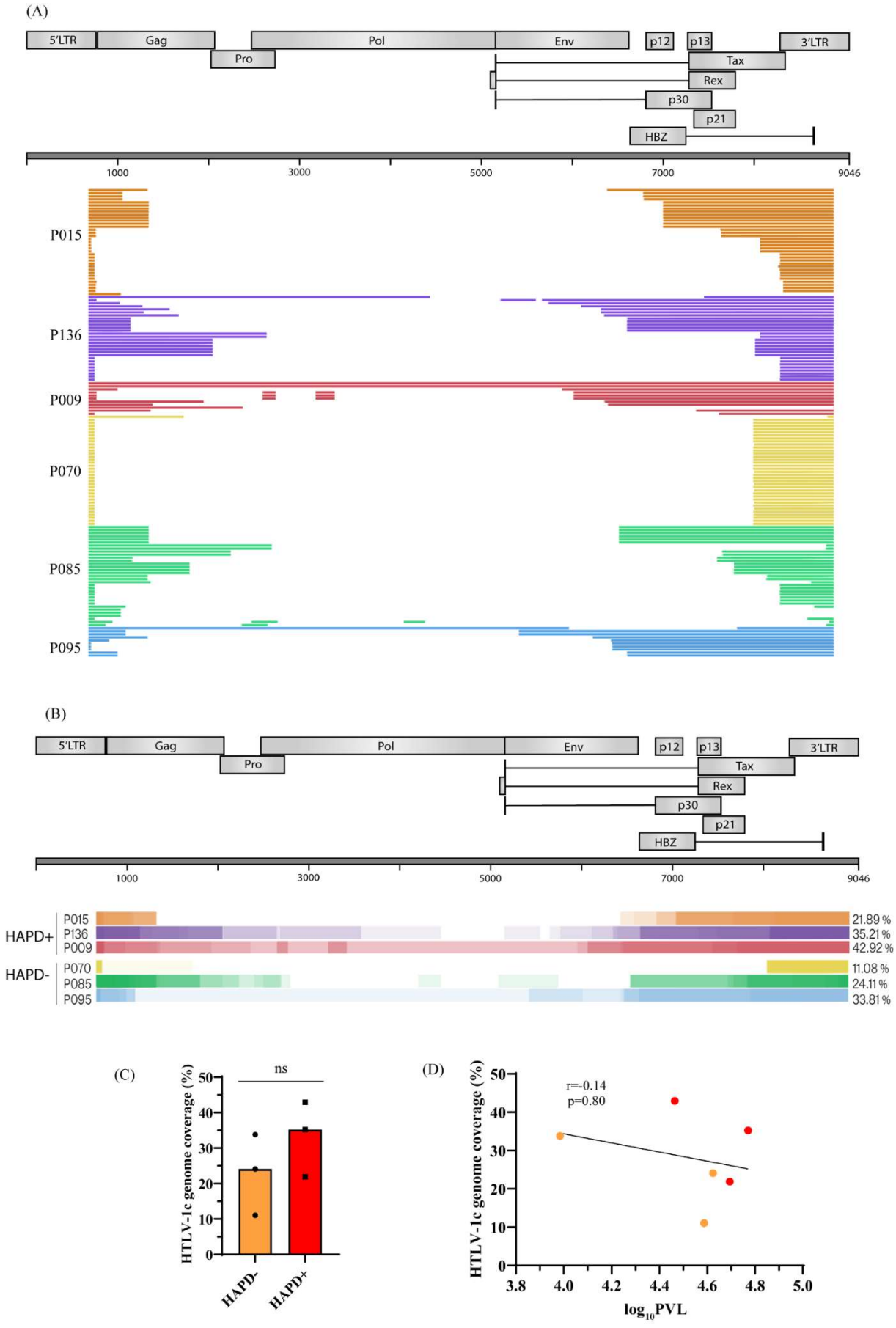

(A)
5'LTR
Gag
Pro
Pol
Env
p12
p13
3'LTR
Tax
Rex
p30
p21
HBZ
1000
3000
5000
7000
9046
P015
P136
P009
P070
P085
P095
(B)
HAPD+
HAPD-
21.89 %
35.21 %
42.92 %
11.08 %
24.11 %
33.81 %
(C)
ns
HTLV-1c genome coverage (%)
HAPD-
HAPD+
(D)
r=-0.14
p=0.80
log10PVL

Figure 4.10 – HAPD+ patients have higher overall HTLV-1c proviral genome coverage in PBMCs than HAPD- donors. gDNA was extracted from primary patient PBMCs (n, 6) and SPA assay was performed by limiting dilution touchdown nested PCR. Individual proviral amplicons were sequenced with the ONT platform and aligned to the HTLV-1c consensus genome. HTLV-1c PVL of each donor was determined by ddPCR. (A) All proviral sequences amplified from bulk PBMCs, stratified by donor, aligned to the HTLV-1c consensus genome. (B) Average genomic coverage of proviral landscape from primary human PBMCs, stratified by donor. Darker colour represents higher retention. (C) Summary of HTLV-1c genome coverage of proviral landscape in PBMCs of HAPD- and HAPD+ patients. Bar graphs show median genomic retention, individual points represent different donors. Statistical significance was assessed using Mann-Whitney test. (D) Correlation of percentage of HTLV-1c genome coverage to PVL. Black trend line represents nonlinear regression straight line of best fit and individual points represent different donors. Correlation significance was assessed using Spearman's correlation test.

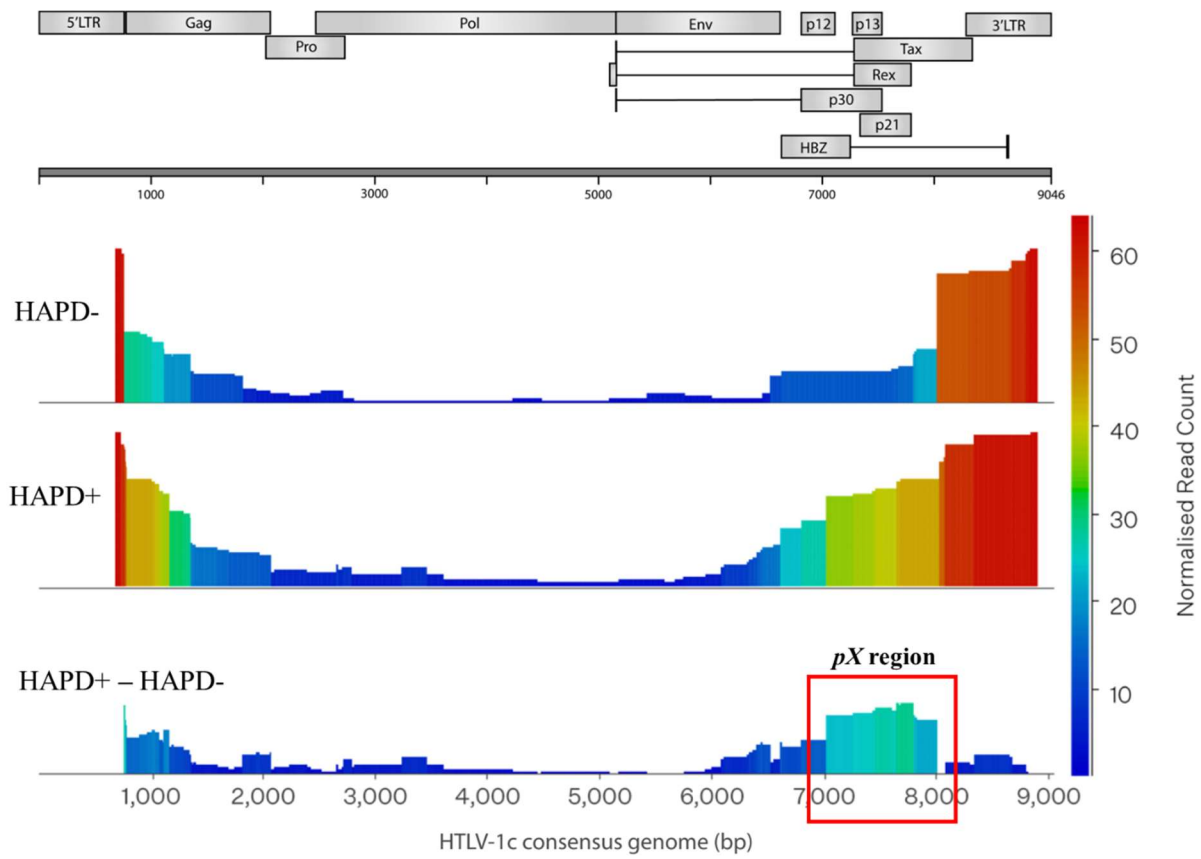


Figure 4.11 – The *pX* region of the HTLV-1c genome has higher retention in the total provirus pool of HAPD+ donors than HAPD-. gDNA was extracted from primary patient PBMCs and SPA assay was performed by limiting dilution touchdown nested PCR. Individual proviral amplicons were sequenced with the ONT platform and aligned to the HTLV-1c consensus genome. The normalised read count of each nucleotide across the HTLV-1c genome in the overall provirus pool of PBMCs was stratified by HAPD- and HAPD+ patients, and the difference calculated between the two groups. HTLV-1c-hg chimeric proviruses were excluded from this analysis.

4.4.4 Exclusively full-length HTLV-1c provirus detected in CCR4- reservoir

The enrichment of HTLV-1c provirus in the expanded phenotypic lung homing cells, and the non-expanded FOXP3 proxy and CCR4- phenotypes, demonstrated earlier in this chapter, prompted the investigation of the proviral landscape at high resolution in these reservoirs. We next mapped the provirus sequences amplified from each of these phenotypes to the HTLV-1c consensus genome, from two donors. The FOXP3+ proxy phenotype had highly defective provirus sequences in both P136 (HAPD+) and P085 (HAPD-) donors, however P136 showed the greatest retention of the *pX* region (**Figure 4.12A**). There were common provirus sequences within donors, but none detected between donors (**Figure 4.12A**). Overall, the proviral reservoir in the FOXP3+ proxy phenotype had greater coverage in the Bex donor P136 (mean, 28.38%; p, 0.17) than the HAPD- donor P085 (mean, 20.78%) (**Figure 4.12B, 4.12C**).

Similarly, defective proviruses were present in both donors in the lung homing proxy CD4+ T cell phenotype (**Figure 4.13A**). Many provirus clones in both P136 (HAPD+) and P085 (HAPD-) retained some, or all, of the *gag* and *pX* coding regions (**Figure 4.13A**), however P136 also retained some coverage in the *pol* and *env* coding regions, while P085 did not (**Figure 4.13A**). Consistently, replicates of the same proviral sequences were detected within donors but not between donors (**Figure 4.13A**). Reflecting the same trend in each proviral reservoir, the lung homing proxy phenotype had greater HTLV-1c genomic coverage in P136 (HAPD+) (mean, 35.56%; p, 0.08) than P085 (HAPD-) (mean, 27.06%) (**Figure 4.13B, 4.13C**).

Surprisingly, in the CCR4- reservoir of P136, we detected only full-length provirus (**Figure 4.14**). Each of these seven proviral sequences have over 99.3% homology with the HTLV-1c consensus sequence (**Figure 4.14**). Unfortunately, we were unable to amplify any provirus from the CCR4- phenotypic cells of P085 due to a lower number of cells processed into the CCR4- reservoir by FACS, and downstream low gDNA yield.

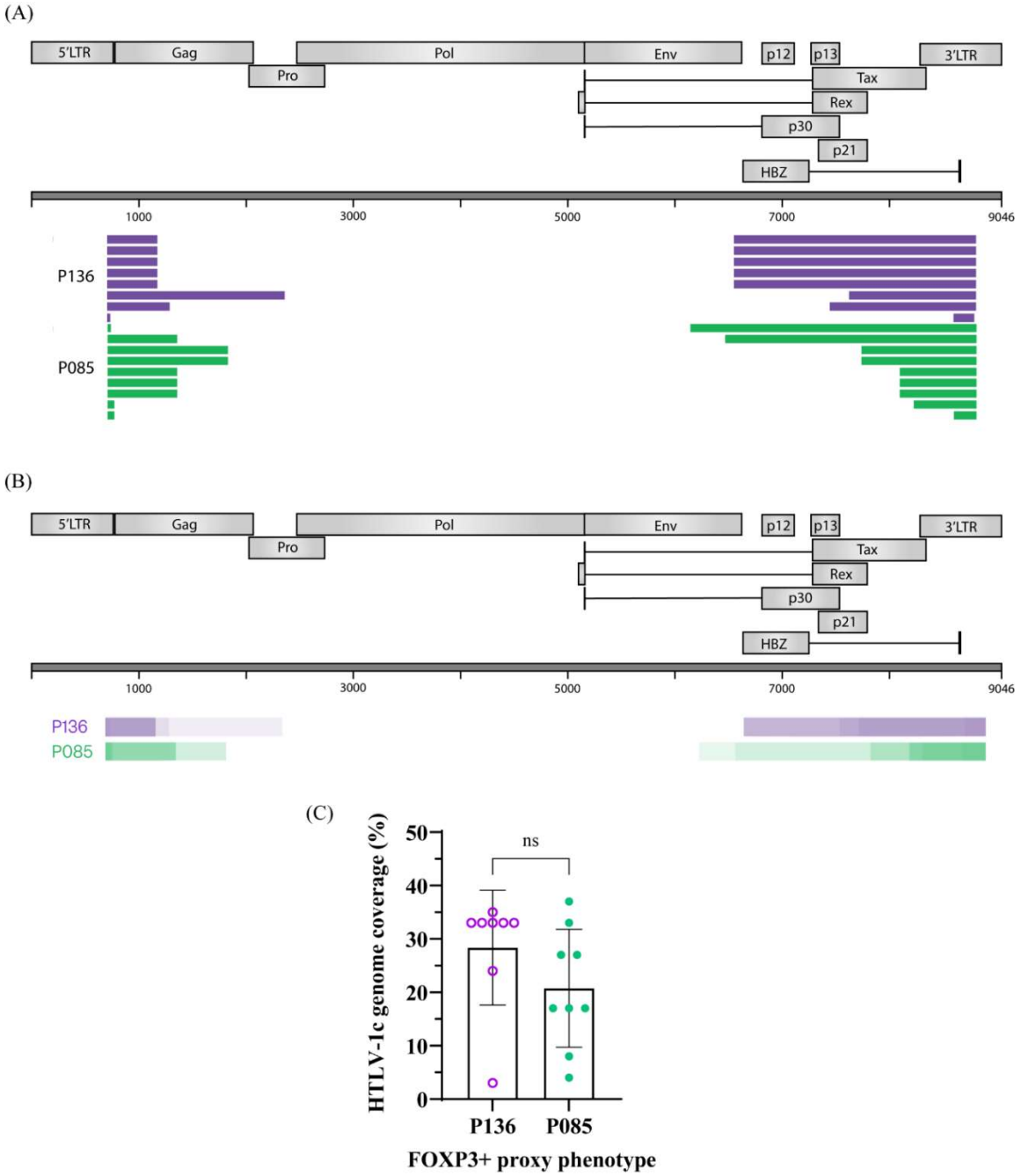


Figure 4.12 – The HTLV-1c proviral landscape is highly defective in FOXP3+ proxy CD4+ T cell phenotype reservoir of primary human PBMCs. Primary patient PBMCs were sorted by FACS into various CD4+ T cell phenotypes, including FOXP3+ proxy (CCR4+CD49d-CD127-). gDNA was extracted from these reservoirs and SPA assay was performed by limiting dilution touchdown nested PCR. Individual proviral amplicons were sequenced with the ONT platform and aligned to the HTLV-1c consensus genome. (A) All proviral sequences amplified from the FOXP3+ proxy phenotype, stratified by donor. (B) Genomic coverage of proviral landscape from FOXP3+ proxy phenotype, stratified by donor; darker colour represents higher retention. (C) Summary of average genome coverage of the proviral landscape in FOXP3+ proxy phenotype. Bar graphs show mean genomic retention with SD error bars. Individual points represent each proviral sequence. Statistical significance was assessed using unpaired t test. (n, 2; P136, HAPD+; P085, HAPD-).

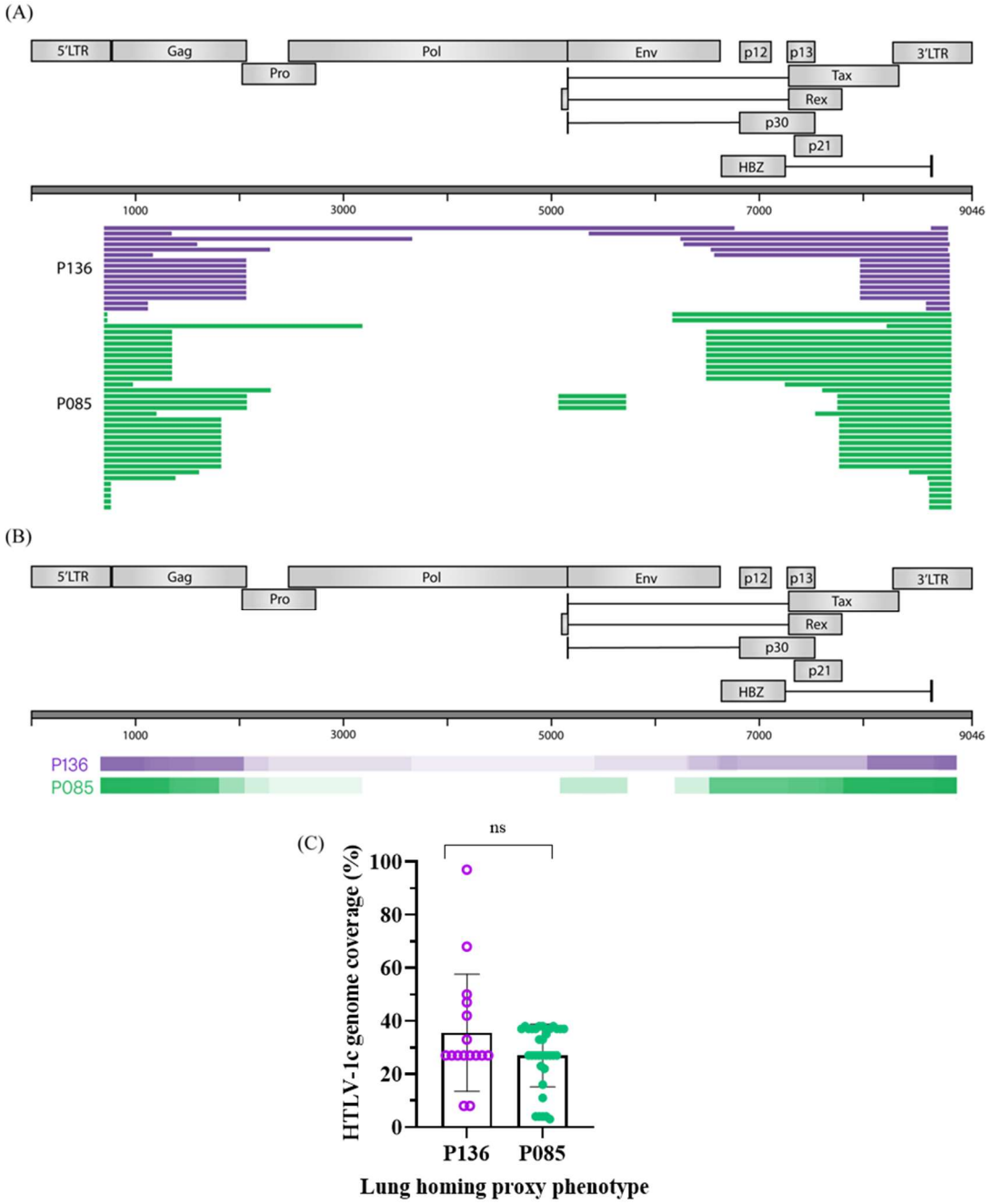


Figure 4.13 –The HTLV-1c proviral landscape is highly defective in lung homing proxy CD4+ T cell phenotype reservoir of primary human PBMCs. Primary patient PBMCs were sorted by FACS into various CD4+ T cell phenotypes, including lung homing proxy (CCR4+CD49d+Integrin-β7-). gDNA was extracted from these reservoirs and SPA assay was performed by limiting dilution touchdown nested PCR. Individual proviral amplicons were sequenced with the ONT platform and aligned to the HTLV-1c consensus genome. (A) All proviral sequences amplified from lung homing proxy phenotype, stratified by donor. (B) Genomic coverage of proviral landscape from lung homing proxy phenotype, stratified by donor; darker colour represents higher retention. (C) Summary of average genome coverage of the proviral landscape in lung homing proxy phenotype. Bar graphs show mean genomic retention with SD error bars. Individual points represent each proviral sequence. Statistical significance was assessed using unpaired t test. (n, 2; P136, HAPD+; P085, HAPD-).

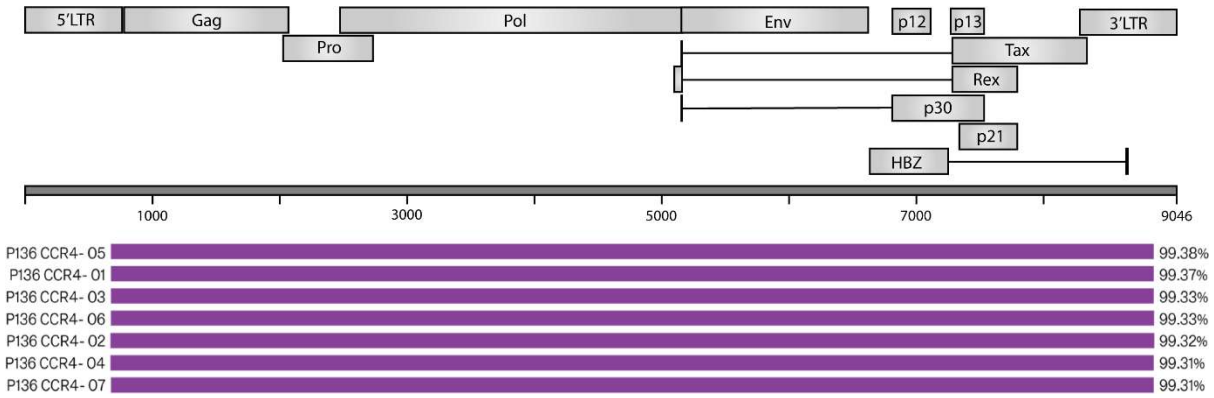


Figure 4.14 – HTLV-1c provirus sequences in CCR4- effector/memory CD4+ T cells of primary human PBMCs are full-length. Primary patient PBMCs were sorted by FACS into various CD4+ T cell phenotypes, including CCR4- E/M (CD45RA-) CD4+ T cells. gDNA was extracted from these reservoirs and SPA assay was performed by limiting dilution touchdown nested PCR. Individual proviral amplicons were sequenced with the ONT platform and aligned to the HTLV-1c consensus genome. All proviral sequences amplified from CCR4- immunophenotype reservoir of P136 (HAPD+) are shown. Percentage shows homology of full-length proviral sequences to HTLV-1c consensus genome.

4.4.5 HTLV-1c defective provirus breakpoints share LTR-like motifs

With a large proportion of the overall HTLV-1c proviral landscape harbouring large internal deletions, we next examined the sequences of the individual defective provirus break points to determine if there were any common motifs present to explain how these deletions may be occurring. Across the board, recurring breakpoints occurred at the 3' end junction of the 5' LTR (758 bp), and likewise, at the 5' end junction of the 3' LTR (8289 bp). In the defective provirus amplified from bulk PBMCs, there were 37 clones with a breakpoint at the 5' LTR, and 1 clone with a breakpoint occurring at the 3' LTR junction. The proviral sequences of the FOXP3+ proxy immunophenotype reservoir had 1 breakpoint occur within 2 nucleotides of the 5' LTR junction (760bp). Similarly, there were 5 defective proviruses with a breakpoint within 2 nucleotides of the 5' LTR (760bp), in the lung homing proxy reservoir.

As the most frequent break points occurred around the LTR junctions (**Figure 4.15A**), we then examined all defective break points that occurred outside of the LTR junction points, exclusive of chimeric and inversion proviruses. We selected the last 30 bases of

the 5’ breakpoint junction, and the first 30 bases of the 3’ breakpoint junction, of every individual provirus sequence, and filtered to remove those sequences which occurred at the LTR junctions themselves (**Figure 4.15B**). We generated a DNA sequence logo to represent the conserved nucleotides across all filtered proviral sequence breakpoints (**Figure 4.15B**). Interestingly, the breakpoints in fact retain sequence homology with the LTRs themselves, as the most conserved nucleotides generate a similar LTR sequence motif.

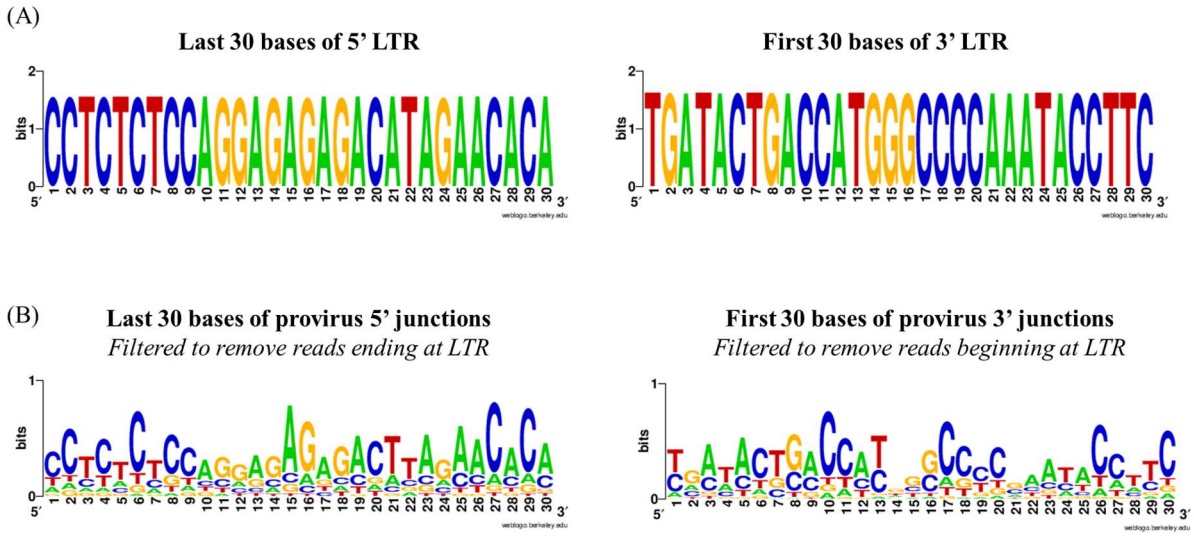


Figure 4.15 – Defective HTLV-1c proviruses from primary human PBMCs display breakpoints with LTR-like motifs. gDNA was extracted from primary human PBMCs infected with HTLV-1c and SPA assay was performed by limiting dilution touchdown nested PCR. Individual proviral amplicons were sequenced with the ONT platform and aligned to the HTLV-1c consensus genome. Sequences flanking the breakpoints of the defective provirus with large internal deletions were selected. Nucleotide signals for all sequences are represented as sequence logos. (A) Sequence of the last 30 bases of the 5’- and the first 30 bases of the 3’- HTLV-1c LTRs. (B) Sequence of the last 30 bases of the 5’- and the first 30 bases of the 3’- HTLV-1c defective provirus segment, filtered to remove any breakpoints that occur at the LTR junctions, for proviruses amplified from primary human PBMCs.

4.4.6 Chimeric HTLV-1c-human genome provirus sequences occur most commonly in centromeric regions

Most surprising and unique was the detection of chimeric defective HTLV-1c proviruses with internal sequences mapping to the hg, making up 21.5% of the sequenced proviral landscape (**Figure 4.9A**). Most of these chimeras were comprised of human genome sequence flanked by the HTLV-1c LTRs only (**Figure 4.16A**). However, there were a few chimeric proviruses which included some HTLV-1c coding sequence: in particular, one containing the entire *pX* region (**Figure 4.16A**).

The chimeric provirus sequences are characterised in **Table 4.2**. Due to the sovereignty of Indigenous genomic information, we have only provided the location of the hg where these sequences are located, and the size of the sequence. The size of the internal hg sequences ranged from 107-3262 bp and were distributed across the genome and chromosomes (**Figure 4.16B**). Most commonly, the hg sequences originated from regions with a repetitive nature: 58.7% of the chimeric human sequence was found in centromeric repeats, while other non-centromeric internal sequences contained an *Alu* element, another kind of repetitive sequence (**Figure 4.16C**). 6.5% of the chimeras contained an intergenic region, while 34.8% contained gene coding regions (**Figure 4.16C**). 7 of these provirus chimeras contain the gene region encoding a replication dependent histone, H2BC12-201 (**Table 4.2**). 2 chimeras contain a coding region for a redox reactive protein of the thioredoxin superfamily, NXN (**Table 4.2**). Other human genes identified in chimeric proviruses included a G protein-coupled inwardly-rectifying potassium channel, KCNJ6; an ephrin type-A receptor, EPHA6; a succinate dehydrogenase complex subunit C (SDHC), AL592295; a RNA recognition and binding protein, LARP1B; a long non-protein coding RNA, Linc02119; an *Alu* SINE monomer, FLAM_C; and an uncharacterised non-coding RNA, XR (**Table 4.2**).

Given the internal hg sequences of the chimeric proviruses often originated from regions with a repetitive nature, we investigated the human DNA sequence patterns at the chimera junctions. We created sequence logos to characterise the first (**Figure 4.17A**) and last 30bp (**Figure 4.17B**) of the hg component. However, no strong sequence motifs were revealed for the hg DNA sequence of the chimeric provirus junction.

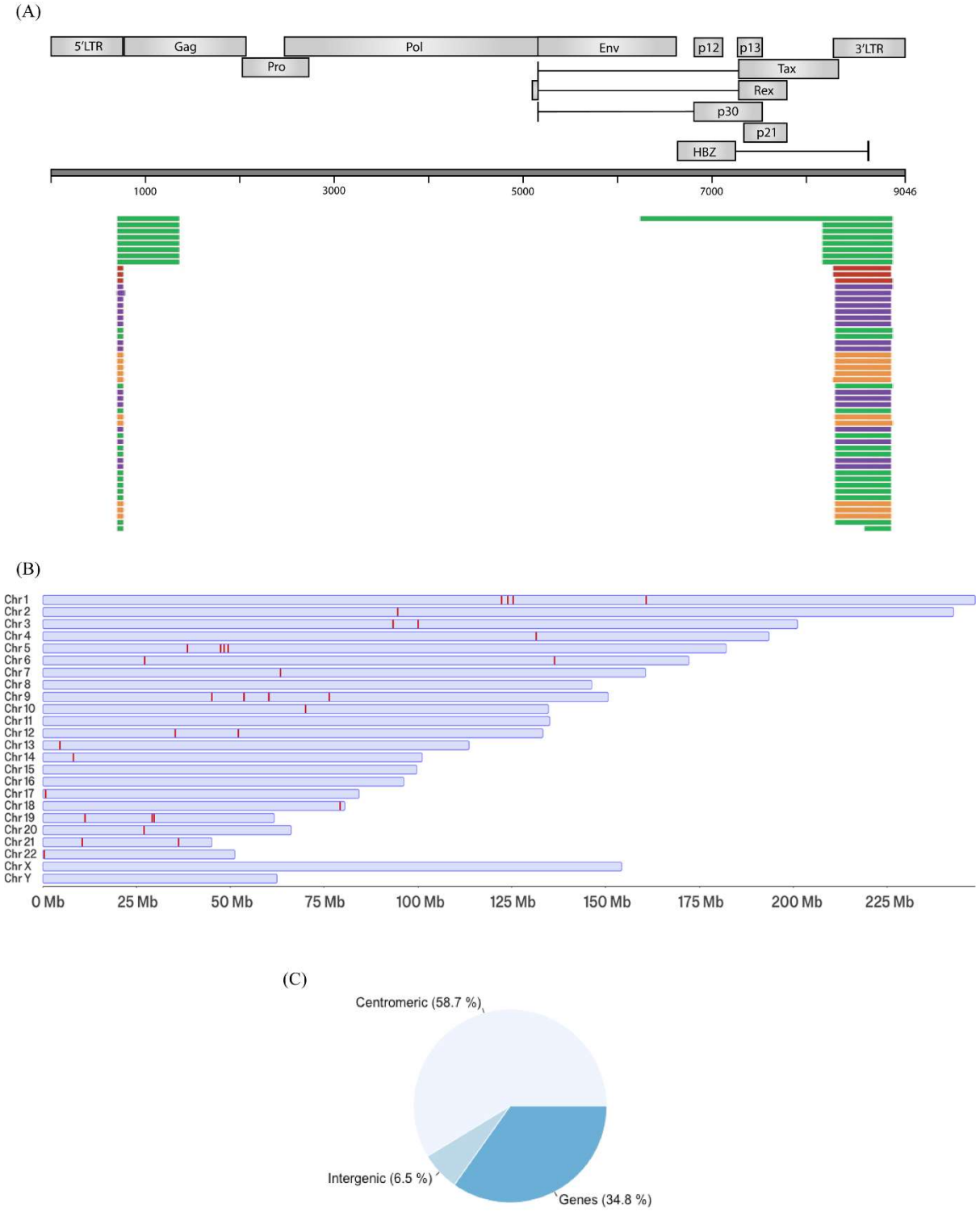


Figure 4.16 – Novel HTLV-1c-human genome chimeric proviruses were amplified from primary human PBMCs. (A) The chimeric provirus sequences aligned to the HTLV-1c genome, containing internal hg sequences (not mapped to HTLV-1c genome). (B) Map of HTLV-1c-human chimeric proviruses to location in hg, indicated by red line. (C) Overall frequencies of HTLV-1c-hg chimeric proviral sequence categories containing: human centromeric, intergenic or gene regions.

Table 4.2 – Summary and annotations of HTLV-1c-human genome chimeric proviruses amplified from primary human PBMCs.

Provirus amplicon	5' HTLV-1c size (bp)	3' HTLV-1c size (bp)	Internal hg size (bp)	hg Coordinates	Repetitive Sequences	Genes	Additional Notes
P009_PBMCs_04	28	82	2022	chr5:38,384,072-38,385,045	hAT-Ac:UCON26;	NR_147009.1; LINC02119	-
P015_PBMCs_11	80	608	918	chr6:136,295,426-136,296,421	AluSx1	XR_007077156.1	-
P015_PBMCs_15	76	606	1077	chr9:52,713,003-54,643,865	hsat3_9_3(B5)	-	Centromeric satellite repeats
P015_PBMCs_23	79	608	327	chr22:280,478-280,803	censat_22_4	-	Centromeric satellite repeats
P015_PBMCs_24	79	608	327	chr22:280,478-280,803	censat_22_4	-	Centromeric satellite repeats
P015_PBMCs_31	79	605	1366	chr19:29,623,263-29,624,964	hor_19_13(S1C1/5/19H1L)	-	Active αSat HOR
P015_PBMCs_32	79	618	1143	chr4:131,418,224-131,419,270	AluSg	LARP1B: NG_047161.1	Over exon
P015_PBMCs_35	79	581	495	chr10:69,919,126-69,919,549	AluY	FLAM_C	-
P015_PBMCs_36	79	608	1009	chr12:52,028,665-52,029,608	AluSg;AluSp;L2C;MLT_24B	-	Intergenic
P015_PBMCs_37	79	608	676	chr21:10,304,525-10,305,134	hsat3_21_17(A4)	-	Classical Human Satellite
P015_PBMCs_38	609	76	1025	chr5:48,143,933-48,144,958	hor_5_6(S1C1/5/19H1L	-	Active αSat HOR
P085_Lung_03	80	609	585	chr18:79,069,536-79,070,500	-	-	No annotations
P085_Lung_27	71	607	688	chr1:125,124,902-125,125,597	hor_1_5(S1C1/5/19H1L)	-	Active αSat HOR
P085_Lung_28	78	606	681	chr1:125,124,902-125,125,587	hor_1_5(S1C1/5/19H1L)	-	Active αSat HOR
P085_Lung_32	80	609	701	chr18:79,069,746-79,070,445	-	-	P085_Lung_03
P085_Lung_34	75	607	686	chr5:47,289,253-47,289,938	ALR/Alpha; hor_5_6(S1C1/5/19H1L)	-	Active αSat HOR
P085_Lung_36	75	625	261	No significant alignment	-	-	hor_19_13(S1C1/5/19H1L)
P085_Lung_06	77	606	686	chr20:26,932,500-26,933,186	hor_20_2(S2C20H1L)	-	Active αSat HOR
P085_PBMC_05	76	606	360	No significant alignment	hsat3	-	Classical Human Satellite
P085_PBMC_08	80	294	107	chr3:93,170,446-93,170,554	hsat1A_3_2	-	Classical Human Satellite
P085_PBMC_09	78	596	695	chr19:28,976,395-28,977,080	hor_19_13(S1C1/5/19H1L)	-	Active αSat HOR
P085_PBMC_20	661	741	392	chr6:27,015,255-27,015,659	-	H2BC12-201	Exon1; CGI
P085_PBMC_24	663	741	389	chr6:27,015,255-27,015,659	-	H2BC12-201	Exon1; CGI
P085_PBMC_25	80	826	1138	chr5:49,351,431-49,352,546	hor_5_6(S1C1/5/19H1L)	-	Active αSat HOR
P085_PBMC_36	77	608	668	chr9:76,145,867-76,146,500	hsat3_9_3(B5)	-	Classical Human Satellite
P085_PBMC_38	73	605	1402	No significant alignment	hsat3_9_3(B5)	-	Classical Human Satellite
P085_FOXP3_04	660	741	391	chr6:27,015,255-27,015,659	-	H2BC12-201	Exon1; CGI
P085_FOXP3_06	662	741	392	chr6:27,015,255-27,015,659	-	H2BC12-201	Exon1; CGI

P085_FOXP3_07	664	741	388	chr6:27,015,255-27,015,659	-	H2BC12-201	Exon1; CGI
P085_FOXP3_08	663	741	390	chr6:27,015,255-27,015,659	-	H2BC12-201	Exon1; CGI
P085_FOXP3_10	77	607	346	chr19:29,142,203-29,142,548	hor_19_13(S1C1/5/19H1L)	-	Active αSat HOR
P085_FOXP3_11	656	741	393	chr6:27,015,255-27,015,659	-	H2BC12-201	Exon1; CGI
P095_PBMCs_04	80	614	3262	chr17:703,650-706,034	AluY;AluSx1;AluJb;MIRb	NXN_201; NM_022463.5	Intron/exon
P095_PBMCs_06	659	2672	435	chr21:36,042,534-36,042,965	Charlie2b;L1MEd	KCNJ6-201; NM_002240.5	Intron 3
P095_PBMCs_13	80	614	2048	chr17:706,029-707,205	L1ME1; AluY; AluJb	NXN	Intron 4
P095_PBMCs_19	65	611	351	chr1:122,173,215-122,173,560	hor_1_5(S1C1/5/19H1L)	-	Active αSat HOR
P136_Lung_03	79	607	1095	chr1:160,563,555-160,564,457	L1M5; AluYh7	AlAL592295.3-201	lncRNA
P136_Lung_16	79	609	1365	chr12:35,200,977-35,202,340	hor_12_2(S1C12H1L)	-	Active αSat HOR
P136_PBMCs_02	79	607	2268	chr3:99,929,608-99,931,291	MER1113; AmnSINE2; AluSQ; MER72	EPHA6-202, NM_001080448.3	exon 3, intron 3/4: Centromeric transition regions
P136_PBMCs_09	80	609	1138	chr13:4,494,711-4,495,852	hsat1A_13_1	-	Classical Human Satellite
P136_PBMCs_10	80	609	1138	chr13:4,494,711-4,495,852	hsat1A_13_1	-	Classical Human Satellite
P136_PBMCs_12	79	607	687	chr2:94,605,711-94,606,396	hor_2_4(S2C2H1L)	-	Active αSat HOR
P136_PBMCs_17	76	607	1029	chr7:63,189,440-63,190,468	hor_7_2(S1C7H1L)	-	Active αSat HOR
P136_PBMCs_23	79	608	2216	chr9:44,979,221-44,980,903	hor_9_1(S2C9H1L)	-	Active αSat HOR
P136_PBMCs_25	80	610	683	chr19:11,065,608-11,066,100	AluY;AluSx1	-	Intergenic
P136_PBMCs_26	79	608	1026	chr1:123,725,408-123,726,434	hor_1_5(S1C1/5/19H1L)	-	Active αSat HOR
P136_PBMCs_27	80	607	1714	chr14:8,106,933-8,131,391	alpha satellite repeats	-	No annotations
P136_PBMCs_30	79	607	1366	chr2:94,598,910-94,600,275	hor_2_4(S2C2H1L)	-	Active αSat HOR; enhancer[366]
P136_PBMCs_33	76	608	955	chr9:60,073,368-60,074,352	hsat3_9_3(B5)	-	Classical Human Satellite
P136_FOXP3_04	79	607	687	chr2:94,605,711-94,606,396	hor_2_4(S2C2H1L)	-	Active αSat HOR
P136_FOXP3_09	79	608	1366	chr2:94,435,266-94,436,632	hor_2_4(S2C2H1L)	-	Active αSat HOR
P136_FOXP3_11	79	608	687	chr2:94,605,711-94,606,396	hor_2_4(S2C2H1L)	-	Active αSat HOR

Chr, chromosome; HOR, higher order repeat; hSat, human satellite; lncRNA, long non-coding RNA; censat, centromeric satellite.

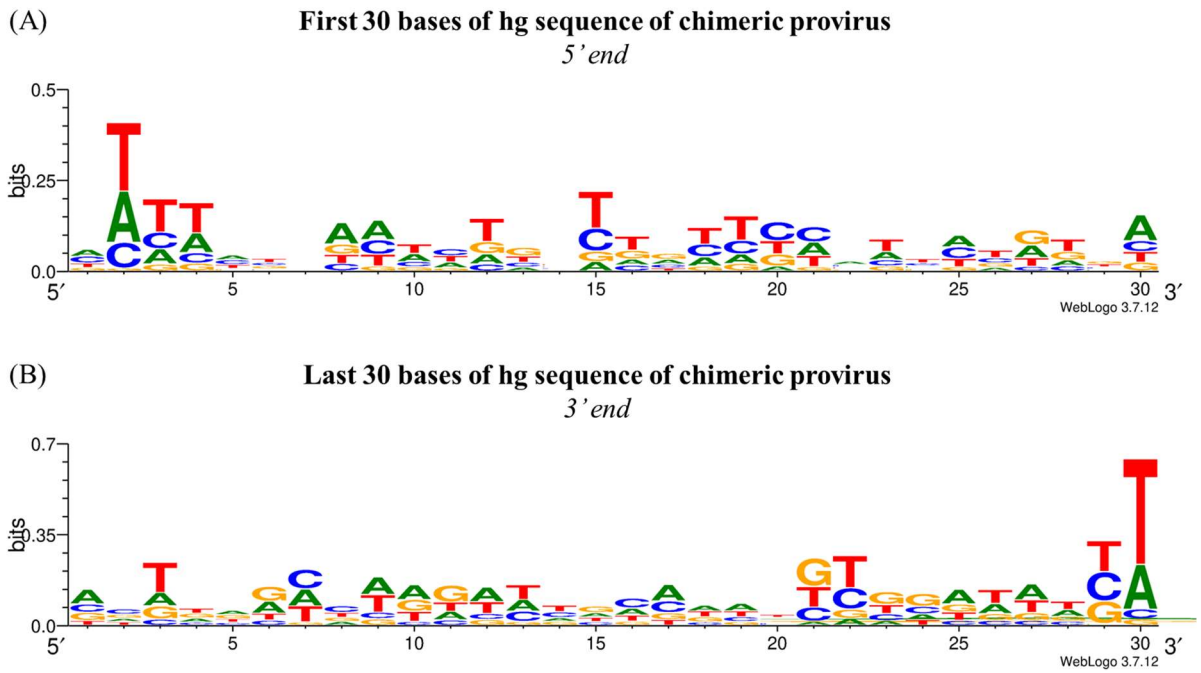


Figure 4.17 – Junction sequences of human genome component of novel HTLV-1c chimeric proviruses from primary patient PBMCs. gDNA was extracted from primary human PBMCs infected with HTLV-1c and SPA assay was performed by limiting dilution touchdown nested PCR. Individual proviral amplicons were sequenced with the ONT platform and aligned to the HTLV-1c consensus genome. The hg sequences of the defective provirus with large internal deletions were selected. Nucleotide signals for all sequences are represented as sequence logos. (A) First 30 bases at the 5' end of the internal hg sequence of HTLV-1c chimeric provirus. (B) Last 30 bases of the 3' end of the internal hg sequence of HTLV-1c chimeric provirus.

4.5 Chapter four discussion

4.5.1 Novel characterisation of the phenotype of HTLV-1c infected cells

While the phenotype of HTLV-1a infected cells and consequential impacts on frequency of circulating immune cells has been investigated extensively since HTLV-1 was discovered, the understanding of subtype-C infection has remained neglected. This is a unique characterisation of the phenotype of HTLV-1c infected cells, including a novel focus on subtype-C HAPD.

HTLV-1c infection in First Nations people in Central Australia may manifest differences in the host-virus interplay which impacts the repertoire of circulating immune cells, as opposed to the more global subtype-A. HTLV-1a infection results in decreased frequency

of circulating naïve CD4+ T cells, potentially due to an interference with thymic function [34, 367]. Moreover, infection with HTLV-1a increases the frequency of circulating memory T lymphocytes, proportional to the PVL [34]. However, our characterisation of HTLV-1c infection tells a different story: there is no virus-induced impact on naïve CD4+ T cell frequency in HTLV-1c infection. In addition, HTLV-1c infected donors and uninfected community controls from Central Australia likely experience frequent antigenic stimulation from a plethora of co-infections such as BSI from gut-associated bacteria, which could result in higher levels of naïve immune cell differentiation into effector/memory phenotypes.

Interestingly, HTLV-1c infection may also impart a distinct effect, as compared to subtype-A, on circulating FOXP3+CD4+ T cells. HTLV-1a infection results in an expansion of FOXP3+CD4+ T cells in ACs, HAM and ATL donors, and those with *S. stercoralis* co-infection, and this expansion is positively correlated to PVL [355]. However, we reported that HTLV-1c infection does not impact the proportion of circulating FOXP3+ proxy phenotype CD4+ T cells in PBMCs. This can be confirmed by using internal FOXP3 staining in future experiments. Nevertheless, these cells in HTLV-1c infection are indeed highly infected with provirus, and a large proportion of these cells express activation markers, both of which may impact the functionality of these cells, which typically have immune modulatory and suppressive functions as Tregs. It may be that in the setting of HTLV-1c infection in Central Australia, Tregs become exhausted and lose their functional capacity. So, while the FOXP3+ proxy phenotype may not be the primary expanded reservoir in subtype-C infection, these cells are still involved in chronic immune activation and dysfunction, likely playing a role in overall pathogenesis.

Perhaps one of the most important findings of this study is the characterisation of the highly expanded, activated, and infected, lung homing proxy phenotype of the E/M CD4+ T cells in HTLV-1c infection. These findings indicate HTLV-1c infection is driving the proliferation and activation of CD4+ T cells that are trafficking to the respiratory tract, and therefore these cells may play a role in pathogenesis, chronic inflammation, tissue damage in the lungs, and ultimately progression to HAPD such as Bex. Moreover, these results advance and support other reports surrounding the lung involvement in HTLV-1 infection (reviewed in [128]): lymphocyte infiltration of lung tissues has been reported to occur during infection [131, 351, 352], HTLV-1a mRNA

expression has been detected in BALF [132, 133], and HTLV-1c DNA was found in the sputum of infected donors suffering from associated bronchiectasis [218].

Unexpectedly, we found the CCR4- phenotype of E/M CD4+ T cells were enriched with HTLV-1c provirus, and a significant proportion of these cells expressed activation markers. This provides some insight as to why clinical trials using mogamulizumab, the anti-CCR4 mAb treatment, did not sufficiently decrease the proviral burden in HTLV-1a patients across the board [110, 111, 344, 345]. Moreover, there may be future implications for trialling mogamulizumab treatment in HTLV-1c infected persons in Central Australia: it would be ineffective at targeting the complete proviral reservoir. Basic research into the HTLV-1c proviral reservoir needs to continue in order to understand how to tackle the infection and develop effective therapeutic strategies and treatments.

While this preliminary study of characterising the immunophenotype of HTLV-1c infected cells made some important advances in understanding the nature of subtype-C infection, further studies should incorporate a larger cohort of uninfected and infected donors. The viability of some samples excluded them from analysis, and so improving the nature and workflow of blood collection, PBMC isolation and storage/cryopreservation in Alice Springs would allow more in-depth analysis of immune cells, including reservoirs not explored such as other T cells or B cells. Additionally, it would be pertinent to include donors infected with HTLV-1a in future studies, including ACs, and ATL and HAM donors, to provide a direct comparison between subtypes and disease states. Unfortunately, we did not have access to donors infected with this subtype, as HTLV-1c is almost exclusively predominant in Central Australia. Nonetheless, the immunophenotype characterisation of HTLV-1c infected cells in First Nations people made significant advances in our understanding of HTLV-1 infection in Central Australia.

4.5.2 Detection of multiple HTLV-1c gene coding regions with ddPCR prompted redesign of HTLV-1 PVL assay in Australia

The previously designed HTLV-1c PVL assay, detecting *gag* or *tax* genomic regions [218], demonstrated that these DNA targets were often present at different levels in bulk PBMCs of infected donors (Hirons et al., unpublished). This led us to incorporate two

other genomic regions: structural gene encoding region for *env*, and regulatory gene encoding region for *hbz,* into the ddPCR assay repertoire.

Overall, *hbz* and *tax* gene regions were most frequently retained in the proviral reservoir across all CD4+ effector/memory phenotypes and bulk PBMCs, while *gag* and *env* gene regions were more frequently deleted. Not surprisingly, this suggests the expression of regulatory genes *tax* and *hbz* play a more pivotal role than the structural genes in HTLV-1c infection and pathogenesis.

A key advantage of this ddPCR assay is that it will detect the genomic target regardless of any deletions in the LTR regions. Furthermore, it is a quick and easy method that can use isolated gDNA to determine the frequency of gene retention and deletion, and potentially the risk for pathogenesis and disease progression. A limitation of this assay is in the normalisation of the probe detections: we are making assumptions about the nature of the provirus such that the highest detected genomic region for a given sample accounts for all integrated proviruses, when in fact this may not exactly be the case. Further studies of these ddPCR quantification methods for gene retention frequency with a larger cohort and longitudinal sampling could elucidate its effectiveness for use as a crude method for monitoring HTLV-1c infection.

Nevertheless, these critical findings prompted an update to the current qPCR targeting *gag* as the PVL assay for HTLV-1c used in Australia in reference laboratories. In collaboration with the NRL, a new qPCR PVL assay targeting the *pX* region is in the process of being developed (Vandegraaff et al., unpublished), which could increase the sensitivity of the assay and provide more thorough monitoring of infection across Australia.

4.5.3 SPA and ONT sequencing methodology provides high resolution of HTLV-1c proviruses

Finding a deeper comprehension of the proviral reservoir is a key component of understanding the host-virus relationship in HTLV-1c infection, and so we set about sequencing individual integrated HTLV-1c proviruses. This thesis chapter built upon an initial HTLV-1c sequencing study that characterised an HTLV-1c consensus genome from Central Australia (Hirons et al., unpublished). Bulk fragment PCR Sanger sequencing with gDNA from HTLV-1c infected donors successfully determined the

sequence of expected amplicons, stitched together to generate a consensus genome sequence (Hirons et al., unpublished). However, this study had several limitations: it did not provide resolution of the sequence or frequency of individual provirus clones, deletion junctions were not determined due to sequencing only of expected fragment size, and a combined provirus sequence that appears to be full-length may in fact just be many different proviruses that together have full coverage. Implementing the SPA assay allowed us to overcome these limitations and further our understanding of the HTLV-1c proviral reservoir, by giving nucleotide resolution of each individual provirus clone. However, the SPA assay is limited in amplifying proviruses that retain the 5'- and 3'-LTR primer binding sites: it will not detect integrated provirus with complete LTR deletions. Therefore, we may miss detection of some defective clones that skews our analysis from true population representation. Furthermore, sequencing clones from the SPA assay unfortunately does not determine the integration sites of each clone: we designed this assay purposefully and carefully to align with our ethics, which honours the sovereignty of First Nations genomic information. However, given the importance of the host-virus interplay in HTLV-1c infection, integration site characterisation and epigenetic mapping will be vital to fully grasp the complexity of this lifelong infection, and determine what is driving proliferation and persistence that contributes to pathogenesis.

We selected the ONT platform for the third generation sequencing of individual proviruses as it has very long read capability [368], so the entire HTLV-1c proviral genome could be sequenced in each read, if full-length proviruses were amplified. We used a ligation barcode kit to generate the libraries, ensuring no PCR errors were introduced to the provirus sequences at this stage. Given ONT sequencing has a higher error rate than other methods such as short read Illumina [368], we used 1000 reads for each individual provirus, where possible, to generate consensus sequences before aligning to the HTLV-1c genome, to eradicate any potential ONT-associated sequencing errors.

Previous HTLV-1 sequencing studies were limited in their in-depth understanding of the nature of individual proviruses in circulation and their implications in infection. The strategy developed and techniques applied in this chapter resulted in a high-resolution characterisation of the HTLV-1 proviral reservoir, and advanced our understanding of

how it may impact the host immune function, cause pathogenesis and trigger disease development.

4.5.4 HTLV-1c provirus graveyard with large internal deletions is capable of driving pathogenesis, while CCR4- cells maintain infection

This thesis chapter presented novel characterisation of the HTLV-1c proviral landscape in six human donors from Central Australia using long-read sequencing. Firstly, we confirmed the presence of very high levels of defective provirus in all HTLV-1c infected individuals, much greater than what was previously reported for subtype-A [152], which could be due to advanced methodologies, the host immune response, or the virus. Even without being replication competent, these defective HTLV-1c proviruses are still capable of driving viral-induced pathogenesis through extensive immune activation, inflammation and dysfunction in the host. Therefore, it is likely that, alongside the proliferation of infected cell reservoirs, viral gene expression is still occurring from these defective proviruses. It is possible there are mRNA transcripts produced which utilise cryptic splice sties when major donors and acceptor sites are deleted. Indeed, in subtype-A infection, defective proviruses can still produce mRNA transcripts with large internal deletions, provided there were no premature stop codons [369]. Furthermore, host-HTLV-1c chimeric mRNA transcripts may be produced by defective proviruses, further contributing to pathogenesis, much like defective HTLV-1a proviruses with long open reading frames that produce host-virus chimeric proteins [369].

The striking retention of the *pX* region in HAPD+ patients when compared to HAPD- donors, including the entire sequence for *hbz* mRNA expression, suggests this is crucial in driving the associated pulmonary disease pathogenesis. The transcriptional activity of the 3' LTR and *hbz* gene may be an important advantage for HTLV-1c proviruses to retain for driving disease, mediated through HBZ protein and functional RNA. HBZ induces CCR4 expression to promote migration of T cells to sites of inflammation and increase cell proliferation [370], contributing to pathogenesis and irreversible lung tissue damage of the host. Furthermore, conservation of an HBZ viral enhancer which lies near the 3' end of the genome between nucleotides 8001-8160 [371] is important for the virus to maintain ongoing expression of the antisense strand. HBZ is also involved in altering cellular mRNA splicing processes by interacting with transcription factors to exclude exons, which may dysregulate the immune cell functional capacity [372].

In addition, our results indicate that donors with HTLV-1c-associated Bex may have higher levels of full-length provirus, and similarly, retain a greater proportion of the genome when defective, compared to HAPD-. Interestingly, this pattern is not reflective of what is reported for HTLV-1a infection: there are similar proportions of full-length provirus between HAM and AC donors, while ATL donors have higher levels of defective provirus [152]. Our results suggest that in HTLV-1c associated Bex there is a greater possibility for ongoing low levels of de novo infection that maintain the proviral reservoir via cell-to-cell transmission, alongside clonal proliferation. Furthermore, the intact or near-full length provirus found in Bex donors likely retain the capability to produce functional viral mRNAs and proteins, which may also contribute to inflammation, pathogenesis and disease development, whereas highly defective provirus may not, depending on where the deleted segments and junctions are located along the integrated genome. Moreover, HTLV-1c-associated Bex donors may experience an ineffective T cell response, leaving more intact or transcriptionally active proviruses in circulation, while HAPD- donors have a more effective response that is able to target cells with replication competent provirus and viral mRNA expression. In particular, the *pX* region products might be inefficiently targeted by the host response, especially in HAPD+ donors. Indeed, this may be influenced by the extent of T cell exhaustion in HAPD+ and HAPD- donors, something that needs to be further investigated in HTLV-1c. Other factors that may also play a role in the proviral landscape include integration site and interactions with the host genome, epigenetic modifications of the provirus, and variable expression of host proteins and cytokines. In fact, it might indeed be a combination of all these components that influences the proviral landscape variability across donors, and future studies should address these factors.

Intriguingly, the proviral reservoir in the CCR4- CD4+ T cell phenotype was composed entirely of full-length provirus, a phenomenon that that has not been reported in vivo. CCR4-CD45RA-CD4+ T cells may harbour a pivotal proviral reservoir that contributes to the lifelong persistence of HTLV-1c infection. Our results indicate that CCR4- cells do not expand during infection, which suggests that defective proviruses may arise during clonal proliferation. CCR4+ expression is commonly associated with cell trafficking to sites of inflammation [370], so these CCR4- cells infected with intact provirus may remain in circulation, where they can facilitate ongoing low levels of cell-to-cell transmission of provirus, maintaining infection. Furthermore, this provides additional

evidence that targeting CCR4 expressing cells using the mAb treatment, mogamulizumab, may not effectively eliminate the proviral reservoir. It is unknown yet whether this occurs exclusively in bronchiectasis donors or if this unique reservoir is found in all donors - something that will be further explored in future studies with additional patient samples.

4.5.5 Defective HTLV-1c proviruses may have arisen from homologous recombinational deletion

This novel study elucidated a defective provirus breakpoint motif in HTLV-1c infection. The use of provirus breakpoint sequences at the LTR or with LTR homology suggests the mechanisms involved in creating the internal proviral deletions are likely stemming from the repetitive nature of these elements. These deletions could be occurring through somatic DNA recombination events during proliferation of infected host cells. Human endogenous retroviruses (HERVs), that compromise between 3-8% of the human genome, are relics of exogenous viruses that were integrated into germ line cells [373], and often undergo recombinational deletion. Originally, HERVs have two LTRs identical in sequence which flank coding regions, much like HTLV-1. Over time, homologous recombination events between the HERV LTRs result in deletions, after which only a solo LTR may remain [374-376]. Studies show that the rate of recombinational deletion among HERVs is inversely proportional to the length of time that the HERV was incorporated into the human genome [377]. One study hypothesizes the increasing rate of mutational divergence of the LTRs the longer they have been integrated, decreases the probability of homologous recombination [377]. So for HTLV-1, which integrates itself into the human genome upon human-to human transmission events and not transmitted vertically via germline cells like HERVs, the rate of recombinational deletion could be high, especially given that infected cells can robustly proliferate in the host. This may explain why we see high levels of defective provirus in lung homing proxy phenotype cells, which significantly expand in HTLV-1c infection correlated to the PVL, hence would be exposed to many DNA recombination opportunities during robust proliferation. On the other hand, the frequency of CCR4-phenotype during HTLV-1c infection does not expand, and therefore less chances of DNA recombination events occurring during proliferation, supported by the fact that only full-length provirus was amplified in this reservoir. Indeed, one study of HTLV-1a

infection indicated that somatic mutations during proliferation rather than integration related mutations, account for most defective provirus [378].

4.5.6 Novel HTLV-1c-human genome chimeric proviruses may impact gene expression and pathogenesis

The discovery of HTLV-1c-hg chimeric provirus sequences may have profound implications for understanding host immune function, viral pathogenesis and risk of infection complications or disease manifestations.

Many of the chimeric proviruses contained hg centromeric sequences. Centromeres are the chromosome regions where the kinetochores assemble for mediating microtubule attachment to DNA for chromosome segregation [379]. While there are no specific DNA sequences that define a centromere, they are typically comprised of repetitive DNA sequences [380, 381]. α satellites are tandem nucleotide repeats in these centromeric and pericentromeric regions, which form a block of satellites called higher order repeats (HORs) [382]. When replication occurs, recombination can take place in the tandem repeats of centromeric regions, due to failed replication mechanisms such as fork stalling and collapse, or through DNA breaks [383]. This suggests that many of the HTLV-1c proviruses were integrated into centromeric regions, followed by a possible recombination event, resulting in an internal α satellite sequence flanked by LTRs to form a chimeric provirus. A recombination event like this may be facilitated by the repetitive nature of both the α satellites and the HTLV-1c LTRs. Furthermore, these chimeric proviruses in the centromeric regions may have been selected for over time, such as occurs in HTLV-1a infection with preference for provirus selection in proximity to centromeres [384]. In addition, the presence of an upstream LTR promoter may induce α satellite mRNA transcripts to be produced, which are conventionally modulated by the stage of cell-cycle and centromere-nucleolar interactions [385].

A number of the chimeric proviruses contained an internal short interspersed element (SINE) sequence. SINEs are short, non-coding, repetitive, retrotransposable sequences between 80-400bp in length [386]. They can invade new genomic sites using RNA intermediates, but as they are non-autonomous, they often utilise the replication machinery encoded for in neighbouring long interspersed elements (LINEs) to achieve this mobility [387]. *Alu* elements are the most abundant type of SINE in the human

genome, with a diverged dimer structure [388]. Although SINEs are transcriptionally active in embryonic stem cells, they are generally transcriptionally repressed through epigenetic modifications [386]. However SINE RNA expression can be triggered during both chemical and biological stress, such as DNA viral infection [389]. Recombination events that lead to internal SINE sequences in the defective HTLV-1c genome may facilitate a hijacking of the HTLV-1c 5'LTR promoter for transcription, and compromise the host cellular control of SINE expression, resulting in increased retrotransposition levels and induction of host genomic instability. Indeed, in vitro HIV-1 infection upregulates active retrotransposition of LINEs and *Alu* elements in a viral protein dependent manner [390].

Chimeric proviruses with internal human coding genes may have profound impacts on the host immune system. The HTLV-1 LTR can impact the expression of nearby host genes, or distal host genes brought into close proximity due to higher order structure with CTCF binding [391]. Therefore, host gene expression could be upregulated in a bidirectional manner if flanked directly by HTLV-1c LTRs. On the other hand, if the chimeric internal hg DNA sequence is only a partial segment of a coding gene, then the HTLV-1c component of the chimera may disrupt successful transcription of the host gene. In addition, novel hg:HTLV-1c mRNAs and proteins may be produced by the chimeric proviruses, instigating cell dysfunction. Indeed, novel chimeric RNAs and proteins have been detected in HIV-1 [392] and HTLV-1a [393, 394] infection. Notably, we detected seven chimeric proviruses with an internal sequence for replication dependent histone H2BC12-201, isolated from a single donor. This raises the possibility that these chimeric proviruses arose from clonal expansion, perhaps being selected for. Histones are a structural component of a nucleosome, which folds into a tertiary chromatin structure [395]. Chromatin structure is a key factor in host regulation of gene expression, replication, genomic stability and DNA repair [395]. Alteration to histone gene expression by HTLV-1c chimeric provirus could exert potentially damaging effects on the host cell, disrupting the cell-cycle, dysregulating host transcription and inducing DNA damage [396]. Indeed, upregulation of histone expression by nature supports more cell proliferation, and often occurs in cancers [397, 398]. Moreover, histones typically experience high levels of post-translational modifications such as phosphorylation, acetylation and methylation, and disruption of this can lead to nucleosome instability and tumours [399].

One such possibility explaining the presence of these insertions of segments of human DNA into defective proviruses is due to introduced experimental errors. Mispriming events in the SPA limiting dilution nested PCRs could result in sequencing errors being introduced. Furthermore, ONT sequencing could potentially introduce sequencing errors of short repeats. However, there are a number of reasons that indicate these novel chimeric proviruses are indeed genuine. Firstly, the host DNA insertions in the defective proviruses have over 98% sequence identity when mapped to the human genome, including all centromeric repeats, intergenic non-coding RNA regions, and coding exons. Introduced nucleotide errors would likely have less homology because they are of course by nature, error-prone. Secondly, some donors had the same defective chimeric provirus sequence, in different cell phenotypes and bulk PBMCs, indicating these likely arose from clonal expansion. In addition, the location of the hg DNA sequences are scattered throughout the entire genome, across all somatic chromosomes. Taken together, these factors indicate that the presence of segments of host DNA inside defective proviruses are not the result of PCR or Nanopore sequencing error, but legitimate biological processes. However, these results are being further validated with additional experiments. We are designing and implementing PCRs to amplify regions at the host:virus interface of chimeric proviruses from donor gDNA.

It will be pertinent to understand the nature of these novel chimeric proviruses better, given they comprise over one-fifth of the proviral landscape in HTLV-1c infection. More studies are needed with a larger cohort of donors to understand the potential effects they may have on host immune responses, viral pathogenesis, and disease association. Additional future experiments can investigate potential chimeric mRNA and proteins that may be produced from these unique proviruses through in vitro transfections with expression plasmids.

4.5.7 Chapter four conclusion

This novel study characterising the phenotype of HTLV-1c infected cells determined CD4+ T cells that have respiratory tract trafficking potential may play an important role in the pathogenesis of HTLV-1c infection and development of associated Bex. These cells are highly expanded, activated, and infected with provirus. Furthermore, the extensive chronic activation of all CD4+ T cells in all HTLV-1c infected donors likely

results in overall T cell dysfunction and diminishes an effective host immune response, contributing to lower survival rate of all HTLV-1c-infected individuals.

The characterisation of the HTLV-1c proviral landscape uncovered several unique and significant findings. Firstly, the proviral landscape in all donors has extensive internal deletions, likely caused by homologous recombinational deletion, supported by the breakpoints either occurring at LTRs or with LTR-like motifs. Secondly, the *pX* region, in particular the *hbz* gene, is preferentially retained in Bex donors, so *hbz* expression is likely driving the development of pulmonary disease. And finally, novel characterisation of HTLV-1c-hg chimeric proviruses may contribute to host immune dysregulation and pathogenesis. This chapter provides direction for future investigations into potential targets and efficacy of therapeutic treatments and vaccines, to ultimately eliminate the reservoir of integrated provirus.

5. Assessing the ability of a novel HTLV-1c humanised mouse model to recapitulate human in vivo infection

5.1 Chapter five background

5.1.1 Mouse models of infection have been widely used in HTLV-1 research

Many different animal model platforms have been developed and utilised over the years for HTLV-1 research, including rodents such as mice, hamsters, and rats; other small mammals like rabbits and ferrets; and NHPs such as macaques and sooty mangabeys [400].

This thesis chapter will focus on mouse models, as they have been used extensively in HTLV-1 research to model disease development, further understanding of the mechanisms of viral transmission and pathogenesis, and have served as pre-clinical platforms for testing therapeutic drugs against HTLV-1 [400]. In addition, mouse models are particularly useful as they are easy to maintain, the costs to sustain models are relatively low, and the number of mice that can be used in a single study is greater than larger animal studies.

5.1.2 Transgenic mouse models elucidated functions of HTLV-1 Tax and HBZ proteins in vivo

Transgenic (Tg) mouse models have facilitated greater understanding of the molecular and cellular functions of Tax and HBZ in HTLV-1 pathogenesis. Tg mice are genetically modified to allow the characterisation of molecular and cellular functions of genes of interest. The most common method for generating Tg mice involves the microinjection of a Tg construct into a fertilized mouse egg [401]. Alternative methods include transfecting a Tg construct into mouse embryonic stem cells, which are subsequently injected into mouse blastocysts, or by delivering the Tg construct into an egg via retrovirus vector [401]. A transgenic construct is comprised of a gene of interest under the control of a selected promoter, and may also include a reporter gene, such as green fluorescent protein, depending on the focus of the study [401]. Various promoters have been used in Tg mice to drive expression of the key HTLV-1 regulatory genes, including the HTLV-1 LTR [402], human granzyme B promoter [403], CD3-epsilon enhancer-promoter [404], and lymphocyte-specific protein tyrosine kinase promoter [405].

Tax-Tg mouse models demonstrated the expression of Tax across a wide range of tissues results in the formation of mesenchymal tumors, adenocarcinomas, and chronic inflammatory conditions [402, 406-408], supporting the notion that Tax contributes to the development of both ATL and HAM. Similarly, HBZ-Tg mice developed widespread inflammation and T cell lymphoma [409]. Furthermore, HBZ expression in Tg mice was shown to suppress IFN-γ production from CD4+ T cells and induce higher susceptibility to co-infections [189]. In addition, Tg mouse models have provided a useful platform for testing the efficacy of HTLV-1 drugs targeting Tax, HBZ, and their associated pathways [410-413].

5.1.3 Immunocompetent mice modelled persistent HTLV-1 infection

Early mouse models investigating HTLV-1 used immunocompetent mice, such as BALB/c strains, to model aspects of in vivo human infection. MT2 cells were injected into neonatal mice, and persistent HTLV-1 infection was established in a wide range of cells. Moreover, polyclonal proliferation was demonstrated in specific reservoirs, namely spleen and lymphatic tissues [414-417]. While easy to maintain, these models had several limiting factors: the immunocompetent mice generated little Ab responses; in vivo transmission of virus was not detected; and no development of HTLV-1-associated disease occurred throughout the course of infection.

To advance studies in immunocompetent mice, chimeric murine/HTLV-1 viruses were developed, which involved modifying the Env glycoprotein sequence to allow efficient infection of murine cells [418]. As such, persistent HTLV-1 infection was established in major tissues and organs of these mice. Further improvements in these chimeric models included detection of oligoclonal proliferation of infected cells, and greater host immune responses [419]. Moreover, chimeric virus infection in DC depleted mouse models demonstrated the importance of DCs in controlling early HTLV-1a infection by cell-free virions [420, 421].

5.1.4 Immunodeficient mouse models progressed understanding of ATL development and therapeutic interventions

In contrast to immunocompetent mice in which persistent infection establishes but does not progress, immunodeficient mice have various defects in their immune system which advances infection. Many strains of immunodeficient mice now exist and have been

extensively reviewed [422], however those most frequently used in HTLV-1 research are summarised in **Table 5.1**. Immunodeficient mice were first employed in xenograft models, facilitating the establishment of HTLV-1 infection more efficiently than the immunocompetent predecessors, and the development of associated disease.

Severe combined immunodeficiency (SCID) mice, utilised to model HTLV-1a-induced tumorigenesis by xenograft, have a mutation in the *Prkdc* gene, coding for an enzyme required in VDJ recombination of the TCR and B cell receptor (BCR) [423]. Early studies infected SCID mice with transformed HTLV-1a cell lines [424, 425], which demonstrated the proliferation and oncogenic potential of these infected cells. However, successful induction of tumorigenesis in SCID mice required the use of human peripheral blood lymphocytes (PBL) from ATL patients as the inoculum, rather than in vitro transformed cell lines [426, 427]. Moreover, the absence of detectable viral mRNA transcripts in SCID mice inoculated with ATL patient cells, as opposed to inoculation with HTLV-1a cell lines [428], suggested persistent infection and immune evasion are crucial to establishing oncogenesis, and the presence of non-transformed ATL cells also contribute to driving ATL in vivo. SCID mice were also used to test therapeutic treatments against ATL cells [429].

Additional combined immunodeficient mice strains were developed over time to improve the applications of using these models. Firstly, the SCID mutation was introduced into non-obese diabetic (NOD) mice lacking innate immunity [430], generating NOD/SCID mice [431]. These NOD/SCID mice served as useful xenograft platforms with which to test a wide range of therapeutic treatments for ATL [432-434] and for preventing de novo HTLV-1 infection [435].

Importantly, further mutations in the IL-2 receptor (IL2R) γ chain [436] were introduced into combined immunodeficient mice to generate three new strains, which are now widely used in HTLV-1 research: NSG (NOD-*scid-IL2Rγ*$^{-/-}$) [437, 438], NOG (NODShi-*scid-IL2Rγ*$^{-/-}$) [439] and BRG (BALB/c.Cg-*Rag2*$^{-/-}$*-IL2Rγ*$^{-/-}$) [440, 441] strains.

NOG mice were used to develop a xenograft model of ATL tumorigenesis, by inoculating the immunodeficient mice with HTLV-1a infected cell lines [442]. Further advancements of the NOG xenograft models involved engraftment with HTLV-1a-infected primary human PBMCs from ACs, demonstrating HTLV-1-induced cell proliferation and de novo infection [443, 444]. In addition, the role of *hbz* RNA and

HBZ protein in the proliferation of infected cells and tissue/organ infiltration was demonstrated in vivo using a NOG mouse model [445].

5.1.5 Humanised mouse models have greatly advanced HTLV-1 research

The development of humanised (hu) mice has revolutionised medical and scientific research, including a wide variety of fields such as infectious diseases, cancers, drug treatments, vaccines, autoimmunity, and indeed HTLV-1 research [446]. They provide a more analogous platform to human infection and can be a useful pre-clinical model to test therapeutics and prevention strategies. In brief, hu-mouse models are generated by engrafting human cells into mice which are already severely immunodeficient (**Table 5.1**), including many previously defined in the HTLV-1 xenograft models, and allowing reconstitution of a human immune system (HIS) [447]. Hu-mouse models can differ depending on the strain of immunodeficient mice, the engraftment technique, and the inoculum of HTLV-1 infected cells.

Regardless of the methods employed, hu-mice have proven to be an invaluable tool to study HTLV-1 infection [400, 448]. SCID mice engrafted with human fetal thymic and liver tissue (hu-SCID), and subsequently inoculated with HTLV-1a-infected human CD34+ cells, suggested the potential for HSCs as a reservoir for HTLV-1 in vivo [449]. The essential role of HTLV-1 accessory protein p12 was confirmed by infecting hu-NSG mice with primary human CD4+ T cells transfected with an HTLV-1 molecular clone [257]. Hu-BRG mice infected with HTLV-1a, after engraftment of an HIS with CD34+ cells, demonstrated disease outcomes such as hepato-splenomegaly and lymphadenopathy to provide a model of ATL development, and suggested the thymus may be important in the development of ATL in vivo [450]. Similarly, ATL was well recapitulated in a hu-NOG mouse model that produced HTLV-1a-specific immune responses and developed lymphocytes with lobulated nuclei over the course of infection, akin to the flower-like cells produced during ATL in human infection [451]. Hu-NSG mice have also been utilised in establishing HTLV-1 co-infection models: in particular, *S. stercoralis* co-infection leading to the frequently fatal hyper-infection [452].

Ultimately, HTLV-1 mouse models, especially humanised models, have provided extensive knowledge surrounding HTLV-1 infection over the years, and improvements to these models are continually being made to better recapitulate human infection. They will continue to be vital in pre-clinical studies as we move towards finding efficient

therapeutic and preventative treatments against HTLV-1 which are so desperately needed.

Table 5.1 – Immunocompromised mouse strains commonly used in HTLV-1 xenograft and humanised mouse models

Strain	Mutations	Features
NOD [430]	Innate immunity (*Il2*; *Hc*, *Cdh23*; *mt-Tr*; *H2*; *Gpr84*; *Del(3)1Lt*; *Del(1)2Lt*; *Cel(3)3Lt*; *Cox7a2l*)	Deficient NK cell and macrophage functions; Reduced complement activity
SCID [423]	Innate immunity; *Prkdc*	Impaired T and B cell development
NOD-SCID [431]	Innate immunity; *Prkdc*	Deficient NK cell and macrophage functions; Impaired T and B cell development; Reduced complement activity
NSG [437, 438]	Innate immunity; *Prkdc; IL-2Rγ* (complete null)	Deficient NK cell, DC and macrophage functions; Impaired T and B cell development; Reduced complement activity; IL2Rγ is not expressed
NOG [439]	Innate immunity; *Prkdc; IL-2Rγ* (truncated signalling domain)	Deficient NK cell, DC and macrophage functions; Impaired T and B cell development; Reduced complement activity; IL2Rγ is expressed and binds cytokines, but cannot signal
BRG [440, 441]	*Rag2; IL2Rγ* (truncated signalling domain)	Deficient T, B and NK cell functions; IL2Rγ is expressed and binds cytokines, but cannot signal

5.2 Chapter five aims

Given the genetic differences of HTLV-1 subtype-A and -C, and all mouse models to date have replicated HTLV-1a infection, we developed, with our collaborators Professor Marc Pellegrini and Dr James Cooney, a humanised mouse model for HTLV-1c infection. This model successfully established HTLV-1c infection in hu-NSG mice, demonstrating rapid infection and lymphocytosis in all animals (Cooney et al., 2023, Cell, in revision). Furthermore, this model has produced promising results in testing combinations of preventative/therapeutic drugs that are already currently available, against HTLV-1 (Cooney et al., 2023, Cell, in revision), and may be a useful pre-clinical

platform to test potential drugs going forwards. Therefore, we sought to assess the molecular and immunological aspects of HTLV-1c infection in this novel model, and its ability to recapitulate human in vivo infection. We aimed to achieve this by mirroring the immunophenotyping and molecular characterisation of the proviral reservoir applied in chapter four of this thesis for human infection, to the HTLV-1c-infected hu-mouse model.

5.3 Characterising the phenotype of HTLV-1c infected cells in humanised mice

5.3.1 Generating an HTLV-1c infected hu-NSG mouse model

A novel hu-mouse model infected with HTLV-1c was developed in collaboration with Professor Marc Pellegrini and Dr James Cooney (WEHI, Melbourne, Australia) (Cooney et al., 2023, Cell, in revision). Immunodeficient NSG mice were chosen for this study to allow for efficient establishment of an HIS and subsequent infection. These mice have defective innate immunity, stemming from the NOD background strain [430]. They also have defective double stranded DNA repair from the SCID background strain [423], such that TCR and BCR recombination is prevented, meaning they are unable to develop antigen-specific T and B lymphocytes [436, 453, 454]. Additionally, these mice have a mutation in IL2Rγ such that it is not expressed nor binds cytokines [437].

To establish the humanised mouse model, mice pups were first sub-lethally irradiated, and then received an intra-facial injection of human CD34+ cord blood stem cells, to allow the engraftment of an HIS. The amount of matched donor CD34+ cells determines the size of an experimental group, to approximately 12. Reconstitution of the human immune system was established from 16 weeks, thereafter mice received an intraperitoneal injection of lethally irradiated primary human PBMCs infected with HTLV-1c. Subsequent infections used ex vivo HTLV-1c-infected hu-NSG splenocytes as the inoculum. Establishment of infection in the hu-NSG mice was confirmed after two weeks in peripheral blood by ddPCR with *gag* primer/probe of extracted gDNA (**Figure 5.1**).

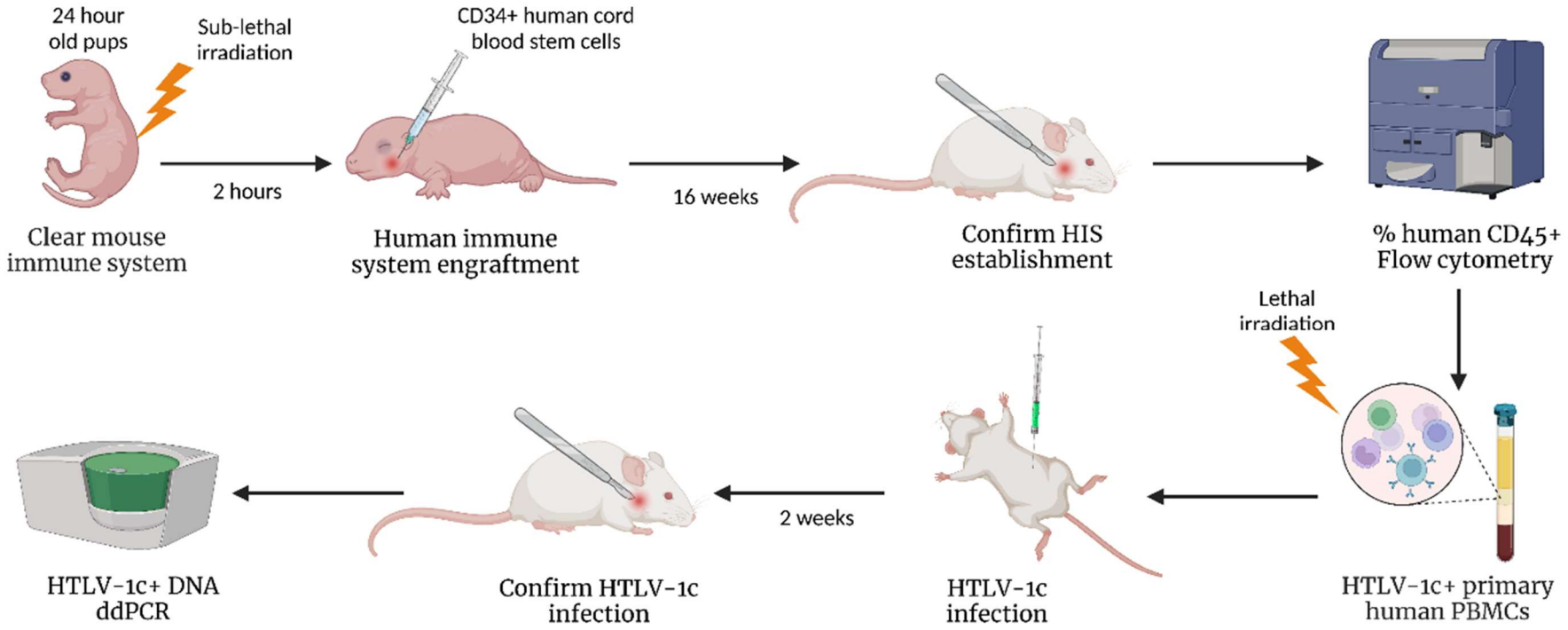


Figure 5.1 – Generation of novel HTLV-1c humanised mouse model. 24-hour old NSG pups are sub-lethally irradiated to clear mouse immune cells. After two hours, the pups receive an intra-facial injection of human donor CD34+ cord blood stem cells, to begin engraftment of a HIS. After 16 weeks the mice are assessed for establishment of a HIS, by determining the frequency of human CD45+ cells by flow cytometry in peripheral blood obtained from a submandibular bleed. Primary human PBMCs infected with HTLV-1c are lethally irradiated and injected intraperitoneally into the hu-NSG mice. After two weeks, an additional submandibular bleed is performed to confirm HTLV-1c infection, where gDNA is isolated from peripheral blood and analysed by ddPCR. Figure was created with Biorender.com.

5.3.2 Immunophenotype of hu-NSG mice infected with HTLV-1c mirrors human in vivo infection

To determine the immunophenotype of hu-NSG mice infected with HTLV-1c, we closely mirrored the methodology used for human in vivo immunophenotyping in chapter four of this thesis. We harvested splenocytes of 8 mice at an early infection timepoint, 2 weeks post infection (wpi); 3 mice at a late infection timepoint, 6 wpi; and compared to 3 uninfected control mice. We incubated the splenocyte suspension samples with the same 10-colour panel of anti-human antibodies detailed in **Table 2.3**, and analysed by flow cytometry (**Figure 4.1**). In addition, we determined the splenocyte PVL of each mouse through the *tax* primer/probe ddPCR assay [218] for all 8 mice at early infection, and 2 out of 3 mice at late infection.

Firstly, we analysed the frequency and activation levels of naïve (CD45RA+) and effector/memory (CD45RA-) CD4+ T cells present in the splenocytes of the three experimental groups. The frequency of naïve CD4+ T cells was not significantly different between uninfected (mean, 16.04%) and HTLV-1c-infected mice from the early timepoint (mean, 32.69%; p, 0.34) (**Figure 5.2A**). Although HTLV-1c infected late timepoint mice exhibited a substantially lower frequency of naïve CD4+ T cells (mean, 0.80%) than the early infection group or controls, this was not significantly different (p, 0.36, 0.09, respectively) (**Figure 5.2A**). While the proportion of E/M CD4+ T cells was much higher in late infection mice (mean, 89.30%) than negative controls (mean, 78.03%), there was no significance (p, 0.51), and similarly, there was no significant difference between early infection mice (mean, 53.39%) and controls (p, 0.13) (**Figure 5.2A**). Between the HTLV-1c infected groups, late infection mice displayed a higher frequency of effector/memory CD4+ T cells in splenocytes when compared to early infection (p, 0.06) (**Figure 5.2A**). Interestingly, the PVL of HTLV-1c infected mice showed a negative correlation trend with the frequency of naïve CD4+ T cells (r, -0.59; p, 0.07), and significantly positive correlation with the frequency of effector/memory CD4+ T cells (r, 0.63; p, 0.05) (**Figure 5.2B**).

Strikingly, the proportion of activated (HLA-DR+CD38+) E/M CD4+ T cell splenocytes was significantly higher in HTLV-1c infected mice at 2 wpi (mean, 21.44%; p, 0.007) and 6 wpi (mean, 36.77%; p, 0.0002), when compared to uninfected control mice (mean, 9.06%) (**Figure 5.3A**). Furthermore, mice at 6 wpi displayed significantly

higher frequencies of activation of E/M CD4+ T cells than mice in the 2 wpi group (p, 0.003) (**Figure 5.3A**). Notably, the PVL of infected mice at both time points is significantly correlated to the proportion of E/M activation (r, 0.83; p, 0.003) (**Figure 5.3B**).

We then examined the key CCR4 expressing immunophenotypes of interest in hu-NSG mice splenocytes: FOXP3+ proxy (CCR4+CD49d-CD127-), lung homing proxy (CCR4+CD49d+Integrin-β7-), and CCR4- E/M CD4+ T cells. Comparable to what was found in human HTLV-1c infection, there was no difference in the frequency of FOXP3+ proxy phenotype between uninfected controls (mean, 30.27%) and early (mean, 27.66%; p, 0.79) or late (mean, 20.34%; p, 0.60) infection mice groups, nor between infection timepoints themselves (p, 0.60) (**Figure 5.2A**). Interestingly, the frequency of FOXP3+ proxy phenotype was significantly negatively correlated to the PVL for HTLV-1c-infected mice, inclusive of both timepoints (r, -0.64; p, 0.05) (**Figure 5.2B**). Meanwhile, the proportion of activated FOXP3+ proxy phenotype is significantly higher in 2 wpi (mean, 18.61%; p, 0.04) and 6 wpi (mean, 32.00%; p, 0.004) hu-mice splenocytes when compared to uninfected controls (median, 7.31%) (**Figure 5.3A**). Furthermore, the proportion of activated FOXP3+ proxy phenotype was significantly higher when comparing late to early infection (p, 0.03) (**Figure 5.3A**). Remarkably, the proportion of activated FOXP3+ proxy phenotype is significantly positively correlated to PVL for all HTLV-1c-infected hu-NSG mice (r, 0.75; p, 0.01) (**Figure 5.3B**).

The lung homing proxy phenotype of E/M CD4+ T cells was significantly expanded in late timepoint HTLV-1c-infected mice (mean, 32.73%; p, 0.02), and trending higher in early infection mice (mean, 18.76%; p, 0.08), when compared to uninfected controls (mean, 6.97%). In addition, 6 wpi mice tended to display higher levels of this phenotype in splenocytes than 2 wpi (p, 0.07) (**Figure 5.2A**). However, the PVL of HTLV-1c infected mice did not correlate to the frequency of lung homing proxy phenotype (r, 0.34; p, 0.33) (**Figure 5.2B**). Consistently, the lung homing proxy phenotype in hu-mice splenocytes displayed greater expression of activation surface markers at both early (mean, 16.83%; p, 0.02) and late (mean, 30.73%; p, 0.001) HTLV-1c infection timepoints when compared to negative controls (median, 5.41%) (**Figure 5.3A**). In addition, the proportion of activated lung homing proxy phenotype cells significantly increases from early to late infection time points (p, 0.01) (**Figure 5.3A**). Notably, the frequency of

activated lung homing proxy phenotype cells is significantly positively correlated to PVL (r, 0.90; p, 0.0004) for all HTLV-1c-infected hu-NSG mice (**Figure 5.3B**).

The frequency of CCR4- phenotype of effector/memory CD4+ T cells in hu-NSG splenocytes was not significantly different when comparing early (mean, 21.68%) and late infection (mean, 26.28%) to negative controls (mean, 32.70%; p, 0.69, 0.69, respectively), nor when comparing infection timepoints (p, 0.69) (**Figure 5.2A**). Meanwhile, the CCR4- frequency in HTLV-1c infected splenocytes was weakly positively correlated to PVL (r, 0.60; p, 0.07) (**Figure 5.2B**). However, consistent with all other CD4+ T cell phenotypes investigated in human and HIS mice, these CCR4- cells displayed substantially higher levels of activation at early (mean, 31.04%; p, 0.0004) and late (mean, 40.47%; p, 0.0004) HTLV-1c infection timepoints when compared to negative controls (mean, 2.19%) (**Figure 5.3A**). Moreover, late infection mice tended to display a higher proportion of CCR4- activation phenotype than early infection (p, 0.06) (**Figure 5.3A**). Finally, the CCR4- activated phenotype frequency of all HTLV-1c infected HIS mice showed a positive correlation trend to PVL (r, 0.60; p, 0.07) (**Figure 5.3B**).

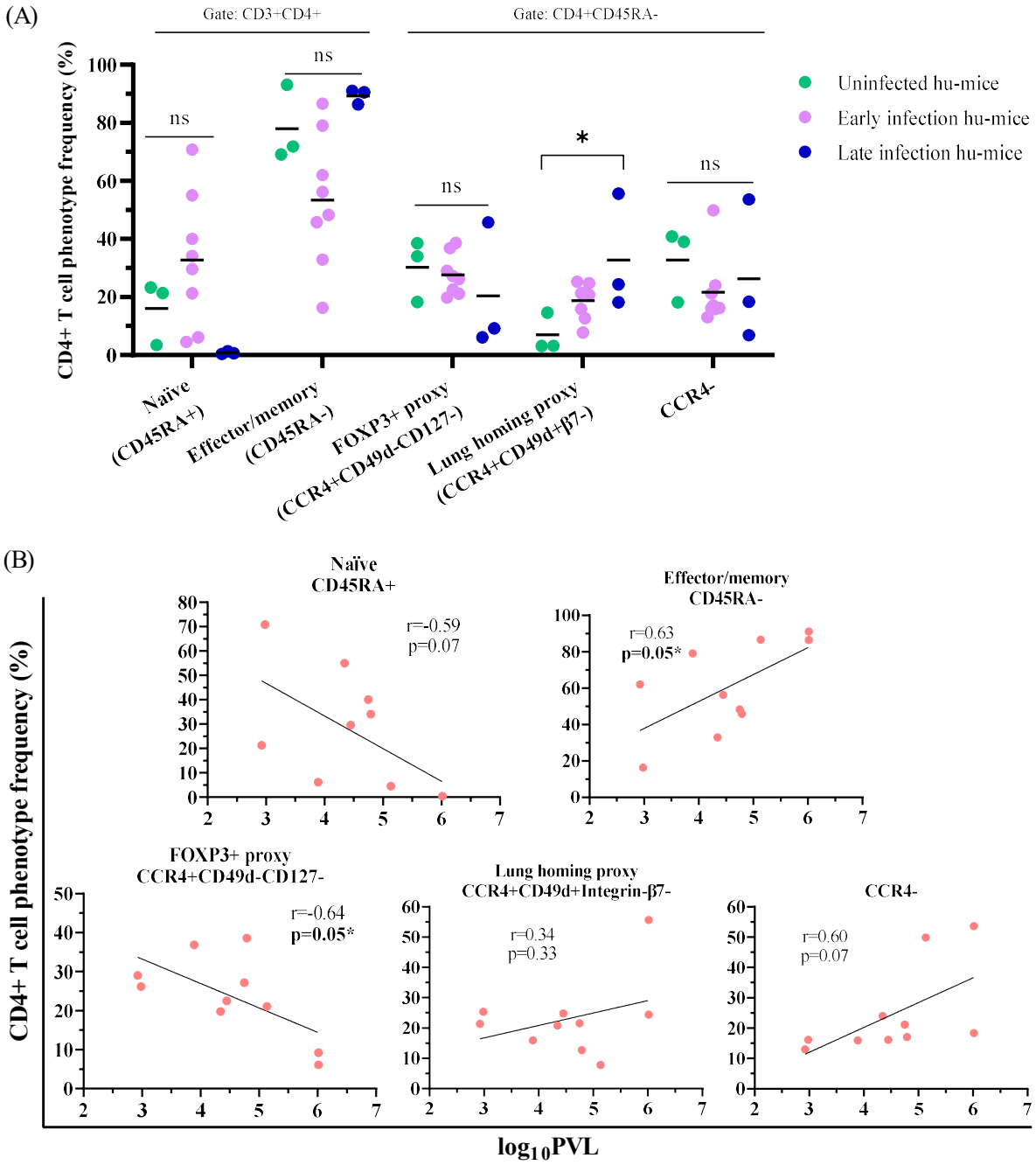


Figure 5.2 – Lung homing proxy phenotype of effector/memory CD4+ T cells expand in humanised mice infected with HTLV-1c. Hu-mice splenocytes were harvested and incubated with a 10-colour cocktail of anti-human mAbs and processed via flow cytometry. gDNA was extracted from bulk splenocytes and PVL was determined with ddPCR. (A) Frequency of CD4+ T cell phenotypes. n, 14 (3 uninfected, 8 early infection (2 wpi); 3 late infection (6 wpi)). (B) Correlation of CD4+ T cell phenotype frequency in HTLV-1c+ hu-mice with PVL. n, 10 (8; 2pwi, 2; 6 wpi). Mean frequency of each test group in (A) is shown by black line. Statistical significance between groups in (A) was assessed using one way ANOVA with two stage linear step-up procedure of Benjamini, Krieger and Yekutieli as an FDR control for multiple comparisons, with adjusted p value (q value) shown. Correlation significance in (B) was assessed using Pearson correlation test. ns, not significant, $p>0.05$; *$p<0.05$.

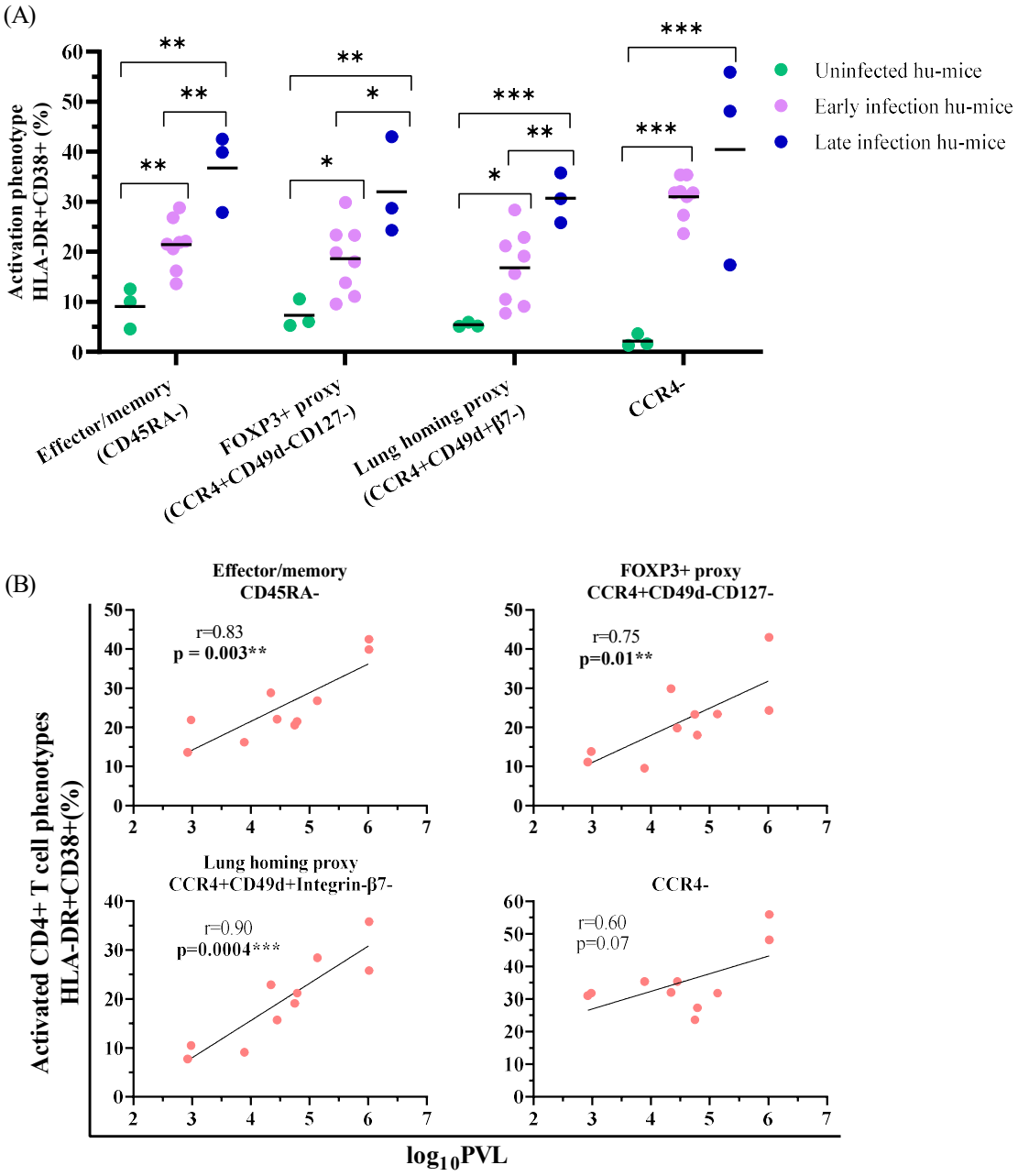


Figure 5.3 – All effector/memory CD4+ T cells are highly activated in humanised mice infected with HTLV-1c and correlated to proviral load. Hu-mice splenocytes were harvested and incubated with a 10-colour cocktail of anti-human mAbs and processed via flow cytometry. gDNA was extracted from bulk splenocytes and PVL was determined with ddPCR. (A) Frequency of activation marker (HLA-DR+CD38+) expression in CD4+ T cell phenotypes. n, 14 (3 uninfected, 8 early infection (2 wpi), 3 late infection (6 wpi). (B) Correlation of activation marker (HLA-DR+CD38+) expression frequency of CD4+ T cell phenotypes and PVL. n, 10 (8, 2wpi; 2, 6wpi). Mean frequency of each test group in (A) is shown by black line. Statistical significance between groups in (A) was assessed using one way ANOVA with two stage linear step-up procedure of Benjamini, Krieger and Yekutieli as an FDR control for multiple comparisons, with adjusted p value (q value) shown. Correlation significance in (B) was assessed using Pearson correlation test. ns, not significant, $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$.

5.3.3 HTLV-1c provirus is enriched in CCR4+ and CCR4- phenotypes of CD4+ T cells in hu-mice splenocytes

With significant activation and immunophenotype expansion in hu-NSG mice analogous to human infection from chapter four, we next investigated whether these splenocyte reservoirs were indeed enriched with HTLV-1c provirus. We sorted sufficient quantities of cells by FACS into each phenotype of five early (2 wpi) and two late (6 wpi) infection timepoint samples and extracted gDNA from these reservoirs. Coinciding with the human donor studies in chapter four of this thesis, we then measured the PVL in each cell reservoir by ddPCR using the *tax* primer/probe target assay [218] (**Figure 5.4A**). Similarly, we $\log_{10}(x)$ transformed these values for better visualisation of the widely distributed dataset.

Consistent with trends observed in primary human PBMCs, during early HTLV-1c infection of hu-mice, provirus was significantly enriched in the lung homing proxy phenotype (CCR4+CD49d+Integrin-B7-) (mean, 4.66; p, 0.01) and CCR4- phenotype (mean, 4.93; p, 0.003), and trending higher in FOXP3+ proxy phenotype (mean, 4.20; p, 0.21), when compared to the bulk splenocyte reservoir (mean, 3.89) (**Figure 5.4B**). However, at the late infection timepoint of hu-NSG mice, given the PVL in all cell reservoirs is greater than 100%, there was no discernible difference between each phenotype as they all surpassed proviral saturation (**Figure 5.4C**), which possibly arises in a setting lacking adaptive immune control of virus.

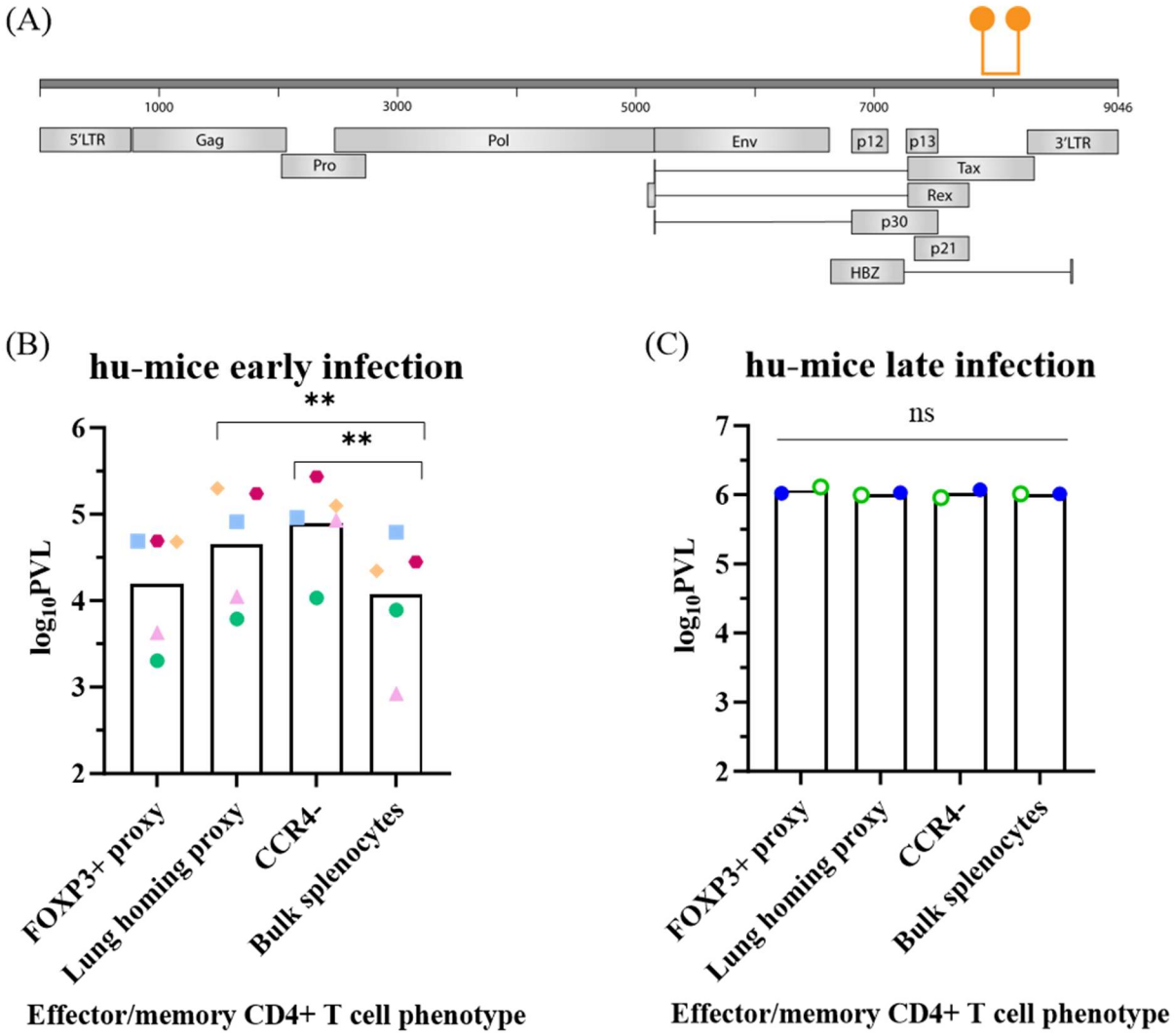


Figure 5.4 – Enrichment of HTLV-1c provirus in CCR4± phenotype CD4+ T cell reservoirs in humanised mouse splenocytes. Hu-mice splenocytes were harvested and incubated with a 10-colour cocktail of anti-human mAbs and processed via flow cytometry. gDNA was extracted from bulk splenocytes and PVL was determined with ddPCR. (A) HTLV-1c genome with ddPCR primers and probes in the *tax* gene region. gDNA was extracted from bulk splenocytes and sorted CD4+ T cells, and PVL was determined by ddPCR in triplicate. (B) Enrichment of HTLV-1c provirus in hu-NSG mice (n, 5) at early infection (2 wpi). (C) Enrichment of HTLV-1c provirus in hu-NSG mice (n, 2) at late infection (6 wpi). Each symbol represents a different mouse and is the mean of triplicate measurements. Bar graphs show the mean PVL for each reservoir. Statistical significance was assessed using RM one-way ANOVA with two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli as an FDR control for multiple comparisons, with adjusted p value (q value) shown. ns, not significant, $p>0.05$; * $p<0.05$; ** $p<0.01$.

5.4 Characterising the proviral landscape of hu-mice infected with HTLV-1c

5.4.1 Envelope deletions accumulate over time in HTLV-1c-infected hu-mice

To continue with the assessment of the HTLV-1c humanised mouse model against human infection, we determined the absolute frequency of four HTLV-1c gene coding regions (*gag*, *env*, *hbz* and *tax*) in CCR4 immunophenotypes harbouring provirus (FOXP3+ proxy, lung homing proxy and CCR4-) and bulk splenocytes (**Figure 5.5A**). We had sufficient template gDNA extracted from sorted splenocytes, for the ddPCR assay to characterise the normalised gene retention of two mice at 6 wpi. At the early infection timepoint we had sufficient template in the bulk splenocytes of 8 mice, however in the sorted phenotypes at 2wpi the HTLV-1c and *RPP30* probes were commonly below the level of detection, and therefore excluded from analysis. Gene retention for each region was determined by normalising the specific probe PVL to the highest detected probe of that sample.

During early HTLV-1c infection in the hu-NSG mouse model, the *tax* gene region was most highly retained (mean, 90.76%) in proviral sequences from bulk splenocytes (**Figure 5.5B**). *Env* and *hbz* gene regions maintained a similar retention in the proviral reservoir (mean, 73.96%, 64.64%, respectively), while *gag* was the most deleted coding region (mean, 55.50%) (**Figure 5.5B**).

At late infection time point, a consistent pattern emerged across all immunophenotype proviral reservoirs – the *env* gene region was least retained. In the FOXP3+ proxy phenotype, *gag*, *hbz* and *tax* regions were retained at similar levels in the proviral reservoir (mean, 100%, 93.10%, 84.55%, respectively), while *env* was detected at a markedly lower frequency (mean, 33.58%) (**Figure 5.5C**). The lung homing proxy phenotype reservoir contained proviruses which consistently retained *gag* and *hbz* coding regions (mean, 100%, 96.07%, respectively) (**Figure 5.5C**). Tax was moderately retained (mean, 71.62%), while the env coding region was only detected in about half of the proviral reservoir (mean, 54.46%) (**Figure 5.5C**). A similar trend was observed in the CCR4- immunophenotype: *gag* and *hbz* regions were detected in nearly all proviruses (mean, 94.18%, 99.92%) (**Figure 5.5C**). While the *tax* gene region was sometimes

deleted in the CCR4- proviral reservoir (mean, 78.93%), consistently the *env* region was least retained (mean, 47.18%) (**Figure 5.5C**). When considering the bulk splenocyte proviral reservoir at late infection, very high levels of retention were reported for *gag* and *hbz* (mean, 98.64%, 98.74%), while *env* and *tax* regions were detected at much lower frequencies (mean, 69.01%, 63.04%, respectively) (**Figure 5.5C**).

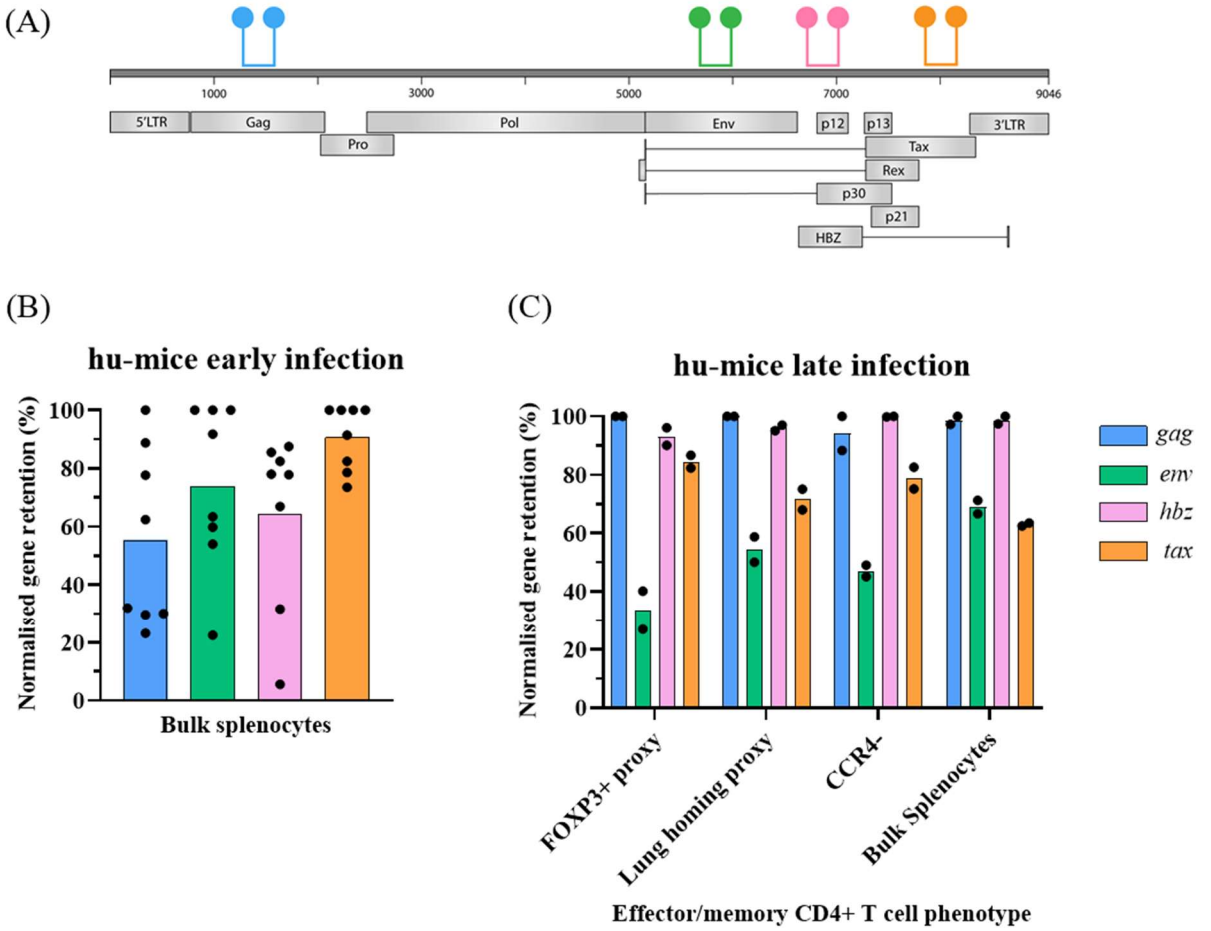


Figure 5.5 – Normalised gene retention of HTLV-1c provirus in humanised mouse splenocytes shows accumulation over time of proviruses with deleted *env*. Hu-mice splenocytes were harvested and incubated with a 10-colour cocktail of anti-human mAbs and processed via flow cytometry. gDNA was extracted from bulk splenocytes and PVL was determined with ddPCR. The PVL of each probe is measured in duplicate, and then normalised to the highest mean PVL of a given probe in that sample. (A) HTLV-1c genome with primers and probes for ddPCR targeting the gene regions of *gag*, *env*, *hbz* and *tax*. (B) Normalised HTLV-1c gene retention frequency in the CD4+ immunophenotype reservoirs of hu-mice at early infection (2 wpi) (n, 8). (C) Normalised HTLV-1c gene retention frequency in the CD4+ immunophenotype reservoirs of hu-mice at late HTLV-1c infection (6 wpi) (n, 2). The mean normalised gene retention is shown by bar graphs, with individual points representing different mice.

5.4.2 In vivo HTLV-1c proviral deletions are more extensive in humanised mice than in humans

To continue the evaluation of the HTLV-1c humanised mouse model, we next generated and analysed high resolution HTLV-1c proviral sequences from infection of hu-NSG mice. We again utilised the SPA assay, the limiting dilution touchdown nested PCR that amplifies individual proviral integrations from 5'- to 3'-LTRs (**Figure 4.8A**), using gDNA isolated from splenocytes harvested from two mice at the late infection timepoint (6 wpi). Proviral amplicons were subsequently purified and sequenced on the ONT platform (**Figure 4.8B**). In addition, we also determined the PVL of each mouse using ddPCR. Unfortunately, we did not recover sufficient cells and gDNA from any mice at 2 wpi to characterise the proviral landscape at this time point.

We mapped every individual proviral sequence from both mice to the HTLV-1c genomic consensus sequence to determine the overall landscape. Firstly, the large majority (85.7%) of proviral sequences in the humanised mouse model had large internal deletions (**Figure 5.6A**), similar to the landscape in human donors. Intriguingly, the remaining 14.3% of proviruses amplified from hu-NSG mice splenocytes were in fact HTLV-1 chimeras with internal hg sequences (**Figure 5.6A**), indeed mirroring human infection.

Next, we determined the normalised read count (out of an arbitrary maximum of 200) of every bp across the HTLV-1c genome for the hu-NSG mice splenocyte derived proviruses, and then compared to the similarly aligned human donor proviral sequences (**Figure 5.6B**). The most retained region of the HTLV-1c provirus in hu-NSG mice is the LTR regions, like the human studies. The 5'-LTR region had a normalised read count of approximately 200, while the 3'-LTR region has a normalised read count that declines from 200 at the far 3'-end, to approximately 100 at the LTR-*pX* border (**Figure 5.6B**). Similarly, the beginning of the *gag* gene region is moderately conserved, demonstrated by a normalised read count of approximately 120. However, the coverage continually drops through the *gag* gene region, so by the end, the normalised read count of the mouse proviral sequences is less than 20 (**Figure 5.6B).** Notably, the extended *pX* genomic region was rarely retained in the hu-mice as compared to the human samples: the normalised read count in the middle of the *tax* exon 2 region was approximately 90, however significantly drops off after this point towards the *hbz* exon 2 to less than 20

(**Figure 5.6B**). Interestingly, the normalised read count decreases more smoothly at the 5'-end in the hu-NSG defective proviral sequences, compared to the more stepwise manner in the human-derived samples, indicating there are less common proviral breakpoint sites across all proviral sequences in the mice splenocytes. Most strikingly, the hu-NSG mice have provirus with larger internal deletions, as demonstrated by the complete absence of internal sequences in the *pro*, *pol* and *env* coding regions, when compared to the provirus sequences amplified from human infection.

When separating out each proviral sequence by mouse, we detected a few common provirus breakpoints both within a single mouse, and also between mice (**Figure 5.7A**). These breakpoints were at 758 bp (Mouse 84 x6 and Mouse 86 x3) and 8289 bp (M084 x7 and M086 x3), the 5' and 3'LTRs, respectively. There were also 8 amplicons with a breakpoint at 1071 bp in *gag* and at 8817 bp within the 3'LTR in mouse 86. There were also 6 instances where the breakpoint occurred within the 3'LTR in mouse 84 (3x 8634 bp and 3x 8876 bp). Overall, the nucleotide retention across the genome was almost identical between mice (**Figure 5.7B**): M084 had 8.90% coverage, while M086 had 8.55% coverage (p, 0.78). (**Figure 5.7C**).

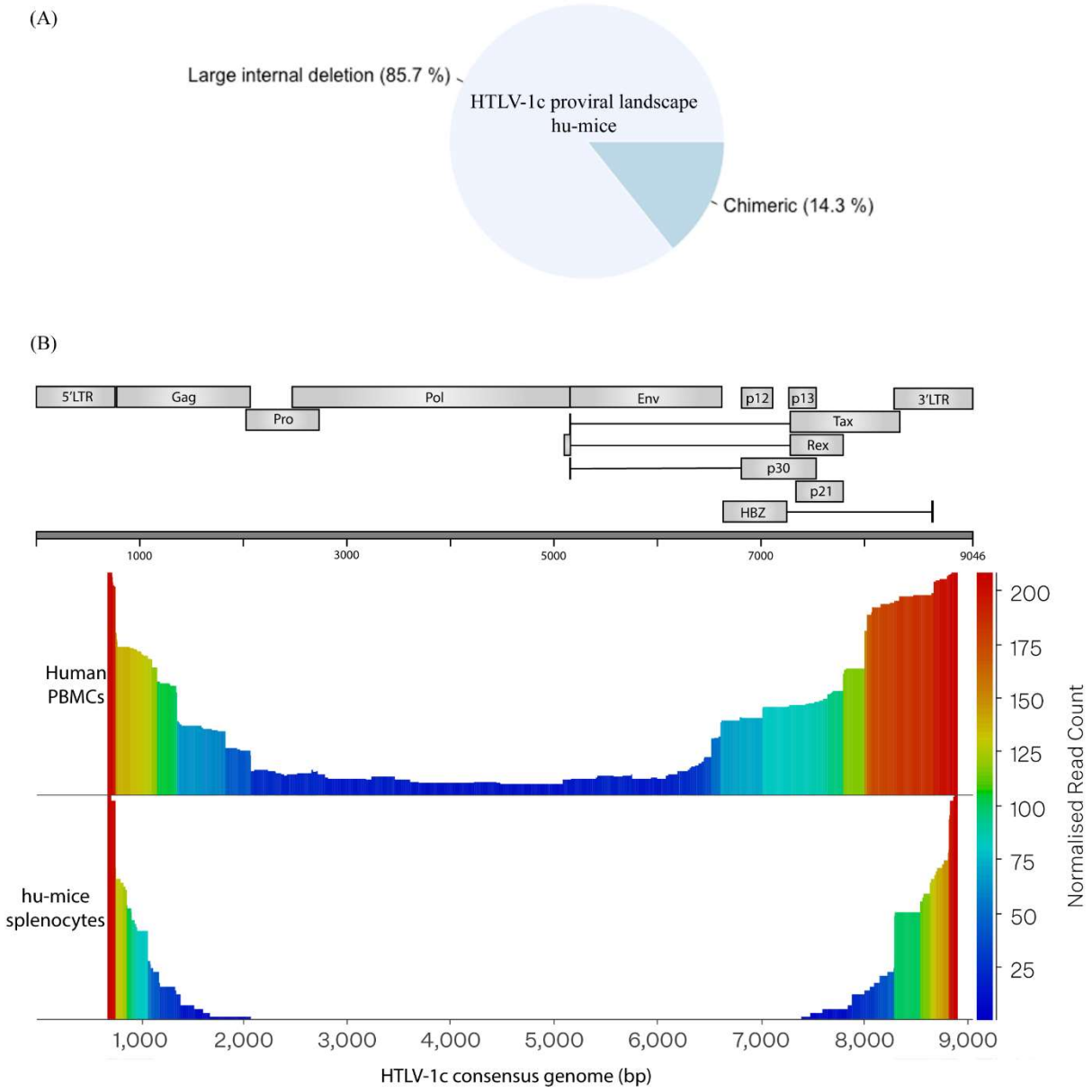


Figure 5.6 – The HTLV-1c proviral landscape in humanised mice splenocytes is highly defective when compared to primary human PBMCs. gDNA was extracted from hu-mice splenocytes infected with HTLV-1c at late time point (6 wpi) and SPA assay was performed by limiting dilution touchdown nested PCR. Individual proviral amplicons were sequenced with the ONT platform and aligned to the HTLV-1c consensus genome. (A) Overall frequencies of proviral sequence categories: Large internal deletions, and HTLV-1c-hg chimeras. (B) Comparison between primary human PBMCs and hu-mice splenocytes of the normalised read count of each nucleotide in the proviral landscape.

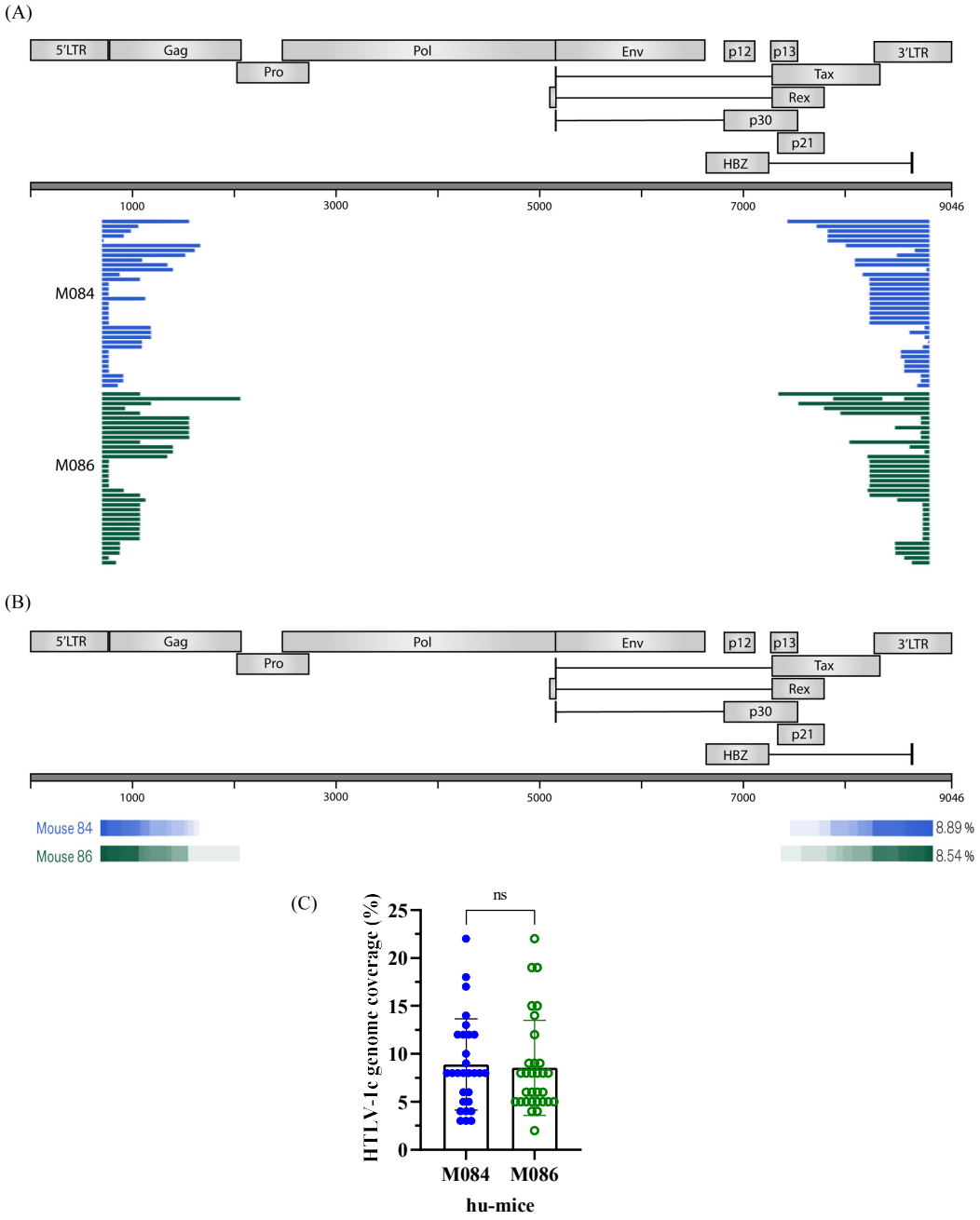


Figure 5.7 – HTLV-1c proviral sequences have large internal deletions in humanised mice splenocytes. gDNA was extracted from hu-mice splenocytes infected with HTLV-1c at late time point (6 wpi) and SPA assay was performed by limiting dilution touchdown nested PCR. Individual proviral amplicons were sequenced with the ONT platform and aligned to the HTLV-1c consensus genome. (A) Individual proviral sequences with large internal deletions aligned to the HTLV-1c consensus genome, separated by hu-mouse (M084 in blue, M086 in green). (B) Summary of the HTLV-1c proviral coverage across the genome for each nucleotide. Deeper colour represents higher retention in the provirus pool, lighter colour represents less retention. (C) Percentage of HTLV-1c genome coverage for each hu-mouse. Bars represent the mean coverage, and each data point represents a single provirus with internal deletion, and error bars ±SD. Statistical significance was assessed using unpaired T-test. ns, not significant, $p>0.05$.

5.4.3 HTLV-1c defective provirus breakpoints in hu-mice do not have common sequence motifs

Given we uncovered a striking breakpoint LTR-like motif in human derived defective proviruses in chapter four, we investigated the sequences at the breakpoints of splenocyte-derived proviruses with large internal deletions. In parallel with the human donor studies, we selected the final 30 bases of the 5' deletion junction, and the first 30 bases of the 3' deletion junction of defective proviruses, and filtered any provirus sequences that had breakpoints occurring exactly at the LTR junctions. Next, we created a sequence logo, which represents the signal strength of each nucleic acid in each base of the sequence, and compared to the HTLV-1c LTR sequences (**Figure 5.8A**). Contrasting to the observed defective provirus breakpoint motif in patient samples in chapter four of this thesis (**Figure 5.8B**), the HTLV-1c breakpoint sequences from hu-mice do not reflect an LTR-like pattern, and in fact do not display any particularly strong sequence motifs (**Figure 5.8C**).

5.4.4 Chimeric HTLV-1c-human genome proviruses detected in hu-mouse model

The unique discovery of chimeric HTLV-1c proviruses in human donors was emulated in the humanised mice splenocytes. The chimeric proviruses amplified from both hu-mice are detailed in **Table 5.2**. Interestingly, the HTLV-1c breakpoints of the chimeric proviruses in mice varied: some occurred at LTR junctions, while others occurred in the *gag* and *pX* regions (**Figure 5.9A**); contrasting to the chimeric proviruses in human donors which almost exclusively demonstrated LTR-LTR junction with internal hg sequence. The chimeric proviruses contained hg sequences ranging in size from 61 – 371 bp, mapping to three different chromosomes in the human genome (**Figure 5.9B**). Five chimeras contained the coding region of DNA damage inducible transcript 3 (DDIT3), three chimeras contained sequences of cytosolic glutamic acid-CTC tRNA (tRNA-Glu anticodon CTC), while one chimera contained intron and exon regions for 40S ribosomal protein subunit 29 (RPS29) and RNA component of signal recognition particle 70SL1 (RN7SL1), respectively (**Table 5.2**).

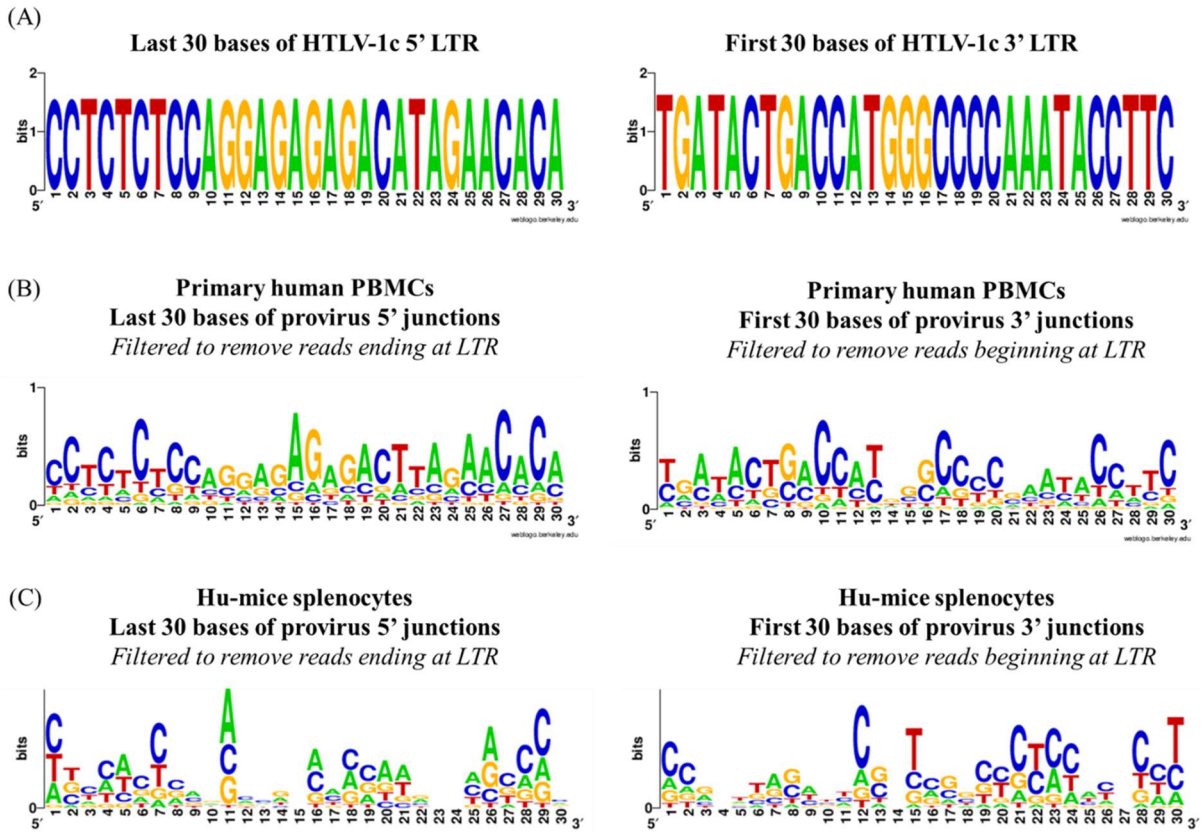


Figure 5.8 – Breakpoint analysis of HTLV-1c defective proviruses from hu-mice model. gDNA was extracted from hu-mice splenocytes infected with HTLV-1c at late time point (6 wpi) and SPA assay was performed by limiting dilution touchdown nested PCR. Individual proviral amplicons were sequenced with the ONT platform and aligned to the HTLV-1c consensus genome. Sequences flanking the breakpoints of the defective provirus with large internal deletions were selected. Nucleotide signals for all sequences are represented as sequence logos. (A) The last 30 bases of the 5' and the first 30 bases of the 3' HTLV-1c LTRs. (B) The last 30 bases of the 5' and the first 30 bases of the 3' HTLV-1c defective provirus segment, filtered to remove any breakpoints that occur at the LTR junctions, for proviruses amplified from primary human PBMCs. (C) The last 30 bases of the 5' and the first 30 bases of the 3' HTLV-1c defective provirus segment, filtered to remove any breakpoints that occur at the LTR junctions, for proviruses amplified from hu-mice splenocytes.

Table 5.2 – Summary and annotations of HTLV-1c-human genome chimeric proviruses amplified from humanised mouse splenocytes.

Provirus amplicon	5' HTLV-1c size (bp)	3' HTLV-1c size (bp)	Internal hg size (bp)	hg Coordinates	Repetitive sequences	Genes	Additional notes
M086 _23	868	103	61	chr12:57,488,680-57,488,739	-	DDIT3-201	-
M086 _02	867	103	61	chr12:57,488,680-57,488,739	-	DDIT3-201	-
M086 _09	866	103	61	chr12:57,488,680-57,488,739	-	DDIT3-201	-
M086 _05	866	103	61	chr12:57,488,680-57,488,739	-	DDIT3-201	-
M086_13	867	103	61	chr12:57,488,680-57,488,739	-	DDIT3-201	-
M086_ 10	79	609	371	No significant alignment	-	NG_046806.1	rDNA
M084_21	360	613	60	chr6:126,968,922-126,968,981	tRNA	-	tRNA-Glu (anticodon CTC) 1-7
M084_09	84	306	69	chr6:28,857,002-28,857,070	tRNA	-	tRNA-Glu (anticodon CTC) 1-6
M084_36	84	306	69	chr6:28,857,002-28,857,070	tRNA	-	tRNA-Glu (anticodon CTC) 1-6
M084_29	91	1120	283	chr14:43,784,833-43,785,115	7SLRNA (srpRNA)	RPS29 (intron); RN7SL1 (exon)	-

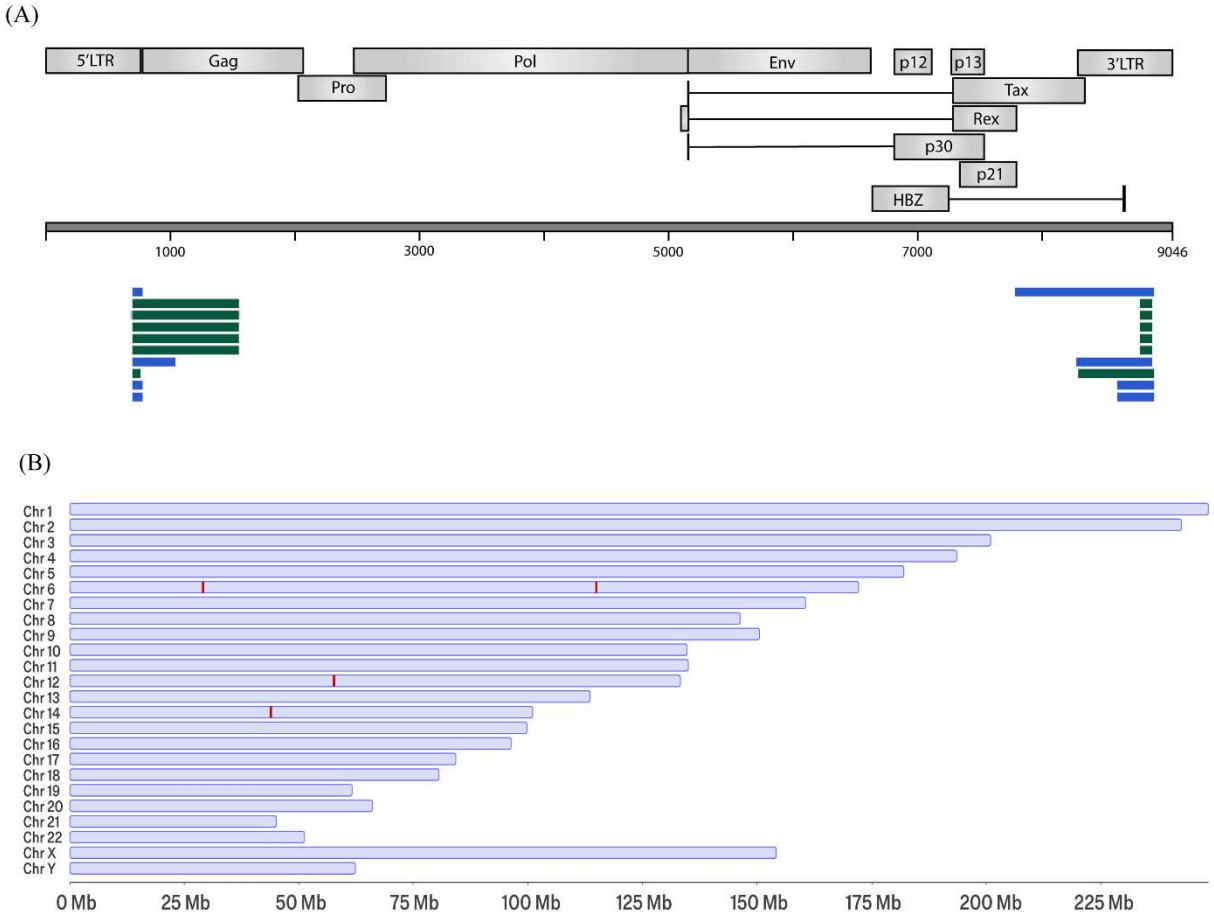


Figure 5.9 – Novel HTLV-1c-human genome chimeric proviruses were amplified from humanised mouse splenocytes. (A) The HTLV-1c-hg chimeric provirus sequences from hu-mice splenocytes, aligned to the HTLV-1c genome, containing internal human sequences (not mapped here to HTLV-1c genome). Blue represents mouse 84, green represents mouse 86. (B) Map of HTLV-1c-human chimeric proviruses to location in hg, indicated by red line.

5.5 Chapter five discussion

This thesis chapter assessed the first HTLV-1c humanised mouse model for its ability to recapitulate in vivo human infection, by emulating the immunophenotyping and proviral landscape studies of primary human PBMCs in the previous chapter four. Overall, the HTLV-1c hu-NSG mouse model retained many aspects of in vivo human infection despite some differences and limitations, which must be noted when assessing the model, and addressed for improvement in future studies.

5.5.1 Immunophenotyping in HTLV-1c hu-mouse model

Remarkably, the immunophenotype of HTLV-1c-infected cells in the hu-mouse model closely mirrored what was found in human donors. HTLV-1c infection does not impart a substantial impact on the frequency of naïve and E/M CD4+ T cells in hu-mice, especially at the early timepoint of infection. Similarly, infection does not greatly impact the proportion of the CCR4- phenotype of E/M CD4+ T cells, in humans or humanised mice.

However, the virus appears to be involved in modulating the frequency of FOXP3+ proxy phenotype in hu-NSG mice, as indicated by the significantly negative correlation to PVL. This supports a similar finding in the preliminary study generating and characterising this HTLV-1c infected hu-NSG mouse model, which showed the proportion of FOXP3+ CD4+ T cells diminished in HTLV-1c infection (Cooney et al., 2023, Cell, in revision). HTLV-1c infection in the humanised mouse model may facilitate downregulation of the expression levels of *foxp3* (and proxy protein markers) on infected or bystander cells, through the expression of viral mRNA and proteins. On the other hand, HTLV-1c infection in the humanised mouse platform may be causing the proliferation of other cell types which inadvertently negatively impacts the proportion of FOXP3+ cells. However it must be considered that in human HTLV-1a infection, FOXP3+ cells do not necessarily equate to functionally adequate regulatory T cells [355].

HTLV-1c infection caused an expansion of the lung homing proxy phenotype in hu-NSG mice, aligning to the trend in human infection, however the proliferation was not as robust. In addition, while the frequency of lung homing proxy phenotype correlated to PVL in human infection, with hu-NSG mice it did not. As these mice do not have

adaptive immunity, this indicates that the significant and proportional expansion of phenotypic lung homing cells in human infection is a result of complex interactions between host immune system and virus, rather than just the virus itself. The immune dysregulation during human HTLV-1c infection presented in chapter three and four, characterised by chronic inflammation and T cell exhaustion, alongside the retention of *hbz* in defective proviruses, likely contributes to the selective expansion of phenotypic lung homing cells, and ultimately the development of HAPD.

Ubiquitously, we saw extensive activation of all CD4+ T cell phenotypes in the HTLV-1c humanised mouse model, consistent with the infection of human donors. However, there was one notable difference: for every CD4+ T cell phenotype in the humanised mouse model, the frequency of activation marker expression was positively correlated to PVL. This is likely because there is no functional adaptive immunity in these mice to oppose the chronic activation and inflammation induced directly by HTLV-1c infection. This is largely consistent with another HTLV-1a humanised mouse study which demonstrated the proportion of activated lymphocytes was positively correlated to *tax* mRNA expression levels [450].

Another consistency between human and humanised mouse HTLV-1c infection lies in the similarity of provirus enrichment patterns across key CD4+ T cell reservoirs. Provirus was significantly enriched in lung homing proxy and CCR4- phenotypes, and similarly higher in FOXP3+ proxy phenotype, in the humanised mouse model at two weeks post infection, suggesting that HTLV-1c shows similar tropism for these cell phenotypes as in humans. On the other hand, at six weeks post infection, the PVL was over 100% in every CD4+ T cell phenotype and bulk splenocytes, so there was no discernible difference between individual reservoirs. The expansion of infected cells in this humanised mouse platform is so profound that by the late infection timepoint there is more than one provirus copy per cell, a phenomenon which has not been reported in human infection.

Therefore, we suggest that the two-week time point is most biologically analogous to human infection in the HTLV-1c humanised mouse model, despite a lack of adaptive immunity, and should be used for future studies when assessing drug treatments or proviral characteristics. At early infection, regardless of the mice being genetically identical, there is variability between individual mice in immune cell frequency and

PVL, replicating the wide range of variability found in a human donor cohort. This likely stems from the initial inoculating pool of HTLV-1c-infected cells injected into the humanised mice, which may have different amounts of defective and replication competent provirus, unlike a cell line which has identical, intact provirus. Furthermore, at early infection, the PVL is between 0.1-5% in bulk splenocytes, which is more closely comparable to what is found in PBMCs of human donors. Going forward, we also suggest it pertinent to include whole blood harvested from hu-mice infected with HTLV-1c, alongside splenocytes, to provide a more direct comparison with human in vivo infection.

5.5.2 Extensive internal proviral deletions and chimeras detected in HTLV-1c-infected humanised mice

Interestingly, the humanised mice showed an extremely high level of defective provirus, with an absence of any breakpoint sequence motifs. While these mice indeed have some levels of humanised innate immunity, which may play a role in how quickly the initial inoculum of irradiated HTLV-1c infected cells transmits infection, they are not capable of amounting an effective adaptive immune response thereafter. This indicates it is not in fact immune selection that is driving the extensive deletions of the provirus in this model, but is more likely due to replicative errors occurring during unmitigated proliferation of infected cells. This highly defective proviral landscape contrasts what was previously reported in an HTLV-1a humanised mouse model, where most proviruses remained full-length [152]. However, this subtype-A study was conducted only using an infected Jurkat cell line for transmission of the virus, and not from irradiated primary human PBMCs from which this subtype-C study originates. Nevertheless, it remains a limitation of the HTLV-1c model that the hu-mice cannot amount an effective immune response, and thus limits the reflection of the host-virus interplay that is so prevalent human infection.

However, the retention of the *pX* region in some hu-mice defective provirus clones, while not to the same extent, reflected the proviral landscape found in human donors, including the genomic region for *hbz*. It is clear this regulatory gene remains important for HTLV-1c pathogenesis, because despite widespread and extensive proviral deletions, infected mice quickly develop disease manifestations such as lymphocytosis and splenomegaly (Cooney et al., 2023, Cell, in revision).

Remarkably, novel chimeric proviruses, which made up 21.5% of the proviral reservoir in human donors, were similarly detected at 14.3% in the humanised mouse model. Some of these chimeric provirus hg sequences include those with a repetitive nature, much like what was found in the human proviral landscape. The presence of chimeric proviruses in both the humanised mouse model and human donors lends itself to the hypothesis that this phenomenon is not occurring through a selective host immune mechanism or response, but most likely from homologous recombination events during proliferation of infected cells.

When an exon comes under the control of the HTLV-1c LTR in a bidirectional manner, serious consequences could arise for the host cell, in human or humanised mice hosts. In particular, HTLV-1c-DDIT3 chimeric provirus, characterised in this chapter, could have a deleterious impact on the host: DDIT3 is a stress-response protein, a dominant-negative inhibitor of many transcription factors. Moreover, DDIT3 was shown to produce chimeric proteins associated with sarcomas in both humans and mice, upon translocation in the genome [455-458], so may likewise have the potential to produce fusion proteins with HTLV-1c components. Furthermore, DDIT3 expression may be modulated by HTLV-1c proteins, as it is known to be upregulated during HIV-1 infection by both Env gp120 [459] and Tat [460, 461] proteins. The understanding of the novel HTLV-1c-hg chimeric proviruses, and their role in persistence and pathogenesis of HTLV-1c infection, is in its infancy, and must be further investigated. The humanised mouse model may be a good tool for preliminary investigations of this nature, particularly when there is limited access to viable primary human samples infected with HTLV-1c from Central Australia.

5.5.3 Future plans to utilise the humanised mouse model to map integration sites and epigenetic modifications

As we would like to investigate the in vivo proviral integration sites of HTLV-1c, but at this stage are unable to sequence First Nations donor DNA due to the sovereignty of Indigenous genomic information and ethics, we plan to utilise the HTLV-1c infected hu-NSG model to circumvent this barrier. The immune systems of these mice are generated through anonymous donor cord blood stem cells, and subsequent infection is through injection of lethally irradiated HTLV-1c patient donor PBMCs, ensuring viral transmission occurs but infected donor cells do not proliferate. Using ONT sequencing

techniques, we will be able to capture the integration site and entire proviral sequence of each infected cell. Furthermore, we plan to investigate the epigenetic modification landscape of infected cells at both early and late timepoints. New techniques have been developed since previous integration site and epigenetic studies of HTLV-1a were published [227, 462-464], and as such this study could be informing for both subtypes. Both our human and humanised mice studies to date have lacked a direct comparison with HTLV-1a samples, and thus relied on comparisons to reported literature, so we hope to use infection of hu-NSG mice with an MT-2 cell line to assess subtype-A alongside -C.

While humanised mouse models have self-regeneration of human HSCs, which in turn can differentiate into lymphocytes, monocytes, macrophages, DCs and natural killer (NK) cells, their main limitation is the inability generate a robust adaptive immune response to infection [465]. Human T and B lymphocytes are educated on mouse thymic stromal cells and thus are restricted to mouse MHC, preventing efficient interactions between MHC and human antigen presenting cells (APC) [465]. One such possibility for improving the immune response of these hu-NSG mice to help control HTLV-1c infection and more closely mimic human in vivo infection, would be to passively immunise them with anti-HTLV-1 antibodies prior to inoculation. This would provide some level of immunity with which to study transmission, replication, and gene expression of HTLV-1c. Moreover, humanised mouse models are continually improving and expanding in their capabilities [466], and so future HTLV-1c work may look to include these new models, such as HUMAMICE [467]. These mice have a highly functioning immune system: they express all human HLA molecules and can produce immunization-induced viral-specific antibodies [467]. Lastly, other animal models such as NHPs may be considered to further validate and advance the findings in this thesis chapter, in a setting of functional adaptive immunity.

5.5.4 Chapter five conclusion

Humanised mouse models are invaluable in gaining understanding of infectious disease transmission and pathogenesis, and providing a platform with which to assess prevention and therapeutic treatments in pre-clinical studies. The HTLV-1c humanised mouse model recapitulates many important aspects of human in vivo infection, with similar immunophenotype and activation patterns of CD4+ T cells, albeit with minimal human

cellular and humoral immune responses of the mice. This HTLV-1c hu-NSG mouse model will continue to be an important tool to understand the proviral reservoir and pathogenesis, and test the efficacy of treatments, to ultimately progress the care and treatment for those infected with HTLV-1c in Australia and prevent further spread of this insidious lifelong disease.

6. General thesis discussion

6.1 The importance of HTLV-1c research in Australia

6.1.1 General comments

This thesis addresses current gaps in knowledge surrounding the molecular virology and immunological responses to HTLV-1c infection in Australia. This research has a high priority given the impact of this stealthy lifelong infection: there are extraordinarily high HTLV-1c incidence rates in Central Australian First Nations communities [254], a high prevalence of HAPD [9, 10, 15, 17], higher health service utilisation [17] and mortality with infection [14, 256], and prominent genetic variation in the *pX* region compared to subtype-A [203].

HTLV-1 subtype-C in Australia has co-existed and co-evolved with humans since human habitation of the continent more than 50,000 years ago [1]. Hence, various selective and evolutionary pressures have shaped the nature and genomics of HTLV-1c virus infection to persist, proliferate and transmit, while the host responses to infection have adapted for transmission without extinguishing the infected populations. Viral pathogenesis may indeed be a coincidental by-product of the virus' ecological selection for transmission and persistence in the host population, which it has done with remarkable success for tens of thousands of years. This is evident when comparing the outcome for HTLV-1 infected T cells, which undergo extraordinary proliferation, with HIV-1 infected T cells that sustain a marked death and depletion. However, this selection HTLV-1 has attained for infected cell proliferation and viral persistence in individuals within culturally isolated communities has likely led to it being overlooked by medical science, since it's discovery in 1980 [391]. Indeed, the World Health Organisation (WHO) has only recently (August 2022) published a publicly available fact sheet on HTLV-1. Nevertheless, as the HTLV-1 research community has uncovered many insidious impacts of infection, it is time that more action is taken to combat this prevalent disease, especially in Australia.

6.1.2 HTLV-1c infection is part of a complex health disparity in Central Australia

Given many HTLV-1 associated diseases take many years to develop, it is difficult to distinguish sub-clinical stages of HTLV-1 associated diseases. This is a limitation of any study that compares HTLV-1 disease and asymptomatic groups, when indeed the ACs

may not be so. This is especially relevant in Central Australia, where the plethora of co-morbidities and infections may confound symptoms and prognostic immune biomarkers. In addition, immunological imbalance and dysfunction in all infected individuals is hardly asymptomatic, and likely one of the reasons why those infected with HTLV-1c have poorer overall health outcomes and higher mortality rates. For these reasons, we have defined the two HTLV-1c seropositive groups throughout this study as HAPD- and HAPD+.

The most common outcomes of HTLV-1c infection in Central Australia are bronchiectasis, other HAPDs, and BSIs, but in addition, less frequently there have been rare, reported cases of ATL and HAM that are the serious conditions with a confirmed HTLV-1 aetiology [17]. We cannot say for certain whether apparently lower rates of slow developing diseases like ATL and HAM among HTLV-1 subtype-C infected First Nations people in Central Australia results from the faster onset of respiratory inflammation that lowers life expectancy such that ATL and HAM do not fully develop within the 10-20 year shorter lifetime [468]. Alternatively, as the rates of mother-to-child transmission of HTLV-1c in Central Australia are unknown, it is possible that transmission during adolescence or adulthood does not allow time enough for ATL to develop. On the other hand, perhaps the summative effect of subtype-C virology, alongside host and environmental factors and triggers, results in the pulmonary disease trend. Nevertheless, this insidious infection that we show here to underpin chronic immune activation is clearly contributing to a very complex problem in Central Australia, of poorer health outcomes and shorter life expectancy of First Nations peoples.

6.2 Novel characteristics of HTLV-1c infection in Central Australia

6.2.1 Chronic T cell activation in all HTLV-1c infection

High levels of T cell activation in all HTLV-1c seropositive individuals were defined in chapters 3 and 4 of this thesis by elevated expression of cell surface activation markers in all effector/memory cell phenotypes. This demonstrates that HTLV-1c infection is not silent, even in the absence of a defined associated disease. Chronic immune activation of

HTLV-1c infected individuals stems from ongoing exposure to HTLV-1c viral mRNA and proteins, as well as persistent stimulation with antigens from co-infections, which are common in the resource-poor setting of remote communities in Central Australia. Chronic activation leads to T cell exhaustion, resulting in loss of T cell effector functions, including in virus-specific cells [282]. Indeed, for several chronic infectious pathogens, such as the Retroviruses and Herpesviruses, high viral load coinciding with low availability of CD4+ Th cells is often correlated with severe exhaustion phenotype, due to diminished effective CTL activity [469]. The higher prevalence of co-morbidities, health service utilisation and all-cause mortality for all HTLV-1c infected individuals in Central Australia likely stems from the HTLV-1c induced chronic T cell activation and exhaustion presented in this thesis.

6.2.2 HTLV-1c bronchiectasis is characterised by immune exhaustion

We presented a unique profile of the host cytokine network in HTLV-1c infection and HAPD, to what is reported for HTLV-1a. In HTLV-1a infection, ATL patients present an immunosuppressive profile characterised by high IL-10 and TGF-β, and low IFN-γ, while HAM donors exhibit pro-inflammatory profiles characterised by high IFN-γ, IL-2 and TNF-α (reviewed in [239]). While we initially hypothesized that in HTLV-1c infection, HAPD+ donors would exhibit a pro-inflammatory profile, something which has been proposed in the literature but not investigated in vivo [128], we instead uncovered a dysregulated cytokine network that has an immunosuppressive phenotype, linked to progressive T cell exhaustion.

The stages of chronic inflammation leading to T cell exhaustion are embodied in the results of our study. When considering the host cytokine response in all HTLV-1c seropositive individuals compared to seronegative donors, we showed many elevated pro-inflammatory cytokines, a profile that is common in many chronic viral infections [282]. We detected elevated levels of TNF-α, involved in cell activation and production of other pro-inflammatory cytokines [319]; IP-10, also involved in activating immune cells and assisting in cell migration; and IL-2, which supports the proliferation and survival of infected cells [320]. Meanwhile, elevated levels of the immunosuppressive IL-10, alongside lower levels of pro-inflammatory IFN-γ and IL-8 cytokines, contrasts the inflammatory response, and is a pattern that begins to emerge in the later chronic phase of viral infections [282]. It may be that HAPD- individuals can maintain a balance

between the pro-inflammatory response to HTLV-1c infection and contrasting immunosuppression, to prevent ensuing T cell exhaustion.

On the other hand, the cytokine networks in HAPD+ donors revealed a complete turn to immune exhaustion, as summarised in **Table 6.1**. We characterised high levels of IL-10 and IL-6, with low levels of IFN-γ and IL-8, associated with Bex, a profile that is consistent with other chronic conditions leading to immune exhaustion, such as long-COVID [470]. The cytokine network presented leads to loss of T cell effector functions, facilitating HAPD pathogenesis. Upregulation of IL-10 production in HAPD+ patients inhibits the function of HTLV-1c-specific effector cells [328], and further dysregulates Th1 response by downregulating IFN-γ production and Th1 polarisation [331]. Moreover, IL-10 inhibits the expression of MHC-II on monocytes and macrophages [329]. IL-10 inhibits the CTL response and promotes T cell exhaustion by inducing the expression of programmed death ligand 1 (PD-L1) on DCs and PD-1 expression on CTLs [471], and similarly, high levels of IL-6 during chronic viral infections [472] . Elevated levels of IL-6 also inhibit the secretion of IFN-γ from CD4+ T cells, further downregulating Th1 polarisation [330]. IL-6 also inhibits AICD [324], promoting the survival of infected cells. Low levels of IL-8 decrease the activation of neutrophils, and therefore the migration of monocytes, which typically produce pro-inflammatory cytokines [321]. So HTLV-1c-infected cells accumulating at lung tissue sites in bronchiectasis no longer face an effective antiviral pro-inflammatory response. Similarly, lower IFN-γ levels in HAPD+ patients decreases the overall inflammatory response and the differentiation of Th1 cells [473]. Indeed, Th1 dysfunction is characterised in chapter three and four of this thesis, likely stemming from this viral-induced T cell exhaustion.

The upregulation of inhibitory receptors such as PD-1 and CTLA-4, among others, is indeed another classic signature of T cell exhaustion, and is something that should be explored in future studies surrounding the pathogenesis of HTLV-1c and HAPD, along with studies of larger cohorts, to confirm these findings.

6.2.3 *hbz* gene drives HTLV-1c-associated bronchiectasis development

In chapter four and five of this thesis, the deep characterisation of the integrated provirus in HTLV-1c patients and a humanised mouse model shows the predominant proviral landscape has extensive internal deletions in all HTLV-1c infected individuals and humanised mice. These deletions are likely caused by homologous recombinational

deletion during cellular proliferation [374-376, 383], and the breakpoints of defective provirus in humans share LTR-like sequence motifs.

Notably, we showed that the defective provirus in HAPD+ donors preferentially retain the *pX* region, and in particular the entire antisense *hbz* gene region, indicating that *hbz* expression of active RNA and protein is implicated in Bex disease, as summarised in **Table 6.1**. HBZ protein upregulates CCR4 expression [370], increasing the migration of infected and activated cells to sites of inflammation, like the respiratory tract. *Hbz* gene retention also promotes the proliferation of infected cells through *hbz* RNA induced upregulation of transcription factor E2-F1 [182] and inhibition of the non-canonical NF-κB pathway through HBZ protein activity [185, 186]. This fosters a positive feedback loop producing an expansion of cells with provirus that retain the *hbz* gene. *Hbz* expression further promotes the survival of these cells, by inhibiting AICD through actions of HBZ protein [474] and upregulating the anti-apoptotic gene *survivin* through *hbz* RNA activity [191]. HBZ also impairs cell-mediated immunity by impairing the functional capacity of Tregs [190] and downregulating the transcription of Th1 pro-inflammatory cytokines (IFN-γ, TNF-α, IL-2) [189]. Finally, HBZ protein induces expression of the coinhibitory molecule TIGIT to decrease the cell-mediated response to HTLV-1c infection [235]. Ultimately, the combined effect of immune exhaustion and provirus retention of *hbz* drives the development of HAPD, as outlined in **Table 6.1**.

Table 6.1 – Summary of the components of HTLV-1c infection that contribute to immune exhaustion and pathogenesis in associated bronchiectasis.

Component	Outcome	Effector mechanism
Retention of *hbz* gene in defective provirus	Cell survival	Inhibits AICD[474]; Upregulates *survivin*[191]
	Impairs cell-mediated immunity	Downregulates transcription of Th1 cytokines[189]; Induces co-inhibitory molecule TIGIT[235]
	Proliferation of infected cells	Inhibits non-canonical NF-κB pathway[185, 186]; Upregulates transcription factor *E2-F1*[182]
	Migration of cells to sites of inflammation	Upregulates CCR4 expression[370]
High IL-6	Cell survival	Inhibits AICD[324]
	Impairs cell-mediated immunity	Impairs CTL response[325]
	Migration of cells to sites of inflammation	Upregulates migration of activated and virus-specific cells[326]
	Positive feedback loop	Produced in tissue damage during viral infections[323]
High IL-10	Impairs cell-mediated immunity	Inhibits function of virus specific cells[328]; Induces expression of PD-L1 on DCs and PD-1 on CTLs[471]; Inhibits expression of MHC-II, ICAM-1 on APCs[329]; Downregulates IFN-γ production and Th1 polarisation[331]
Low IFN-γ	Impairs cell-mediated immunity	Decreases Th1 differentiation[473]; Inhibits CTL response[310, 475]
Low IL-8	Impairs cell-mediated immunity	Decreases activation of neutrophils and production of pro-inflammatory cytokines[321]

AICD, activation induced cell death; TIGIT, T cell immunoreceptor with immunoglobulin and ITIM domain; NF-κB, Nuclear factor kappa B; PD-L1, programmed death ligand 1; DC, dendritic cell; CTL, cytotoxic lymphocyte; PD-1, programmed cell death protein 1; MHC-II, major histocompatibility complex class II; ICAM-1, intracellular adhesion molecule 1; IFN-γ, interferon gamma; Th1, Type 1 helper T cell.

6.2.4 CCR4- CD4+ T cells may be a persistent reservoir of HTLV-1c provirus

The detection of exclusively full-length infectious provirus in the CCR4- effector/memory CD4+ T cell subpopulation in blood of one patient suggests that these cells do not undergo cell division and possibly lack the molecular mechanisms that drive proviral deletions. This thesis showed that these cells do not expand in HTLV-1c infection, and thus retain the full-length provirus because they have a low rate of division and low occurrence of the homologous recombinational deletion events. Importantly, these cells could therefore become the dominant cell population needed for viral transmission. CCR4- cells do not express the tissue homing phenotypic markers and are likely to remain circulating in peripheral blood where they provide a transmission source for infectious virus. Transmissible virus harboured in peripherally circulating T cells may avoid contact with the tissue microenvironments implicated in disease like the lungs, where the cellular anti-HTLV-1c immune response is focused. Further studies are needed to determine the molecular mechanisms that allow the accumulation of intact proviruses. In addition, it will be pertinent to investigate whether these cells are present in transmission fluid, such as sexual secretions and breast milk, as the events that lead to the transfer of HTLV-1-infected CCR4- cells may be central to the ecology of viral transmission.

6.2.5 Novel chimeric HTLV-1c-hg provirus may contribute to immune dysfunction and pathogenesis

The characterisation of novel HTLV-1c-hg chimeric provirus in human donors and humanised mice in chapters four and five of this thesis presents a new component of the host-virus interplay that may contribute to pathogenesis. It has previously been reported that HTLV-1 provirus can interact with proximal or distant host genes through CTCF binding, resulting in a disruption to host and viral gene expression (reviewed in [391]). Similarly, direct interactions between the HTLV-1c and hg components of these chimeric proviruses characterised in this thesis, may have profound impacts on the regulation of gene expression and immune cell dysfunction. A deeper characterisation of these chimeric proviruses, and the potential chimeric mRNAs and proteins produced, is required to fully understand the implications of this discovery.

6.3 Translational impacts of HTLV-1c pathogenesis research

6.3.1 Improvements to HTLV-1 PVL assay in Australian reference laboratories

The preferential retention of the *pX* region in the defective proviral landscape informed the development of a new standard testing regimens in collaboration with the National Reference Laboratory (NRL) in Melbourne. Diagnostic testing at the NRL has historically used enzyme immunoassay and western blot, as well as confirmatory PVL when required, measured by qPCR targeting the *gag* region. The development of a new *pX* region target in collaboration with the NRL has resulted in increased sensitivity of the HTLV-1 PVL assay (Vandegraaff et al., unpublished), and it will be implemented as the standard confirmatory test in reference laboratories in Australia. For cases that are serology indeterminant, this PVL assay targeting the *pX* region will increase the level of accurate diagnoses, and ultimately patient support and public health surveillance of HTLV-1c in Australia.

6.3.2 Viral therapeutics should target *hbz* expression

HTLV-1, unlike other retroviruses, is comparatively genetically stable, likely owing to the majority of viral replication occurring through proliferation of infected cells rather than reverse transcription during de novo infection from cell-free and cell-associated virions. The relative genomic stability suggests that future therapeutic and prevention treatments may have significant cross-clade efficacy. It is therefore imperative that approaches consider the nuances of the molecular virology, pathogenesis, and elicited host responses of the different HTLV-1 subtypes – investigations that are addressed in this thesis for subtype-C.

This thesis shows that the HTLV-1c proviral landscape in both HAPD- and HAPD+ donors acquires extensive internal deletions, highlighting that viral therapeutics should be designed to target the *pX* genes that predominantly endure. In particular, strategies should focus on targeting *hbz* gene expression and the spliced *hbz* mRNA and sHBZ protein, as this coding region is most frequently preserved even in the defective proviruses. Indeed, the outcome of HTLV-1a infection is linked to the ability of HLA-I to present HBZ epitopes to CTLs [476]. Therefore, if HTLV-1 infected cells with

proviruses retaining *hbz* can be eliminated through therapeutic targeting, then pathogenesis and disease progression may be mitigated. Targeting sense strand gene expression may not be as effective, as exon 1 of the doubly spliced RNA transcripts, such as *tax*, *rex* and *p30*, are mostly deleted in this setting. Moreover, the effect of these deletions on the expression of singly spliced forward RNA transcripts *p12/p8* and *p13* have not yet been investigated, however subtype-C contains a start codon mutation for *p12/p8* (Hirons et al., unpublished) that is located the forward sense genomic region of the retained antisense *hbz*. Further studies are needed to determine the molecular mechanisms, speed and immune impact of the accumulating defective proviruses.

6.3.3 The humanised mouse model of HTLV-c infection will be an important tool for pre-clinical testing of preventative and therapeutic drugs

Humanised mouse models will continue to be an important tool for pre-clinical testing of drugs targeting HTLV-1 or blocking transmission. The humanised mouse model evaluated in chapter five of this thesis recapitulates many aspects of human in vivo infection, including the preferential expansion of phenotypic lung homing cells, but none in FOXP3 proxy or CCR4- CD4+ T cell reservoirs. Furthermore, these reservoirs are highly infected with HTLV-1c provirus, similar to human infection. In addition, the extensive immune cell activation we reported for human infection is also present in the humanised mouse model. While the limitations of this model stem from the lack of adaptive immunity, and therefore host-virus interplay, it is still a valuable tool for pre-clinical testing of therapeutic and prevention strategies against HTLV-1c.

Notably, the hu-NSG model evaluated in this thesis has been used to demonstrate that two drugs which are already safety-approved for HIV therapy, tenofovir and dolutegravir, show some efficacy in reducing or preventing transmission (Cooney et al., 2023, Cell, in revision). Moreover, targeting the antiapoptotic myeloid cell leukemia 1 (MCL-1) in combination with antiretrovirals in the hu-NSG mouse model has shown promise in reducing PVL in established infection (Cooney, et al, 2023, Cell, in revision). Other animal models that more closely relate to human infection, such as macaques, may be considered going forwards for validation of these promising results, where infection is naturally mitigated by a functioning host immune response to HTLV-1.

6.3.4 Changes are required to improve the HTLV-1 public health strategy in Australia

Given the human transmission modes of HTLV-1 are like other blood borne and sexually transmitted viral infections, such as HIV-1, it seems pertinent that very similar approaches could easily be employed to reduce the spread of such a consequential infection, without even having to develop HTLV-1-specific therapeutic and prevention treatments. This is especially important in Australia, given it is home to one of the highest known endemic incidence rates worldwide [254]. Confronting Australia's concentrated HTLV-1c infections requires several coordinated responses. Firstly, more widespread dissemination of information surrounding HTLV-1c is essential, including to the general public, especially those residing in at-risk communities; to healthcare practitioners, particularly those who service individuals from endemic areas; and to policy makers. Moreover, this information needs to be tailored to each audience and delivered in a culturally appropriate and sensitive manner. Secondly, simple behavioural or lifestyle approaches that can reduce the incidence of HTLV-1c transmission need to be clearly explained, encouraged, and have community and government support frameworks implemented to achieve them. For example, condoms could be made available from health care centres in endemic areas, or baby feeding bottles and formula, or provide equipment to freeze and thaw expressed breastmilk to kill HTLV-1 infected cells, as basic measures to help reduce transmission risk. Indeed, the implementation of similar simple public health measures in Japan saw an 80% reduction in HTLV-1a transmission [82, 477, 478]. In addition, antenatal screening for HTLV-1c must be included in patient care, to further reduce vertical transmission, especially in endemic areas, such as what has been proposed in Brazil [479]. Lastly, there needs to be a support network in place for those that are diagnosed with HTLV-1 in Australia, and a clear set of guidelines of the sequential steps that follow such a diagnosis, for health care practitioners and patients alike. Australia is well behind on appropriate public health strategy, investment in research and patient care – something that must be changed to avoid viral infections spreading from foci with high endemicity in remote locations into larger population centres, reducing transmission within the endemic regions, and improving the lives of those infected with HTLV-1c.

6.4 Concluding thesis remarks

This thesis presented a unique characterisation of the proviral landscape and host immune response of HTLV-1c infection, knowledge that will underpin future design and development of effective prevention and therapeutic treatments. HTLV-1 has evolved to have strategic expression of viral genes in cells of the immune system, and consequent modulation of host immune phenotype and function. Successful viral transmission and lifetime persistence in the host have shaped the dynamic proviral landscape outlined in this thesis. The remarkable ability of HTLV-1c to avoid immune clearance of full-length infectious provirus while maintaining or increasing the reservoir of defective genomes that modulate immune function and link to disease provides an obstacle that must be overcome when developing future treatment and prevention strategies. The level of HTLV-1c provirus established in the host is not silent, even when most of the genes that make up the infectious virus are deleted over time, imparting a persistent impact on the host immune function. The immune system dysregulation and exhaustion demonstrated in HTLV-1c infection is likely influenced by a complex intertwined web of factors including PVL, co-morbidities and co-infections, effectiveness and robustness of host immune response, host genetic characteristics, the repertoire of intact and defective proviruses and consequential viral gene expression, proviral integration site and proximal or distal host genes, and other factors yet to be identified. Regardless, the identification of significant increased unexplained mortality and reduction in life span of all HTLV-1 positive individuals shows that even in apparently asymptomatic individuals, HTLV-1 infection is not silent, and the virus exerts complex and subtle impacts on long-term health.

As Central Australia faces some of the highest endemic rates of HTLV-1 infection of anywhere in the world, mitigation strategies should be placed atop the national agenda. This is a country with a renowned healthcare system and lucrative scientific research precincts, so tackling the prevalent HTLV-1c infection in Australia is absolutely within our collective reach. This must be approached with careful thought and consideration: working in partnership with First Nations communities and individuals, to combine clinical, scientific and cultural knowledge to deliver treatment, testing and support on country, for those affected by HTLV-1c.

References

1. Gessain A, Cassar O: **Epidemiological aspects and world distribution of HTLV-1 infection.** *Front Microbiol* 2012, **3:**388.
2. Yoshida M, Miyoshi I, Hinuma Y: **Isolation and characterization of retrovirus from cell lines of human adult T-cell leukemia and its implication in the disease.** *Proc Natl Acad Sci USA* 1982, **79:**2031-2035.
3. Kondo T, Kono H, Miyamoto N, Yoshida R, Toki H, Matsumoto I, Hara M, Inoue H, Inatsuki A, Funatsu T, et al.: **Age- and sex-specific cumulative rate and risk of ATLL for HTLV-I carriers.** *Int J Cancer* 1989, **43:**1061-1064.
4. Koga Y, Iwanaga M, Soda M, Inokuchi N, Sasaki D, Hasegawa H, Yanagihara K, Yamaguchi K, Kamihira S, Yamada Y: **Trends in HTLV-1 prevalence and incidence of adult T-cell leukemia/lymphoma in Nagasaki, Japan.** *J Med Virol* 2010, **82:**668-674.
5. Gessain A, Vernant JC, Maurs L, Barin F, Gout O, Calender A, De Thé G: **Antibodies to human T-lymphotrophic virus type-I in patients with tropical spastic paraparesis.** *The Lancet* 1985, **326:**407-410.
6. Kaplan JE, Osame M, Kubota H, Igata A, Nishitani H, Maeda Y, Khabbaz RF, Janssen RS: **The risk of development of HTLV-I-associated myelopathy/tropical spastic paraparesis among persons infected with HTLV-I.** *J Acquir Immune Defic Syndr (1988)* 1990, **3:**1096-1101.
7. Maloney EM, Cleghorn FR, Morgan OS, Rodgers-Johnson P, Cranston B, Jack N, Blattner WA, Bartholomew C, Manns A: **Incidence of HTLV-I-associated myelopathy/tropical spastic paraparesis (HAM/TSP) in Jamaica and Trinidad.** *J Acquir Immune Defic Syndr Hum Retrovirol* 1998, **17:**167-170.
8. Mochizuki M, Watanabe T, Yamaguchi K, Takatsuki K, Yoshimura K, Shirao M, Nakashima S, Mori S, Araki S, Miyata N: **HTLV-I uveitis: a distinct clinical entity caused by HTLV-I.** *Jpn J Cancer Res* 1992, **83:**236-239.
9. Einsiedel L, Cassar O, Goeman E, Spelman T, Au V, Hatami S, Joseph S, Gessain A: **Higher human T-lymphotropic virus type 1 subtype C proviral loads are associated with bronchiectasis in indigenous australians: results of a case-control study.** *Open Forum Infect Dis* 2014, **1:**ofu023.
10. Einsiedel L, Fernandes L, Spelman T, Steinfort D, Gotuzzo E: **Bronchiectasis is associated with human T-lymphotropic virus 1 infection in an Indigenous Australian population.** *Clin Infect Dis* 2012, **54:**43-50.
11. Einsiedel LJ, Pepperill C, Wilson K: **Crusted scabies: a clinical marker of human T-lymphotropic virus type 1 infection in central Australia.** *Med J Aust* 2014, **200:**633-634.
12. Mollison LC, Lo ST, Marning G: **HTLV-I and scabies in Australian aborigines.** *Lancet* 1993, **341:**1281-1282.
13. Einsiedel L, Cassar O, Spelman T, Joseph S, Gessain A: **Higher HTLV-1c proviral loads are associated with blood stream infections in an Indigenous Australian population.** *J Clin Virol* 2016, **78:**93-98.
14. Schierhout G, McGregor S, Gessain A, Einsiedel L, Martinello M, Kaldor J: **Association between HTLV-1 infection and adverse health outcomes: a systematic review and meta-analysis of epidemiological studies.** *Lancet Infect Dis* 2020, **20:**133-143.

15. Einsiedel LJ, Pham H, Woodman RJ, Pepperill C, Taylor KA: **The prevalence and clinical associations of HTLV-1 infection in a remote Indigenous community.** *Med J Aust* 2016, **205:**305-309.
16. Einsiedel L, Woodman RJ, Flynn M, Wilson K, Cassar O, Gessain A: **Human T-Lymphotropic Virus type 1 infection in an Indigenous Australian population: epidemiological insights from a hospital-based cohort study.** *BMC Public Health* 2016, **16:**787.
17. Einsiedel L, Spelman T, Goeman E, Cassar O, Arundell M, Gessain A: **Clinical associations of Human T-Lymphotropic Virus type 1 infection in an indigenous Australian population.** *PLoS Negl Trop Dis* 2014, **8:**e2643.
18. Poiesz BJ, Ruscetti FW, Gazdar AF, Bunn PA, Minna JD, Gallo RC: **Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma.** *Proc Natl Acad Sci USA* 1980, **77:**7415-7419.
19. Le Blanc I, Grange MP, Delamarre L, Rosenberg AR, Blot V, Pique C, Dokhélar MC: **HTLV-1 structural proteins.** *Virus Res* 2001, **78:**5-16.
20. Nam SH, Copeland TD, Hatanaka M, Oroszlan S: **Characterization of ribosomal frameshifting for expression of pol gene products of human T-cell leukemia virus type I.** *J Virol* 1993, **67:**196-203.
21. Seiki M, Hattori S, Hirayama Y, Yoshida M: **Human adult T-cell leukemia virus: complete nucleotide sequence of the provirus genome integrated in leukemia cell DNA.** *Proc Natl Acad Sci USA* 1983, **80:**3618-3622.
22. Center RJ, Kobe B, Wilson KA, Teh T, Howlett GJ, Kemp BE, Poumbourios P: **Crystallization of a trimeric human T cell leukemia virus type 1 gp21 ectodomain fragment as a chimera with maltose-binding protein.** *Protein Sci* 1998, **7:**1612-1619.
23. Wallin M, Ekström M, Garoff H: **Isomerization of the intersubunit disulphide-bond in Env controls retrovirus fusion.** *Embo j* 2004, **23:**54-65.
24. Hattori S, Kiyokawa T, Imagawa K, Shimizu F, Hashimura E, Seiki M, Yoshida M: **Identification of gag and env gene products of human T-cell leukemia virus (HTLV).** *Virology* 1984, **136:**338-347.
25. Christensen AM, Massiah MA, Turner BG, Sundquist WI, Summers MF: **Three-dimensional structure of the HTLV-II matrix protein and comparative analysis of matrix proteins from the different classes of pathogenic human retroviruses.** *J Mol Biol* 1996, **264:**1117-1131.
26. Khorasanizadeh S, Campos-Olivas R, Summers MF: **Solution structure of the capsid protein from the human T-cell leukemia virus type-I.** *J Mol Biol* 1999, **291:**491-505.
27. Mitchell MS, Tözsér J, Princler G, Lloyd PA, Auth A, Derse D: **Synthesis, processing, and composition of the virion-associated HTLV-1 reverse transcriptase.** *J Biol Chem* 2006, **281:**3964-3971.
28. Ruggero K, Guffanti A, Corradin A, Sharma VK, De Bellis G, Corti G, Grassi A, Zanovello P, Bronte V, Ciminale V, D'Agostino DM: **Small noncoding RNAs in cells transformed by human T-cell leukemia virus type 1: a role for a tRNA fragment as a primer for reverse transcriptase.** *J Virol* 2014, **88:**3612-3622.
29. Richardson JH, Edwards AJ, Cruickshank JK, Rudge P, Dalgleish AG: **In vivo cellular tropism of human T-cell leukemia virus type 1.** *J Virol* 1990, **64:**5682-5687.

30. Jones KS, Petrow-Sadowski C, Bertolette DC, Huang Y, Ruscetti FW: **Heparan sulfate proteoglycans mediate attachment and entry of human T-cell leukemia virus type 1 virions into CD4+ T cells.** *J Virol* 2005, **79:**12692-12702.
31. Piñon JD, Klasse PJ, Jassal SR, Welson S, Weber J, Brighty DW, Sattentau QJ: **Human T-cell leukemia virus type 1 envelope glycoprotein gp46 interacts with cell surface heparan sulfate proteoglycans.** *J Virol* 2003, **77:**9922-9930.
32. Ghez D, Lepelletier Y, Lambert S, Fourneau JM, Blot V, Janvier S, Arnulf B, van Endert PM, Heveker N, Pique C, Hermine O: **Neuropilin-1 is involved in human T-cell lymphotropic virus type 1 entry.** *J Virol* 2006, **80:**6844-6854.
33. Manel N, Kim FJ, Kinet S, Taylor N, Sitbon M, Battini J-L: **The Ubiquitous Glucose Transporter GLUT-1 Is a Receptor for HTLV.** *Cell* 2003, **115:**449-459.
34. Yasunaga J, Sakai T, Nosaka K, Etoh K, Tamiya S, Koga S, Mita S, Uchino M, Mitsuya H, Matsuoka M: **Impaired production of naive T lymphocytes in human T-cell leukemia virus type I-infected individuals: its implications in the immunodeficient state.** *Blood* 2001, **97:**3177-3183.
35. Nagai M, Brennan MB, Sakai JA, Mora CA, Jacobson S: **CD8(+) T cells are an in vivo reservoir for human T-cell lymphotropic virus type I.** *Blood* 2001, **98:**1858-1861.
36. Macatonia SE, Cruickshank JK, Rudge P, Knight SC: **Dendritic cells from patients with tropical spastic paraparesis are infected with HTLV-1 and stimulate autologous lymphocyte proliferation.** *AIDS Res Hum Retroviruses* 1992, **8:**1699-1706.
37. Koyanagi Y, Itoyama Y, Nakamura N, Takamatsu K, Kira J, Iwamasa T, Goto I, Yamamoto N: ***In vivo*** **infection of human T-cell leukemia virus type I in non-T cells.** *Virology* 1993, **196:**25-33.
38. Jones KS, Fugo K, Petrow-Sadowski C, Huang Y, Bertolette DC, Lisinski I, Cushman SW, Jacobson S, Ruscetti FW: **Human T-cell leukemia virus type 1 (HTLV-1) and HTLV-2 use different receptor complexes to enter T cells.** *J Virol* 2006, **80:**8291-8302.
39. Lambert S, Bouttier M, Vassy R, Seigneuret M, Petrow-Sadowski C, Janvier S, Heveker N, Ruscetti FW, Perret G, Jones KS, Pique C: **HTLV-1 uses HSPG and neuropilin-1 for entry by molecular mimicry of VEGF165.** *Blood* 2009, **113:**5176-5185.
40. Ghez D, Lepelletier Y, Jones KS, Pique C, Hermine O: **Current concepts regarding the HTLV-1 receptor complex.** *Retrovirology* 2010, **7:**99.
41. Maertens GN, Hare S, Cherepanov P: **The mechanism of retroviral integration from X-ray structures of its key intermediates.** *Nature* 2010, **468:**326-329.
42. Matreyek KA, Engelman A: **Viral and cellular requirements for the nuclear entry of retroviral preintegration nucleoprotein complexes.** *Viruses* 2013, **5:**2483-2511.
43. Li M, Mizuuchi M, Burke TR, Jr., Craigie R: **Retroviral DNA integration: reaction pathway and critical intermediates.** *Embo j* 2006, **25:**1295-1304.
44. Demontis MA, Sadiq MT, Golz S, Taylor GP: **HTLV-1 viral RNA is detected rarely in plasma of HTLV-1 infected subjects.** *J Med Virol* 2015, **87:**2130-2134.
45. Jeang KT, Boros I, Brady J, Radonovich M, Khoury G: **Characterization of cellular factors that interact with the human T-cell leukemia virus type I p40x-responsive 21-base-pair sequence.** *J Virol* 1988, **62:**4499-4509.
46. Harrod R, Tang Y, Nicot C, Lu HS, Vassilev A, Nakatani Y, Giam CZ: **An exposed KID-like domain in human T-cell lymphotropic virus type 1 Tax is**

responsible for the recruitment of coactivators CBP/p300. *Mol Cell Biol* 1998, **18:**5052-5061.
47. Olivares E, Landry DM, Cáceres CJ, Pino K, Rossi F, Navarrete C, Huidobro-Toro JP, Thompson SR, López-Lastra M: **The 5' untranslated region of the human T-cell lymphotropic virus type 1 mRNA enables cap-independent translation initiation.** *J Virol* 2014, **88:**5936-5955.
48. Rayne F, Bouamr F, Lalanne J, Mamoun RZ: **The NH2-terminal domain of the human T-cell leukemia virus type 1 capsid protein is involved in particle formation.** *J Virol* 2001, **75:**5277-5287.
49. Konvalinka J, Kräusslich HG, Müller B: **Retroviral proteases and their roles in virion maturation.** *Virology* 2015, **479-480:**403-417.
50. Mellors JW, Muñoz A, Giorgi JV, Margolick JB, Tassoni CJ, Gupta P, Kingsley LA, Todd JA, Saah AJ, Detels R, et al: **Plasma viral load and CD4+ lymphocytes as prognostic markers of HIV-1 infection.** *Ann Intern Med* 1997, **126:**946-954.
51. Ina Y, Gojobori T: **Molecular evolution of human T-cell leukemia virus.** *J Mol Evol* 1990, **31:**493-499.
52. Wolfe ND, Heneine W, Carr JK, Garcia AD, Shanmugam V, Tamoufe U, Torimiro JN, Prosser AT, Lebreton M, Mpoudi-Ngole E, et al: **Emergence of unique primate T-lymphotropic viruses among central African bushmeat hunters.** *Proc Natl Acad Sci U S A* 2005, **102:**7994-7999.
53. Richard L, Mouinga-Ondémé A, Betsem E, Filippone C, Nerrienet E, Kazanji M, Gessain A: **Zoonotic Transmission of Two New Strains of Human T-lymphotropic Virus Type 4 in Hunters Bitten by a Gorilla in Central Africa.** *Clin Infect Dis* 2016, **63:**800-803.
54. Gessain A, Gallo RC, Franchini G: **Low degree of human T-cell leukemia/lymphoma virus type I genetic drift *in vivo* as a means of monitoring viral transmission and movement of ancient human populations.** *J Virol* 1992, **66:**2288-2295.
55. Slattery JP, Franchini G, Gessain A: **Genomic evolution, patterns of global dissemination, and interspecies transmission of human and simian T-cell leukemia/lymphotropic viruses.** *Genome Res* 1999, **9:**525-540.
56. Vidal AU, Gessain A, Yoshida M, Tekaia F, Garin B, Guillemain B, Schulz T, Farid R, De Thé G: **Phylogenetic classification of human T cell leukaemia/lymphoma virus type I genotypes in five major molecular and geographical subtypes.** *J Gen Virol* 1994, **75 (Pt 12):**3655-3666.
57. Hinuma Y, Nagata K, Hanaoka M, Nakai M, Matsumoto T, Kinoshita KI, Shirakawa S, Miyoshi I: **Adult T-cell leukemia: antigen in an ATL cell line and detection of antibodies to the antigen in human sera.** *Proc Natl Acad Sci USA* 1981, **78:**6476-6480.
58. Blattner WA, Gibbs WN, Saxinger C, Robert-Guroff M, Clark J, Lofters W, Hanchard B, Campbell M, Gallo RC: **Human T-cell leukaemia/lymphoma virus-associated lymphoreticular neoplasia in Jamaica.** *Lancet* 1983, **2:**61-64.
59. Biggar RJ, Saxinger C, Gardiner C, Collins WE, Levine PH, Clark JW, Nkrumah FK, Blattner WA: **Type-I HTLV antibody in urban and rural Ghana, West Africa.** *Int J Cancer* 1984, **34:**215-219.
60. Mahieux R, Ibrahim F, Mauclere P, Herve V, Michel P, Tekaia F, Chappey C, Garin B, Van Der Ryst E, Guillemain B, et al: **Molecular epidemiology of 58 new African human T-cell leukemia virus type 1 (HTLV-1) strains:**

identification of a new and distinct HTLV-1 molecular subtype in Central Africa and in Pygmies. *J Virol* 1997, **71:**1317-1333.
61. Miura T, Fukunaga T, Igarashi T, Yamashita M, Ido E, Funahashi S, Ishida T, Washio K, Ueda S, Hashimoto K, et al.: **Phylogenetic subtypes of human T-lymphotropic virus type I and their relations to the anthropological background.** *Proc Natl Acad Sci U S A* 1994, **91:**1124-1127.
62. Carneiro-Proietti AB, Ribas JG, Catalan-Soares BC, Martins ML, Brito-Melo GE, Martins-Filho OA, Pinheiro SR, Araújo Ade Q, Galvão-Castro B, de Oliveira MS, et al: **Infection and disease caused by the human T cell lymphotropic viruses type I and II in Brazil.** *Rev Soc Bras Med Trop* 2002, **35:**499-508.
63. Stienlauf S, Yahalom V, Schwartz E, Shinar E, Segal G, Sidi Y: **Epidemiology of human T-cell lymphotropic virus type 1 infection in blood donors, Israel.** *Emerg Infect Dis* 2009, **15:**1116-1118.
64. Safai B, Huang JL, Boeri E, Farid R, Raafat J, Schutzer P, Ahkami R, Franchini G: **Prevalence of HTLV type I infection in Iran: a serological and genetic study.** *AIDS Res Hum Retroviruses* 1996, **12:**1185-1190.
65. Salemi M, Van Dooren S, Audenaert E, Delaporte E, Goubau P, Desmyter J, Vandamme AM: **Two new human T-lymphotropic virus type I phylogenetic subtypes in seroindeterminates, a Mbuti pygmy and a Gabonese, have closest relatives among African STLV-I strains.** *Virology* 1998, **246:**277-287.
66. Bastian I, Gardner J, Webb D, Gardner I: **Isolation of a human T-lymphotropic virus type I strain from Australian aboriginals.** *J Virol* 1993, **67:**843-851.
67. Cassar O, Einsiedel L, Afonso PV, Gessain A: **Human T-cell lymphotropic virus type 1 subtype C molecular variants among indigenous australians: new insights into the molecular epidemiology of HTLV-1 in Australo-Melanesia.** *PLoS Negl Trop Dis* 2013, **7:**e2418.
68. Gessain A, Boeri E, Yanagihara R, Gallo RC, Franchini G: **Complete nucleotide sequence of a highly divergent human T-cell leukemia (lymphotropic) virus type I (HTLV-I) variant from melanesia: genetic and phylogenetic relationship to HTLV-I strains from other geographical regions.** *J Virol* 1993, **67:**1015-1023.
69. Cassar O, Capuano C, Bassot S, Charavay F, Duprez R, Afonso PV, Abel M, Walter H, Mera W, Martin PM, et al: **Human T lymphotropic virus type 1 subtype C melanesian genetic variants of the Vanuatu Archipelago and Solomon Islands share a common ancestor.** *J Infect Dis* 2007, **196:**510-521.
70. Saksena NK, Sherman MP, Yanagihara R, Dube DK, Poiesz BJ: **LTR sequence and phylogenetic analyses of a newly discovered variant of HTLV-I isolated from the Hagahai of Papua New Guinea.** *Virology* 1992, **189:**1-9.
71. Murphy EL, Figueroa JP, Gibbs WN, Brathwaite A, Holding-Cobham M, Waters D, Cranston B, Hanchard B, Blattner WA: **Sexual transmission of human T-lymphotropic virus type I (HTLV-I).** *Ann Intern Med* 1989, **111:**555-560.
72. Stuver SO, Tachibana N, Okayama A, Shioiri S, Tsunetoshi Y, Tsuda K, Mueller NE: **Heterosexual transmission of human T cell leukemia/lymphoma virus type I among married couples in southwestern Japan: an initial report from the Miyazaki Cohort Study.** *J Infect Dis* 1993, **167:**57-65.
73. Roucoux DF, Wang B, Smith D, Nass CC, Smith J, Hutching ST, Newman B, Lee TH, Chafets DM, Murphy EL: **A prospective study of sexual transmission**

of human T lymphotropic virus (HTLV)-I and HTLV-II. *J Infect Dis* 2005, **191:**1490-1497.
74. Kajiyama W, Kashiwagi S, Ikematsu H, Hayashi J, Nomura H, Okochi K: **Intrafamilial transmission of adult T cell leukemia virus.** *J Infect Dis* 1986, **154:**851-857.
75. La Rosa AM, Zunt JR, Peinado J, Lama JR, Ton TG, Suarez L, Pun M, Cabezas C, Sanchez J: **Retroviral infection in Peruvian men who have sex with men.** *Clin Infect Dis* 2009, **49:**112-117.
76. Moriuchi M, Moriuchi H: **Seminal fluid enhances replication of human T-cell leukemia virus type 1: implications for sexual transmission.** *J Virol* 2004, **78:**12709-12711.
77. Paiva A, Casseb J: **Sexual transmission of human T-cell lymphotropic virus type 1.** *Rev Soc Bras Med Trop* 2014, **47:**265-274.
78. Kaplan JE, Khabbaz RF, Murphy EL, Hermansen S, Roberts C, Lal R, Heneine W, Wright D, Matijas L, Thomson R, et al: **Male-to-female transmission of human T-cell lymphotropic virus types I and II: association with viral load. The Retrovirus Epidemiology Donor Study Group.** *J Acquir Immune Defic Syndr Hum Retrovirol* 1996, **12:**193-201.
79. Mayer KH, Venkatesh KK: **Interactions of HIV, other sexually transmitted diseases, and genital tract inflammation facilitating local pathogen transmission and acquisition.** *Am J Reprod Immunol* 2011, **65:**308-316.
80. Kinoshita K, Hino S, Amagaski T, Ikeda S, Yamada Y, Suzuyama J, Momita S, Toriya K, Kamihira S, Ichimaru M: **Demonstration of adult T-cell leukemia virus antigen in milk from three sero-positive mothers.** *Gan* 1984, **75:**103-105.
81. Hino S, Yamaguchi K, Katamine S, Sugiyama H, Amagasaki T, Kinoshita K, Yoshida Y, Doi H, Tsuji Y, Miyamoto T: **Mother-to-child transmission of human T-cell leukemia virus type-I.** *Jpn J Cancer Res* 1985, **76:**474-480.
82. Hino S: **Establishment of the milk-borne transmission as a key factor for the peculiar endemicity of human T-lymphotropic virus type 1 (HTLV-1): the ATL Prevention Program Nagasaki.** *Proc Jpn Acad Ser B Phys Biol Sci* 2011, **87:**152-166.
83. Takahashi K, Takezaki T, Oki T, Kawakami K, Yashiki S, Fujiyoshi T, Usuku K, Mueller N, Osame M, Miyata K, et al.: **Inhibitory effect of maternal antibody on mother-to-child transmission of human T-lymphotropic virus type I. The Mother-to-Child Transmission Study Group.** *Int J Cancer* 1991, **49:**673-677.
84. Ureta-Vidal A, Angelin-Duclos C, Tortevoye P, Murphy E, Lepère JF, Buigues RP, Jolly N, Joubert M, Carles G, Pouliquen JF, et al: **Mother-to-child transmission of human T-cell-leukemia/lymphoma virus type I: implication of high antiviral antibody titer and high proviral load in carrier mothers.** *Int J Cancer* 1999, **82:**832-836.
85. Li HC, Biggar RJ, Miley WJ, Maloney EM, Cranston B, Hanchard B, Hisada M: **Provirus load in breast milk and risk of mother-to-child transmission of human T lymphotropic virus type I.** *J Infect Dis* 2004, **190:**1275-1278.
86. Fujino T, Fujiyoshi T, Yashiki S, Sonoda S, Otsuka H, Nagata Y: **HTLV-I transmission from mother to fetus via placenta.** *The Lancet* 1992, **340:**1157.
87. Fujino T, Nagata Y: **HTLV-I transmission from mother to child.** *J Reprod Immunol* 2000, **47:**197-206.

88. Okochi K, Sato H, Hinuma Y: **A retrospective study on transmission of adult T cell leukemia virus by blood transfusion: seroconversion in recipients.** *Vox Sang* 1984, **46:**245-253.
89. Toro C, Benito R, Aguilera A, Bassani S, Rodríguez C, Calderón E, Caballero E, Alvarez P, García J, Rodríguez-Iglesias M, et al: **Infection with human T lymphotropic virus type I in organ transplant donors and recipients in Spain.** *J Med Virol* 2005, **76:**268-270.
90. Toro C, Rodés B, Poveda E, Soriano V: **Rapid development of subacute myelopathy in three organ transplant recipients after transmission of human T-cell lymphotropic virus type I from a single donor.** *Transplantation* 2003, **75:**102-104.
91. Kikuchi H, Ohtsuka E, Ono K, Nakayama T, Saburi Y, Tezono K, Ogata M, Iwahashi M, Nasu M: **Allogeneic bone marrow transplantation-related transmission of human T lymphotropic virus type I (HTLV-I).** *Bone Marrow Transplant* 2000, **26:**1235-1237.
92. Caswell RJ, Nall P, Boothby M, Taylor GP: **Rapid onset and progression of myelopathy following an STI: a case for screening?** *Sex Transm Infect* 2019, **95:**244-245.
93. Gout O, Baulac M, Gessain A, Semah F, Saal F, Périès J, Cabrol C, Foucault-Fretz C, Laplane D, Sigaux F, et al.: **Rapid development of myelopathy after HTLV-I infection acquired by transfusion during cardiac transplantation.** *N Engl J Med* 1990, **322:**383-388.
94. Emmanouilides CE, Territo M: **HTLV-I-associated myelopathy following allogeneic bone marrow transplantation.** *Bone Marrow Transplant* 1999, **24:**205-206.
95. Khabbaz RF, Onorato IM, Cannon RO, Hartley TM, Roberts B, Hosein B, Kaplan JE: **Seroprevalence of HTLV-1 and HTLV-2 among intravenous drug users and persons in clinics for sexually transmitted diseases.** *N Engl J Med* 1992, **326:**375-380.
96. Feigal E, Murphy E, Vranizan K, Bacchetti P, Chaisson R, Drummond JE, Blattner W, McGrath M, Greenspan J, Moss A: **Human T cell lymphotropic virus types I and II in intravenous drug users in San Francisco: risk factors associated with seropositivity.** *J Infect Dis* 1991, **164:**36-42.
97. Tang AR, Taylor GP, Dhasmana D: **Self-Flagellation as Possible Route of Human T-Cell Lymphotropic Virus Type-1 Transmission.** *Emerg Infect Dis* 2019, **25:**811-813.
98. Styles CE, Hoad VC, Denham-Ricks P, Brown D, Seed CR: **Self-Flagellation as Possible Route of Human T-Cell Lymphotropic Virus Type 1 Transmission.** *Emerg Infect Dis* 2019, **25:**1996-1997.
99. Mahieux R, Gessain A: **Adult T-cell leukemia/lymphoma and HTLV-1.** *Curr Hematol Malig Rep* 2007, **2:**257-264.
100. Shimoyama M: **Diagnostic criteria and classification of clinical subtypes of adult T-cell leukaemia-lymphoma. A report from the Lymphoma Study Group (1984-87).** *Br J Haematol* 1991, **79:**428-437.
101. Katsuya H, Ishitsuka K, Utsunomiya A, Hanada S, Eto T, Moriuchi Y, Saburi Y, Miyahara M, Sueoka E, Uike N, et al: **Treatment and survival among 1594 patients with ATL.** *Blood* 2015, **126:**2570-2577.
102. Bangham CRM, Ratner L: **How does HTLV-1 cause adult T-cell leukaemia/lymphoma (ATL)?** *Curr Opin Virol* 2015, **14:**93-100.

103. Gill PS, Harrington W, Jr., Kaplan MH, Ribeiro RC, Bennett JM, Liebman HA, Bernstein-Singer M, Espina BM, Cabral L, Allen S, et al.: **Treatment of adult T-cell leukemia-lymphoma with a combination of interferon alfa and zidovudine.** *N Engl J Med* 1995, **332:**1744-1748.
104. Hermine O, Bouscary D, Gessain A, Turlure P, Leblond V, Franck N, Buzyn-Veil A, Rio B, Macintyre E, Dreyfus F, et al.: **Brief report: treatment of adult T-cell leukemia-lymphoma with zidovudine and interferon alfa.** *N Engl J Med* 1995, **332:**1749-1751.
105. Bazarbachi A, Plumelle Y, Carlos Ramos J, Tortevoye P, Otrock Z, Taylor G, Gessain A, Harrington W, Panelatti G, Hermine O: **Meta-analysis on the use of zidovudine and interferon-alfa in adult T-cell leukemia/lymphoma showing improved survival in the leukemic subtypes.** *J Clin Oncol* 2010, **28:**4177-4183.
106. Takasaki Y, Iwanaga M, Imaizumi Y, Tawara M, Joh T, Kohno T, Yamada Y, Kamihira S, Ikeda S, Miyazaki Y, et al: **Long-term study of indolent adult T-cell leukemia-lymphoma.** *Blood* 2010, **115:**4337-4343.
107. Hishizawa M, Kanda J, Utsunomiya A, Taniguchi S, Eto T, Moriuchi Y, Tanosaki R, Kawano F, Miyazaki Y, Masuda M, et al: **Transplantation of allogeneic hematopoietic stem cells for adult T-cell leukemia: a nationwide retrospective study.** *Blood* 2010, **116:**1369-1376.
108. Bazarbachi A, Cwynarski K, Boumendil A, Finel H, Fields P, Raj K, Nagler A, Mohty M, Sureda A, Dreger P, Hermine O: **Outcome of patients with HTLV-1-associated adult T-cell leukemia/lymphoma after SCT: a retrospective study by the EBMT LWP.** *Bone Marrow Transplant* 2014, **49:**1266-1268.
109. Iqbal M, Reljic T, Klocksieben F, Sher T, Ayala E, Murthy H, Bazarbachi A, Kumar A, Kharfan-Dabaja MA: **Efficacy of Allogeneic Hematopoietic Cell Transplantation in Human T Cell Lymphotropic Virus Type 1-Associated Adult T Cell Leukemia/Lymphoma: Results of a Systematic Review/Meta-Analysis.** *Biol Blood Marrow Transplant* 2019, **25:**1695-1700.
110. Ishida T, Joh T, Uike N, Yamamoto K, Utsunomiya A, Yoshida S, Saburi Y, Miyamoto T, Takemoto S, Suzushima H, et al: **Defucosylated anti-CCR4 monoclonal antibody (KW-0761) for relapsed adult T-cell leukemia-lymphoma: a multicenter phase II study.** *J Clin Oncol* 2012, **30:**837-842.
111. Ishida T, Jo T, Takemoto S, Suzushima H, Uozumi K, Yamamoto K, Uike N, Saburi Y, Nosaka K, Utsunomiya A, et al: **Dose-intensified chemotherapy alone or in combination with mogamulizumab in newly diagnosed aggressive adult T-cell leukaemia-lymphoma: a randomized phase II study.** *Br J Haematol* 2015, **169:**672-682.
112. Waldmann TA, White JD, Goldman CK, Top L, Grant A, Bamford R, Roessler E, Horak ID, Zaknoen S, Kasten-Sportes C, et al.: **The interleukin-2 receptor: a target for monoclonal antibody treatment of human T-cell lymphotrophic virus I-induced adult T-cell leukemia.** *Blood* 1993, **82:**1701-1712.
113. Berkowitz JL, Janik JE, Stewart DM, Jaffe ES, Stetler-Stevenson M, Shih JH, Fleisher TA, Turner M, Urquhart NE, Wharfe GH, et al: **Safety, efficacy, and pharmacokinetics/pharmacodynamics of daclizumab (anti-CD25) in patients with adult T-cell leukemia/lymphoma.** *Clin Immunol* 2014, **155:**176-187.
114. Waldmann TA, White JD, Carrasquillo JA, Reynolds JC, Paik CH, Gansow OA, Brechbiel MW, Jaffe ES, Fleisher TA, Goldman CK, et al: **Radioimmunotherapy of interleukin-2R alpha-expressing adult T-cell leukemia with Yttrium-90-labeled anti-Tac.** *Blood* 1995, **86:**4063-4075.

115. Moura IC, Lepelletier Y, Arnulf B, England P, Baude C, Beaumont C, Bazarbachi A, Benhamou M, Monteiro RC, Hermine O: **A neutralizing monoclonal antibody (mAb A24) directed against the transferrin receptor induces apoptosis of tumor T lymphocytes from ATL patients.** *Blood* 2004, **103:**1838-1845.
116. El Hajj H, Tsukasaki K, Cheminant M, Bazarbachi A, Watanabe T, Hermine O: **Novel Treatments of Adult T Cell Leukemia Lymphoma.** *Front Microbiol* 2020, **11:**1062.
117. Bangham CR, Araujo A, Yamano Y, Taylor GP: **HTLV-1-associated myelopathy/tropical spastic paraparesis.** *Nat Rev Dis Primers* 2015, **1:**15012.
118. Iwasaki Y: **Pathology of chronic myelopathy associated with HTLV-I infection (HAM/TSP).** *J Neurol Sci* 1990, **96:**103-123.
119. Umehara F, Izumo S, Nakagawa M, Ronquillo AT, Takahashi K, Matsumuro K, Sato E, Osame M: **Immunocytochemical analysis of the cellular infiltrate in the spinal cord lesions in HTLV-I-associated myelopathy.** *J Neuropathol Exp Neurol* 1993, **52:**424-430.
120. Nagai M, Yamano Y, Brennan MB, Mora CA, Jacobson S: **Increased HTLV-I proviral load and preferential expansion of HTLV-I Tax-specific CD8+ T cells in cerebrospinal fluid from patients with HAM/TSP.** *Ann Neurol* 2001, **50:**807-812.
121. Coler-Reilly ALG, Sato T, Matsuzaki T, Nakagawa M, Niino M, Nagai M, Nakamura T, Takenouchi N, Araya N, Yagishita N, et al: **Effectiveness of Daily Prednisolone to Slow Progression of Human T-Lymphotropic Virus Type 1-Associated Myelopathy/Tropical Spastic Paraparesis: A Multicenter Retrospective Cohort Study.** *Neurotherapeutics* 2017, **14:**1084-1094.
122. Yamauchi J, Tanabe K, Sato T, Nakagawa M, Matsuura E, Tsuboi Y, Tamaki K, Sakima H, Ishihara S, Ohta Y, et al: **Efficacy of Corticosteroid Therapy for HTLV-1-Associated Myelopathy: A Randomized Controlled Trial (HAMLET-P).** *Viruses* 2022, **14**.
123. Machuca A, Rodés B, Soriano V: **The effect of antiretroviral therapy on HTLV infection.** *Virus Res* 2001, **78:**93-100.
124. Izumo S, Goto I, Itoyama Y, Okajima T, Watanabe S, Kuroda Y, Araki S, Mori M, Nagataki S, Matsukura S, et al: **Interferon-alpha is effective in HTLV-I-associated myelopathy: a multicenter, randomized, double-blind, controlled trial.** *Neurology* 1996, **46:**1016-1021.
125. Seegulam ME, Ratner L: **Integrase inhibitors effective against human T-cell leukemia virus type 1.** *Antimicrob Agents Chemother* 2011, **55:**2011-2017.
126. Sheremata WA, Benedict D, Squilacote DC, Sazant A, DeFreitas E: **High-dose zidovudine induction in HTLV-I-associated myelopathy: safety and possible efficacy.** *Neurology* 1993, **43:**2125-2129.
127. Dias ARN, Falcão LFM, Falcão ASC, Normando VMF, Quaresma JAS: **Human T Lymphotropic Virus and Pulmonary Diseases.** *Front Microbiol* 2018, **9:**1879.
128. Einsiedel L, Chiong F, Jersmann H, Taylor GP: **Human T-cell leukaemia virus type 1 associated pulmonary disease: clinical and pathological features of an under-recognised complication of HTLV-1 infection.** *Retrovirology* 2021, **18:**1.
129. Okada F, Ando Y, Yoshitake S, Yotsumoto S, Matsumoto S, Wakisaka M, Maeda T, Mori H: **Pulmonary CT findings in 320 carriers of human T-lymphotropic virus type 1.** *Radiology* 2006, **240:**559-564.

130. Yamashiro T, Kamiya H, Miyara T, Gibo S, Ogawa K, Akamine T, Moromizato H, Yara S, Murayama S: **CT scans of the chest in carriers of human T-cell lymphotropic virus type 1: presence of interstitial pneumonia.** *Acad Radiol* 2012, **19:**952-957.
131. Desgranges C, Bechet JM, Couderc LJ, Caubarrere I, Vernant JC: **Detection of HTLV-1 DNA by polymerase chain reaction in alveolar lymphocytes of patients with tropical spastic paraparesis.** *J Infect Dis* 1989, **160:**162-163.
132. Yamazato Y, Miyazato A, Kawakami K, Yara S, Kaneshima H, Saito A: **High expression of p40(tax) and pro-inflammatory cytokines and chemokines in the lungs of human T-lymphotropic virus type 1-related bronchopulmonary disorders.** *Chest* 2003, **124:**2283-2292.
133. Seki M, Higashiyama Y, Kadota J, Mukae H, Yanagihara K, Tomono K, Kohno S: **Elevated levels of soluble adhesion molecules in sera and BAL fluid of individuals infected with human T-cell lymphotropic virus type 1.** *Chest* 2000, **118:**1754-1761.
134. Chiong F, Jersmann H, Wilson K, Einsiedel L: **HTLV-1c associated bronchiolitis in an Aboriginal man from central Australia.** *IDCases* 2020, **19:**e00714.
135. Taniwaki M, Yamasaki M, Matsumoto Y, Matsumoto N, Ohashi N, Hattori N: **Corticosteroid therapy for organizing pneumonia in a human T-cell lymphotropic virus type-1 carrier.** *Pulmonology* 2019, **25:**193-195.
136. Mochizuki M, Watanabe T, Yamaguchi K, Tajima K, Yoshimura K, Nakashima S, Shirao M, Araki S, Miyata N, Mori S, et al.: **Uveitis associated with human T lymphotropic virus type I: seroepidemiologic, clinical, and virologic studies.** *J Infect Dis* 1992, **166:**943-944.
137. Mochizuki M, Watanabe T, Yamaguchi K, Yoshimura K, Nakashima S, Shirao M, Araki S, Takatsuki K, Mori S, Miyata N: **Uveitis associated with human T-cell lymphotropic virus type I.** *Am J Ophthalmol* 1992, **114:**123-129.
138. Kamoi K, Mochizuki M: **HTLV-1 uveitis.** *Front Microbiol* 2012, **3:**270.
139. Forrester JV: **Uveitis: pathogenesis.** *Lancet* 1991, **338:**1498-1501.
140. Yoshimura K, Mochizuki M, Araki S, Miyata N, Yamaguchi K, Tajima K, Watanabe T: **Clinical and immunologic features of human T-cell lymphotropic virus type I uveitis.** *Am J Ophthalmol* 1993, **116:**156-163.
141. Sagawa K, Mochizuki M, Masuoka K, Katagiri K, Katayama T, Maeda T, Tanimoto A, Sugita S, Watanabe T, Itoh K: **Immunopathological mechanisms of human T cell lymphotropic virus type 1 (HTLV-I) uveitis. Detection of HTLV-I-infected T cells in the eye and their constitutive cytokine production.** *J Clin Invest* 1995, **95:**852-858.
142. Ono A, Mochizuki M, Yamaguchi K, Miyata N, Watanabe T: **Immunologic and virologic characterization of the primary infiltrating cells in the aqueous humor of human T-cell leukemia virus type-1 uveitis. Accumulation of the human T-cell leukemia virus type-1-infected cells and constitutive expression of viral and interleukin-6 messenger ribonucleic acids.** *Invest Ophthalmol Vis Sci* 1997, **38:**676-689.
143. Ono A, Ikeda E, Mochizuki M, Matsuoka M, Yamaguchi K, Sawada T, Yamane S, Tokudome S, Watanabe T: **Provirus load in patients with human T-cell leukemia virus type 1 uveitis correlates with precedent Graves' disease and disease activities.** *Jpn J Cancer Res* 1998, **89:**608-614.

144. Ye L, Taylor GP, Rosadas C: **Human T-Cell Lymphotropic Virus Type 1 and Strongyloides stercoralis Co-infection: A Systematic Review and Meta-Analysis.** *Front Med (Lausanne)* 2022, **9:**832430.
145. Vasquez-Rios G, Pineda-Reyes R, Pineda-Reyes J, Marin R, Ruiz EF, Terashima A: **Strongyloides stercoralis hyperinfection syndrome: a deeper understanding of a neglected disease.** *J Parasit Dis* 2019, **43:**167-175.
146. McGill NK, Vyas J, Shimauchi T, Tokura Y, Piguet V: **HTLV-1-associated infective dermatitis: updates on the pathogenesis.** *Exp Dermatol* 2012, **21:**815-821.
147. Turpin J, Yurick D, Khoury G, Pham H, Locarnini S, Melamed A, Witkover A, Wilson K, Purcell D, Bangham CRM, Einsiedel L: **Impact of hepatitis B virus coinfection on human T-lymphotropic virus type 1 clonality in an Indigenous population of Central Australia.** *J Infect Dis* 2019, **219:**562-567.
148. Alves C, Dourado L: **Endocrine and metabolic disorders in HTLV-1 infected patients.** *Braz J Infect Dis* 2010, **14:**613-620.
149. Talukder MRR, Walley R, Pham H, Schinke S, Woodman R, Wilson K, Sajiv C, Einsiedel L: **Higher human T-cell leukaemia virus type 1 (HTLV-1) proviral load is associated with end-stage kidney disease in Indigenous Australians: Results of a case-control study in central Australia.** *J Med Virol* 2019, **91:**1866-1872.
150. Talukder MR, Clauss CS, Cherian S, Woodman R, Einsiedel L: **Risk factors for HTLV-1, acute kidney injury, and urinary tract infection among aboriginal adults with end stage kidney disease in central Australia.** *J Med Virol* 2021, **93:**6362-6370.
151. Ma G, Yasunaga J, Matsuoka M: **Multifaceted functions and roles of HBZ in HTLV-1 pathogenesis.** *Retrovirology* 2016, **13:**16.
152. Katsuya H, Islam S, Tan BJY, Ito J, Miyazato P, Matsuo M, Inada Y, Iwase SC, Uchiyama Y, Hata H, et al: **The Nature of the HTLV-1 Provirus in Naturally Infected Individuals Analyzed by the Viral DNA-Capture-Seq Approach.** *Cell Rep* 2019, **29:**724-735.e724.
153. Tamiya S, Matsuoka M, Etoh K, Watanabe T, Kamihira S, Yamaguchi K, Takatsuki K: **Two types of defective human T-lymphotropic virus type I provirus in adult T-cell leukemia.** *Blood* 1996, **88:**3065-3073.
154. Miyazaki M, Yasunaga J, Taniguchi Y, Tamiya S, Nakahata T, Matsuoka M: **Preferential selection of human T-cell leukemia virus type 1 provirus lacking the 5' long terminal repeat during oncogenesis.** *J Virol* 2007, **81:**5714-5723.
155. Yurick D: **Human T-lymphotrophic virus type-1: molecular characterisation of HTLV-1 subtype-C and the role of proviral load in pathogenesis in a remote Indigenous Australian setting.** University of Melbourne, Department of Microbiology and Immunology; 2017.
156. Moles R, Sarkis S, Galli V, Omsland M, Purcell DFJ, Yurick D, Khoury G, Pise-Masison CA, Franchini G: **p30 protein: a critical regulator of HTLV-1 viral latency and host immunity.** *Retrovirology* 2019, **16:**42.
157. Georgieva ER: **Non-structural proteins from human T-cell leukemia virus type 1 in cellular membranes-mechanisms for viral survivability and proliferation.** *Int J Mol Sci* 2018, **19**.
158. Harrod R: **Silencers of HTLV-1 and HTLV-2: the pX-encoded latency-maintenance factors.** *Retrovirology* 2019, **16:**25.

159. Felber BK, Paskalis H, Kleinman-Ewing C, Wong-Staal F, Pavlakis GN: **The pX protein of HTLV-I is a transcriptional activator of its long terminal repeats.** *Science* 1985, **229:**675-679.
160. Baranger AM, Palmer CR, Hamm MK, Giebler HA, Brauweiler A, Nyborg JK, Schepartz A: **Mechanism of DNA-binding enhancement by the human T-cell leukaemia virus transactivator Tax.** *Nature* 1995, **376:**606-608.
161. Adya N, Giam CZ: **Distinct regions in human T-cell lymphotropic virus type I tax mediate interactions with activator protein CREB and basal transcription factors.** *J Virol* 1995, **69:**1834-1841.
162. Giebler HA, Loring JE, van Orden K, Colgin MA, Garrus JE, Escudero KW, Brauweiler A, Nyborg JK: **Anchoring of CREB binding protein to the human T-cell leukemia virus type 1 promoter: a molecular mechanism of Tax transactivation.** *Mol Cell Biol* 1997, **17:**5156-5164.
163. Bex F, Yin MJ, Burny A, Gaynor RB: **Differential transcriptional activation by human T-cell leukemia virus type 1 Tax mutants is mediated by distinct interactions with CREB binding protein and p300.** *Mol Cell Biol* 1998, **18:**2392-2405.
164. Leung K, Nabel GJ: **HTLV-1 transactivator induces interleukin-2 receptor expression through an NF-kappa B-like factor.** *Nature* 1988, **333:**776-778.
165. Fujii M, Niki T, Mori T, Matsuda T, Matsui M, Nomura N, Seiki M: **HTLV-1 Tax induces expression of various immediate early serum responsive genes.** *Oncogene* 1991, **6:**1023-1029.
166. Azran I, Schavinsky-Khrapunsky Y, Aboud M: **Role of Tax protein in human T-cell leukemia virus type-I leukemogenicity.** *Retrovirology* 2004, **1:**20.
167. Li M, Kesic M, Yin H, Yu L, Green PL: **Kinetic analysis of human T-cell leukemia virus type 1 gene expression in cell culture and infected animals.** *J Virol* 2009, **83:**3788-3797.
168. Rende F, Cavallari I, Corradin A, Silic-Benussi M, Toulza F, Toffolo GM, Tanaka Y, Jacobson S, Taylor GP, D'Agostino DM: **Kinetics and intracellular compartmentalization of HTLV-1 gene expression: nuclear retention of HBZ mRNAs.** *Blood* 2011, **117**.
169. Kannagi M, Harada S, Maruyama I, Inoko H, Igarashi H, Kuwashima G, Sato S, Morita M, Kidokoro M, Sugimoto M, et al.: **Predominant recognition of human T cell leukemia virus type I (HTLV-I) pX gene products by human CD8+ cytotoxic T cells directed against HTLV-I-infected cells.** *Int Immunol* 1991, **3:**761-767.
170. Pique C, Connan F, Levilain JP, Choppin J, Dokhelar MC: **Among all human T-cell leukemia virus type 1 proteins, tax, polymerase, and envelope proteins are predicted as preferential targets for the HLA-A2-restricted cytotoxic T-cell response.** *J Virol* 1996, **70:**4919-4926.
171. Goon PK, Biancardi A, Fast N, Igakura T, Hanon E, Mosley AJ, Asquith B, Gould KG, Marshall S, Taylor GP, Bangham CR: **Human T cell lymphotropic virus (HTLV) type-1-specific CD8+ T cells: frequency and immunodominance hierarchy.** *J Infect Dis* 2004, **189:**2294-2298.
172. Furukawa Y, Kubota R, Tara M, Izumo S, Osame M: **Existence of escape mutant in HTLV-I tax during the development of adult T-cell leukemia.** *Blood* 2001, **97:**987-993.
173. Koiwa T, Hamano-Usami A, Ishida T, Okayama A, Yamaguchi K, Kamihira S, Watanabe T: **5'-long terminal repeat-selective CpG methylation of latent**

human T-cell leukemia virus type 1 provirus *in vitro* and *in vivo*. *J Virol* 2002, **76:**9389-9397.

174. Billman MR, Rueda D, Bangham CRM: **Single-cell heterogeneity and cell-cycle-related viral gene bursts in the human leukaemia virus HTLV-1.** *Wellcome Open Res* 2017, **2:**87.
175. Mahgoub M, Yasunaga JI, Iwami S, Nakaoka S, Koizumi Y, Shimura K, Matsuoka M: **Sporadic on/off switching of HTLV-1 Tax expression is crucial to maintain the whole population of virus-induced leukemic cells.** *Proc Natl Acad Sci U S A* 2018, **115:**E1269-e1278.
176. Kulkarni A, Mateus M, Thinnes CC, McCullagh JS, Schofield CJ, Taylor GP, Bangham CRM: **Glucose Metabolism and Oxygen Availability Govern Reactivation of the Latent Human Retrovirus HTLV-1.** *Cell Chem Biol* 2017, **24:**1377-1387.e1373.
177. Kulkarni A, Taylor GP, Klose RJ, Schofield CJ, Bangham CR: **Histone H2A monoubiquitylation and p38-MAPKs regulate immediate-early gene-like reactivation of latent retrovirus HTLV-1.** *JCI Insight* 2018, **3**.
178. Melamed A, Laydon DJ, Gillet NA, Tanaka Y, Taylor GP, Bangham CR: **Genome-wide determinants of proviral targeting, clonal abundance and expression in natural HTLV-1 infection.** *PLoS Pathog* 2013, **9:**e1003271.
179. Jacobson S, Shida H, McFarlin DE, Fauci AS, Koenig S: **Circulating CD8+ cytotoxic T lymphocytes specific for HTLV-I pX in patients with HTLV-I associated neurological disease.** *Nature* 1990, **348:**245-248.
180. Parker CE, Daenke S, Nightingale S, Bangham CR: **Activated, HTLV-1-specific cytotoxic T-lymphocytes are found in healthy seropositives as well as in patients with tropical spastic paraparesis.** *Virology* 1992, **188:**628-636.
181. Gaudray G, Gachon F, Basbous J, Biard-Piechaczyk M, Devaux C, Mesnard JM: **The complementary strand of the human T-cell leukemia virus type 1 RNA genome encodes a bZIP transcription factor that down-regulates viral transcription.** *J Virol* 2002, **76:**12813-12822.
182. Satou Y, Yasunaga J, Yoshida M, Matsuoka M: **HTLV-I basic leucine zipper factor gene mRNA supports proliferation of adult T cell leukemia cells.** *Proc Natl Acad Sci USA* 2006, **103:**720-725.
183. Saito M, Matsuzaki T, Satou Y, Yasunaga J, Saito K, Arimura K, Matsuoka M, Ohara Y: ***In vivo* expression of the HBZ gene of HTLV-1 correlates with proviral load, inflammatory markers and disease severity in HTLV-1 associated myelopathy/tropical spastic paraparesis (HAM/TSP).** *Retrovirology* 2009, **6:**19.
184. Clerc I, Polakowski N, André-Arpin C, Cook P, Barbeau B, Mesnard J-M, Lemasson I: **An interaction between the human T cell leukemia virus type 1 basic leucine zipper factor (HBZ) and the KIX domain of p300/CBP contributes to the down-regulation of Tax-dependent viral transcription by HBZ.** *J Biol Chem* 2008, **283:**23903-23913.
185. Zhao T, Yasunaga J, Satou Y, Nakao M, Takahashi M, Fujii M, Matsuoka M: **Human T-cell leukemia virus type 1 bZIP factor selectively suppresses the classical pathway of NF-kappaB.** *Blood* 2009, **113:**2755-2764.
186. Zhi H, Yang L, Kuo YL, Ho YK, Shih HM, Giam CZ: **NF-kappaB hyper-activation by HTLV-1 tax induces cellular senescence, but can be alleviated by the viral anti-sense protein HBZ.** *PLoS Pathog* 2011, **7:**e1002025.

187. Rowan AG, Suemori K, Fujiwara H, Yasukawa M, Tanaka Y, Taylor GP, Bangham CR: **Cytotoxic T lymphocyte lysis of HTLV-1 infected cells is limited by weak HBZ protein expression, but non-specifically enhanced on induction of Tax expression.** *Retrovirology* 2014, **11:**116.
188. Philip S, Zahoor MA, Zhi H, Ho YK, Giam CZ: **Regulation of human T-lymphotropic virus type I latency and reactivation by HBZ and Rex.** *PLoS Pathog* 2014, **10:**e1004040.
189. Sugata K, Satou Y, Yasunaga J, Hara H, Ohshima K, Utsunomiya A, Mitsuyama M, Matsuoka M: **HTLV-1 bZIP factor impairs cell-mediated immunity by suppressing production of Th1 cytokines.** *Blood* 2012, **119:**434-444.
190. Zhao T, Satou Y, Sugata K, Miyazato P, Green PL, Imamura T, Matsuoka M: **HTLV-1 bZIP factor enhances TGF-β signaling through p300 coactivator.** *Blood* 2011, **118:**1865-1876.
191. Mitobe Y, Yasunaga J, Furuta R, Matsuoka M: **HTLV-1 bZIP factor RNA and protein impart distinct functions on T-cell proliferation and survival.** *Cancer Res* 2015, **75**.
192. Ye J, Silverman L, Lairmore MD, Green PL: **HTLV-1 Rex is required for viral spread and persistence in vivo but is dispensable for cellular immortalization in vitro.** *Blood* 2003, **102:**3963-3969.
193. D'Agostino DM, Cavallari I, Romanelli MG, Ciminale V: **Post-transcriptional regulation of HTLV gene expression: Rex to the rescue.** *Front Microbiol* 2019, **10:**1958.
194. Toyoshima H, Itoh M, Inoue J, Seiki M, Takaku F, Yoshida M: **Secondary structure of the human T-cell leukemia virus type 1 rex-responsive element is essential for rex regulation of RNA processing and transport of unspliced RNAs.** *J Virol* 1990, **64:**2825-2832.
195. Kim JH, Kaufman PA, Hanly SM, Rimsky LT, Greene WC: **Rex transregulation of human T-cell leukemia virus type II gene expression.** *J Virol* 1991, **65:**405-414.
196. Neville M, Stutz F, Lee L, Davis LI, Rosbash M: **The importin-beta family member Crm1p bridges the interaction between Rev and the nuclear pore complex during nuclear export.** *Curr Biol* 1997, **7:**767-775.
197. Younis I, Green PL: **The human T-cell leukemia virus Rex protein.** *Front Biosci* 2005, **10:**431-445.
198. Nakano K, Ando T, Yamagishi M, Yokoyama K, Ishida T, Ohsugi T, Tanaka Y, Brighty DW, Watanabe T: **Viral interference with host mRNA surveillance, the nonsense-mediated mRNA decay (NMD) pathway, through a new function of HTLV-1 Rex: implications for retroviral replication.** *Microbes Infect* 2013, **15:**491-505.
199. Prochasson L, Jalinot P, Mocquet V: **The Complex Relationship between HTLV-1 and Nonsense-Mediated mRNA Decay (NMD).** *Pathogens* 2020, **9**.
200. Lee PJ, Yang S, Sun Y, Guo JU: **Regulation of nonsense-mediated mRNA decay in neural development and disease.** *J Mol Cell Biol* 2021, **13:**269-281.
201. Koralnik IJ, Fullen J, Franchini G: **The p12I, p13II, and p30II proteins encoded by human T-cell leukemia/lymphotropic virus type I open reading frames I and II are localized in three different cellular compartments.** *J Virol* 1993, **67:**2360-2366.
202. Ding W, Albrecht B, Luo R, Zhang W, Stanley JR, Newbound GC, Lairmore MD: **Endoplasmic reticulum and cis-Golgi localization of human T-**

lymphotropic virus type 1 p12(I): association with calreticulin and calnexin. *J Virol* 2001, **75:**7672-7682.
203. Sarkis S, Galli V, Moles R, Yurick D, Khoury G, Purcell DFJ, Franchini G, Pise-Masison CA: **Role of HTLV-1 orf-I encoded proteins in viral transmission and persistence.** *Retrovirology* 2019, **16:**43.
204. Albrecht B, D'Souza CD, Ding W, Tridandapani S, Coggeshall KM, Lairmore MD: **Activation of nuclear factor of activated T cells by human T-lymphotropic virus type 1 accessory protein p12(I).** *J Virol* 2002, **76:**3493-3501.
205. Ding W, Albrecht B, Kelley RE, Muthusamy N, Kim SJ, Altschuld RA, Lairmore MD: **Human T-cell lymphotropic virus type 1 p12(I) expression increases cytoplasmic calcium to enhance the activation of nuclear factor of activated T cells.** *J Virol* 2002, **76:**10374-10382.
206. Nicot C, Mulloy JC, Ferrari MG, Johnson JM, Fu K, Fukumoto R, Trovato R, Fullen J, Leonard WJ, Franchini G: **HTLV-1 p12(I) protein enhances STAT5 activation and decreases the interleukin-2 requirement for proliferation of primary human peripheral blood mononuclear cells.** *Blood* 2001, **98:**823-829.
207. Johnson JM, Nicot C, Fullen J, Ciminale V, Casareto L, Mulloy JC, Jacobson S, Franchini G: **Free major histocompatibility complex class I heavy chain is preferentially targeted for degradation by human T-cell leukemia/lymphotropic virus type 1 p12(I) protein.** *J Virol* 2001, **75:**6086-6094.
208. Banerjee P, Feuer G, Barker E: **Human T-cell leukemia virus type 1 (HTLV-1) p12I down-modulates ICAM-1 and -2 and reduces adherence of natural killer cells, thereby protecting HTLV-1-infected primary CD4+ T cells from autologous natural killer cell-mediated cytotoxicity despite the reduction of major histocompatibility complex class I molecules on infected cells.** *J Virol* 2007, **81:**9707-9717.
209. Koralnik IJ, Gessain A, Klotman ME, Lo Monico A, Berneman ZN, Franchini G: **Protein isoforms encoded by the pX region of human T-cell leukemia/lymphotropic virus type I.** *Proc Natl Acad Sci USA* 1992, **89:**8813-8817.
210. Fukumoto R, Andresen V, Bialuk I, Cecchinato V, Walser JC, Valeri VW, Nauroth JM, Gessain A, Nicot C, Franchini G: **In vivo genetic mutations define predominant functions of the human T-cell leukemia/lymphoma virus p12I protein.** *Blood* 2009, **113:**3726-3734.
211. Van Prooyen N, Gold H, Andresen V, Schwartz O, Jones K, Ruscetti F, Lockett S, Gudla P, Venzon D, Franchini G: **Human T-cell leukemia virus type 1 p8 protein increases cellular conduits and virus transmission.** *Proc Natl Acad Sci USA* 2010, **107:**20738-20743.
212. Kim SJ, Nair AM, Fernandez S, Mathes L, Lairmore MD: **Enhancement of LFA-1-mediated T cell adhesion by human T lymphotropic virus type 1 p12I1.** *J Immunol* 2006, **176:**5463-5470.
213. Fukumoto R, Dundr M, Nicot C, Adams A, Valeri VW, Samelson LE, Franchini G: **Inhibition of T-cell receptor signal transduction and viral expression by the linker for activation of T cells-interacting p12(I) protein of human T-cell leukemia/lymphoma virus type 1.** *J Virol* 2007, **81:**9088-9099.
214. Pise-Masison CA, de Castro-Amarante MF, Enose-Akahata Y, Buchmann RC, Fenizia C, Washington Parks R, Edwards D, Fiocchi M, Alcantara LC, Jr., Bialuk I, et al: **Co-dependence of HTLV-1 p12 and p8 functions in virus persistence.** *PLoS Pathog* 2014, **10:**e1004454.

215. Igakura T, Stinchcombe JC, Goon PK, Taylor GP, Weber JN, Griffiths GM, Tanaka Y, Osame M, Bangham CR: **Spread of HTLV-I between lymphocytes by virus-induced polarization of the cytoskeleton.** *Science* 2003, **299:**1713-1716.
216. Pais-Correia AM, Sachse M, Guadagnini S, Robbiati V, Lasserre R, Gessain A, Gout O, Alcover A, Thoulouze MI: **Biofilm-like extracellular viral assemblies mediate HTLV-1 cell-to-cell transmission at virological synapses.** *Nat Med* 2010, **16:**83-89.
217. Pinto DO, DeMarino C, Pleet ML, Cowen M, Branscome H, Al Sharif S, Jones J, Dutartre H, Lepene B, Liotta LA, et al: **HTLV-1 extracellular vesicles promote cell-to-cell contact.** *Front Microbiol* 2019, **10:**2147.
218. Yurick D, Khoury G, Clemens B, Loh L, Pham H, Kedzierska K, Einsiedel L, Purcell D: **Multiplex droplet digital PCR assay for quantification of human T-cell leukemia virus type 1 subtype C DNA proviral load and T cells from blood and respiratory exudates sampled in a remote setting.** *J Clin Microbiol* 2019, **57**.
219. Gillet NA, Malani N, Melamed A, Gormley N, Carter R, Bentley D, Berry C, Bushman FD, Taylor GP, Bangham CR: **The host genomic environment of the provirus determines the abundance of HTLV-1-infected T-cell clones.** *Blood* 2011, **117:**3113-3122.
220. Cook LB, Melamed A, Niederer H, Valganon M, Laydon D, Foroni L, Taylor GP, Matsuoka M, Bangham CR: **The role of HTLV-1 clonality, proviral structure, and genomic integration site in adult T-cell leukemia/lymphoma.** *Blood* 2014, **123:**3925-3931.
221. Furuta R, Yasunaga JI, Miura M, Sugata K, Saito A, Akari H, Ueno T, Takenouchi N, Fujisawa JI, Koh KR, et al: **Human T-cell leukemia virus type 1 infects multiple lineage hematopoietic cells in vivo.** *PLoS Pathog* 2017, **13:**e1006722.
222. Satou Y, Miyazato P, Ishihara K, Yaguchi H, Melamed A, Miura M, Fukuda A, Nosaka K, Watanabe T, Rowan AG, et al: **The retrovirus HTLV-1 inserts an ectopic CTCF-binding site into the human genome.** *Proc Natl Acad Sci USA* 2016, **113:**3054-3059.
223. Melamed A, Yaguchi H, Miura M, Witkover A, Fitzgerald TW, Birney E, Bangham CR: **The human leukemia virus HTLV-1 alters the structure and transcription of host chromatin in cis.** *Elife* 2018, **7**.
224. Cook L, Melamed A, Yaguchi H, Bangham CR: **The impact of HTLV-1 on the cellular genome.** *Curr Opin Virol* 2017, **26:**125-131.
225. Jothi R, Cuddapah S, Barski A, Cui K, Zhao K: **Genome-wide identification of in vivo protein-DNA binding sites from ChIP-Seq data.** *Nucleic Acids Res* 2008, **36:**5221-5231.
226. Laydon DJ, Melamed A, Sim A, Gillet NA, Sim K, Darko S, Kroll JS, Douek DC, Price DA, Bangham CR, Asquith B: **Quantification of HTLV-1 clonality and TCR diversity.** *PLoS Comput Biol* 2014, **10:**e1003646.
227. Meekings KN, Leipzig J, Bushman FD, Taylor GP, Bangham CR: **HTLV-1 integration into transcriptionally active genomic regions is associated with proviral expression and with HAM/TSP.** *PLoS Pathog* 2008, **4:**e1000027.
228. Rosewick N, Durkin K, Artesi M, Marcais A, Hahaut V, Griebel P, Arsic N, Avettand-Fenoel V, Burny A, Charlier C, et al: **Cis-perturbation of cancer drivers by the HTLV-1/BLV proviruses is an early determinant of leukemogenesis.** *Nat Commun* 2017, **8:**15264.

229. Miura M, Miyazato P, Satou Y, Tanaka Y, Bangham CRM: **Epigenetic changes around the pX region and spontaneous HTLV-1 transcription are CTCF-independent.** *Wellcome Open Res* 2018, **3:**105.
230. Kataoka K, Nagata Y, Kitanaka A, Shiraishi Y, Shimamura T, Yasunaga J, Totoki Y, Chiba K, Sato-Otsubo A, Nagae G, et al: **Integrated molecular analysis of adult T cell leukemia/lymphoma.** *Nat Genet* 2015, **47:**1304-1315.
231. Asquith B, Zhang Y, Mosley AJ, de Lara CM, Wallace DL, Worth A, Kaftantzi L, Meekings K, Griffin GE, Tanaka Y, et al: **In vivo T lymphocyte dynamics in humans and the impact of human T-lymphotropic virus 1 infection.** *Proc Natl Acad Sci U S A* 2007, **104:**8035-8040.
232. Laydon DJ, Sunkara V, Boelen L, Bangham CRM, Asquith B: **The relative contributions of infectious and mitotic spread to HTLV-1 persistence.** *PLoS Comput Biol* 2020, **16:**e1007470.
233. Martin F, Fedina A, Youshya S, Taylor GP: **A 15-year prospective longitudinal study of disease progression in patients with HTLV-1 associated myelopathy in the UK.** *J Neurol Neurosurg Psychiatry* 2010, **81:**1336-1340.
234. Gabet AS, Mortreux F, Talarmin A, Plumelle Y, Leclercq I, Leroy A, Gessain A, Clity E, Joubert M, Wattel E: **High circulating proviral load with oligoclonal expansion of HTLV-1 bearing T cells in HTLV-1 carriers with strongyloidiasis.** *Oncogene* 2000, **19:**4954-4960.
235. Yasuma K, Yasunaga J, Takemoto K, Sugata K, Mitobe Y, Takenouchi N, Nakagawa M, Suzuki Y, Matsuoka M: **HTLV-1 bZIP factor impairs anti-viral immunity by inducing co-inhibitory molecule, T cell immunoglobulin and ITIM domain (TIGIT).** *PLoS Pathog* 2016, **12:**e1005372.
236. Satou Y, Matsuoka M: **HTLV-1 and the host immune system: how the virus disrupts immune regulation, leading to HTLV-1 associated diseases.** *J Clin Exp Hematop* 2010, **50:**1-8.
237. Quaresma JA, Yoshikawa GT, Koyama RV, Dias GA, Fujihara S, Fuzii HT: **HTLV-1, immune response and autoimmunity.** *Viruses* 2015, **8**.
238. Bangham CR: **HTLV-1 infection: role of CTL efficiency.** *Blood* 2008, **112:**2176-2177.
239. Futsch N, Prates G, Mahieux R, Casseb J, Dutartre H: **Cytokine networks dysregulation during HTLV-1 infection and associated diseases.** *Viruses* 2018, **10**.
240. Mori N, Gill PS, Mougdil T, Murakami S, Eto S, Prager D: **Interleukin-10 gene expression in adult T-cell leukemia.** *Blood* 1996, **88:**1035-1045.
241. Datta A, Sinha-Datta U, Dhillon NK, Buch S, Nicot C: **The HTLV-I p30 interferes with TLR4 signaling and modulates the release of pro- and anti-inflammatory cytokines from human macrophages.** *J Biol Chem* 2006, **281:**23414-23424.
242. Toulza F, Nosaka K, Tanaka Y, Schioppa T, Balkwill F, Taylor GP, Bangham CR: **Human T-lymphotropic virus type 1-induced CC chemokine ligand 22 maintains a high frequency of functional FoxP3+ regulatory T cells.** *J Immunol* 2010, **185:**183-189.
243. Sawada L, Nagano Y, Hasegawa A, Kanai H, Nogami K, Ito S, Sato T, Yamano Y, Tanaka Y, Masuda T, Kannagi M: **IL-10-mediated signals act as a switch for lymphoproliferation in Human T-cell leukemia virus type-1 infection by activating the STAT3 and IRF4 pathways.** *PLoS Pathog* 2017, **13:**e1006597.

244. Kannagi M, Hasegawa A, Nagano Y, Kimpara S, Suehiro Y: **Impact of host immunity on HTLV-1 pathogenesis: potential of Tax-targeted immunotherapy against ATL.** *Retrovirology* 2019, **16:**23.
245. Sonar SA, Shaikh S, Joshi N, Atre AN, Lal G: **IFN-γ promotes transendothelial migration of CD4(+) T cells across the blood-brain barrier.** *Immunol Cell Biol* 2017, **95:**843-853.
246. Jeffery KJ, Usuku K, Hall SE, Matsumoto W, Taylor GP, Procter J, Bunce M, Ogg GS, Welsh KI, Weber JN, et al: **HLA alleles determine human T-lymphotropic virus-I (HTLV-I) proviral load and the risk of HTLV-I-associated myelopathy.** *Proc Natl Acad Sci U S A* 1999, **96:**3848-3853.
247. Catalan-Soares BC, Carneiro-Proietti AB, Da Fonseca FG, Correa-Oliveira R, Peralva-Lima D, Portela R, Ribas JG, Gonçalves DU, Proietti FA: **HLA class I alleles in HTLV-1-associated myelopathy and asymptomatic carriers from the Brazilian cohort GIPH.** *Med Microbiol Immunol* 2009, **198:**1-3.
248. Talledo M, López G, Huyghe JR, Verdonck K, Adaui V, González E, Best I, Clark D, Vanham G, Gotuzzo E, et al: **Evaluation of host genetic and viral factors as surrogate markers for HTLV-1-associated myelopathy/tropical spastic paraparesis in Peruvian HTLV-1-infected patients.** *J Med Virol* 2010, **82:**460-466.
249. Deschamps R, Béra O, Belrose G, Lezin A, Bellance R, Signate A, Cabre P, Smadja D, Césaire R, Olindo S: **Absence of consistent association between human leukocyte antigen-I and -II alleles and human T-lymphotropic virus type 1 (HTLV-1)-associated myelopathy/tropical spastic paraparesis risk in an HTLV-1 French Afro-Caribbean population.** *Int J Infect Dis* 2010, **14:**e986-990.
250. Treviño A, Vicario JL, Lopez M, Parra P, Benito R, Ortiz de Lejarazu R, Ramos JM, Del Romero J, de Mendoza C, Soriano V: **Association between HLA alleles and HAM/TSP in individuals infected with HTLV-1.** *J Neurol* 2013, **260:**2551-2555.
251. Sabouri AH, Saito M, Lloyd AL, Vine AM, Witkover AW, Furukawa Y, Izumo S, Arimura K, Marshall SE, Usuku K, et al: **Polymorphism in the interleukin-10 promoter affects both provirus load and the risk of human T lymphotropic virus type I-associated myelopathy/tropical spastic paraparesis.** *J Infect Dis* 2004, **190:**1279-1285.
252. Vine AM, Witkover AD, Lloyd AL, Jeffery KJ, Siddiqui A, Marshall SE, Bunce M, Eiraku N, Izumo S, Usuku K, et al: **Polygenic control of human T lymphotropic virus type I (HTLV-I) provirus load and the risk of HTLV-I-associated myelopathy/tropical spastic paraparesis.** *J Infect Dis* 2002, **186:**932-939.
253. Gadelha SR, Junior Alcântara LC, Costa GC, Acosta AX, Rios D, Kashima S, Covas DT, Galvão-Castro B: **Correlation between polymorphisms at interleukin-6 but not at interleukin-10 promoter and the risk of human T lymphotropic virus type I-associated myelopathy/tropical spastic paraparesis in Brazilian individuals.** *J Med Virol* 2008, **80:**2141-2146.
254. Einsiedel L, Pham H, Talukder MR, Taylor K, Wilson K, Kaldor J, Gessain A, Woodman R: **Very high prevalence of infection with the human T cell leukaemia virus type 1c in remote Australian Aboriginal communities: Results of a large cross-sectional community survey.** *PLoS Negl Trop Dis* 2021, **15:**e0009915.

255. Einsiedel LJ, Woodman RJ: **Two nations: racial disparities in bloodstream infections recorded at Alice Springs Hospital, central Australia, 2001-2005.** *Med J Aust* 2010, **192:**567-571.
256. Einsiedel L, Pham H, Wilson K, Walley R, Turpin J, Bangham C, Gessain A, Woodman RJ: **Human T-Lymphotropic Virus type 1c subtype proviral loads, chronic lung disease and survival in a prospective cohort of Indigenous Australians.** *PLoS Negl Trop Dis* 2018, **12:**e0006281.
257. Galli V, Nixon CC, Strbo N, Artesi M, de Castro-Amarante MF, McKinnon K, Fujikawa D, Omsland M, Washington-Parks R, Romero L, et al: **Essential role of human T cell leukemia virus type 1 orf-I in lethal proliferation of CD4(+) cells in humanized mice.** *J Virol* 2019, **93**.
258. Bartoe JT, Albrecht B, Collins ND, Robek MD, Ratner L, Green PL, Lairmore MD: **Functional role of pX open reading frame II of human T-lymphotropic virus type 1 in maintenance of viral loads in vivo.** *J Virol* 2000, **74:**1094-1100.
259. Valeri VW, Hryniewicz A, Andresen V, Jones K, Fenizia C, Bialuk I, Chung HK, Fukumoto R, Parks RW, Ferrari MG, et al: **Requirement of the human T-cell leukemia virus p12 and p30 products for infectivity of human dendritic cells and macaques but not rabbits.** *Blood* 2010, **116:**3809-3817.
260. Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, Bright IJ, Lucero MY, Hiddessen AL, Legler TC, et al: **High-throughput droplet digital PCR system for absolute quantitation of DNA copy number.** *Anal Chem* 2011, **83:**8604-8610.
261. Dan JM, Lindestam Arlehamn CS, Weiskopf D, da Silva Antunes R, Havenar-Daughton C, Reiss SM, Brigger M, Bothwell M, Sette A, Crotty S: **A Cytokine-Independent Approach To Identify Antigen-Specific Human Germinal Center T Follicular Helper Cells and Rare Antigen-Specific CD4+ T Cells in Blood.** *J Immunol* 2016, **197:**983-993.
262. Juno JA, Tan H-X, Lee WS, Reynaldi A, Kelly HG, Wragg K, Esterbauer R, Kent HE, Batten CJ, Mordant FL, et al: **Humoral and circulating follicular helper T cell responses in recovered patients with COVID-19.** *Nature Medicine* 2020, **26:**1428-1434.
263. Mann HB, Whitney DR: **On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other.** *The Annals of Mathematical Statistics* 1947, **18:**50-60.
264. Kruskal WH, Wallis WA: **Use of Ranks in One-Criterion Variance Analysis.** *Journal of the American Statistical Association* 1952, **47:**583-621.
265. Fisher RA: **Statistical Methods for Research Workers.** *Journal of the Royal Statistical Society* 1925, **89:**144-145.
266. Benjamini Y, Krieger AM, Yekutieli D: **Adaptive Linear Step-up Procedures That Control the False Discovery Rate.** *Biometrika* 2006, **93:**491-507.
267. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, et al: **Scikit-learn: Machine Learning in Python.** *Journal of Machine Learning Research* 2011**:**2825-2830.
268. Pearson K: **LIII. On lines and planes of closest fit to systems of points in space.** *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 1901, **2:**559-572.
269. Hotelling H: **Analysis of a complex of statistical variables into principal components.** *Journal of educational psychology* 1933, **24:**417.

270. Hiener B, Horsburgh BA, Eden JS, Barton K, Schlub TE, Lee E, von Stockenstrom S, Odevall L, Milush JM, Liegler T, et al: **Identification of Genetically Intact HIV-1 Proviruses in Specific CD4(+) T Cells from Effectively Treated Participants.** *Cell Rep* 2017, **21:**813-822.
271. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG: **Primer3--new capabilities and interfaces.** *Nucleic Acids Res* 2012, **40:**e115.
272. Manzari V, Wong-Staal F, Franchini G, Colombini S, Gelmann EP, Oroszlan S, Staal S, Gallo RC: **Human T-cell leukemia-lymphoma virus (HTLV): cloning of an integrated defective provirus and flanking cellular sequences.** *Proc Natl Acad Sci U S A* 1983, **80:**1574-1578.
273. Konishi H, Kobayashi N, Hatanaka M: **Defective human T-cell leukemia virus in adult T-cell leukemia patients.** *Mol Biol Med* 1984, **2:**273-283.
274. Mikheyev AS, Tin MM: **A first look at the Oxford Nanopore MinION sequencer.** *Mol Ecol Resour* 2014, **14:**1097-1102.
275. Jain M, Olsen HE, Paten B, Akeson M: **The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community.** *Genome Biol* 2016, **17:**239.
276. Frith MC, Mitsuhashi S, Katoh K: **lamassemble: Multiple Alignment and Consensus Sequence of Long Reads.** *Methods Mol Biol* 2021, **2231:**135-145.
277. Perešíni P, Boža V, Brejová B, Vinař T: **Nanopore base calling on the edge.** *Bioinformatics* 2021, **37:**4661-4667.
278. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL: **NCBI BLAST: a better web interface.** *Nucleic Acids Res* 2008, **36:**W5-9.
279. Crooks GE, Hon G, Chandonia JM, Brenner SE: **WebLogo: a sequence logo generator.** *Genome Res* 2004, **14:**1188-1190.
280. Kent WJ: **BLAT--the BLAST-like alignment tool.** *Genome Res* 2002, **12:**656-664.
281. MacLeod MK, Clambey ET, Kappler JW, Marrack P: **CD4 memory T cells: what are they and what can they do?** *Semin Immunol* 2009, **21:**53-61.
282. Gunasinghe SD, Peres NG, Goyette J, Gaus K: **Biomechanics of T Cell Dysfunctions in Chronic Diseases.** *Frontiers in Immunology* 2021, **12**.
283. Walker JA, McKenzie ANJ: **T(H)2 cell development and function.** *Nat Rev Immunol* 2018, **18:**121-133.
284. Damsker JM, Hansen AM, Caspi RR: **Th1 and Th17 cells: adversaries and collaborators.** *Ann N Y Acad Sci* 2010, **1183:**211-221.
285. Shi G, Cox CA, Vistica BP, Tan C, Wawrousek EF, Gery I: **Phenotype switching by inflammation-inducing polarized Th17 cells, but not by Th1 cells.** *J Immunol* 2008, **181:**7205-7213.
286. Morita R, Schmitt N, Bentebibel SE, Ranganathan R, Bourdery L, Zurawski G, Foucat E, Dullaers M, Oh S, Sabzghabaei N, et al: **Human blood CXCR5(+)CD4(+) T cells are counterparts of T follicular cells and contain specific subsets that differentially support antibody secretion.** *Immunity* 2011, **34:**108-121.
287. Crotty S: **T Follicular Helper Cell Biology: A Decade of Discovery and Diseases.** *Immunity* 2019, **50:**1132-1148.
288. Goon PK, Hanon E, Igakura T, Tanaka Y, Weber JN, Taylor GP, Bangham CR: **High frequencies of Th1-type CD4(+) T cells specific to HTLV-1 Env and Tax proteins in patients with HTLV-1-associated myelopathy/tropical spastic paraparesis.** *Blood* 2002, **99:**3335-3341.

289. Goon PK, Igakura T, Hanon E, Mosley AJ, Asquith B, Gould KG, Taylor GP, Weber JN, Bangham CR: **High circulating frequencies of tumor necrosis factor alpha- and interleukin-2-secreting human T-lymphotropic virus type 1 (HTLV-1)-specific CD4+ T cells in patients with HTLV-1-associated neurological disease.** *J Virol* 2003, **77:**9716-9722.
290. Goon PK, Igakura T, Hanon E, Mosley AJ, Barfield A, Barnard AL, Kaftantzi L, Tanaka Y, Taylor GP, Weber JN, Bangham CR: **Human T cell lymphotropic virus type I (HTLV-I)-specific CD4+ T cells: immunodominance hierarchy and preferential infection with HTLV-I.** *J Immunol* 2004, **172:**1735-1743.
291. Kaech SM, Ahmed R: **Memory CD8+ T cell differentiation: initial antigen encounter triggers a developmental program in naïve cells.** *Nat Immunol* 2001, **2:**415-422.
292. Cassioli C, Baldari CT: **The Expanding Arsenal of Cytotoxic T Cells.** *Front Immunol* 2022, **13:**883010.
293. Kubota R, Nagai M, Kawanishi T, Osame M, Jacobson S: **Increased HTLV type 1 tax specific CD8+ cells in HTLV type 1-asociated myelopathy/tropical spastic paraparesis: correlation with HTLV type 1 proviral load.** *AIDS Res Hum Retroviruses* 2000, **16:**1705-1709.
294. Sabouri AH, Usuku K, Hayashi D, Izumo S, Ohara Y, Osame M, Saito M: **Impaired function of human T-lymphotropic virus type 1 (HTLV-1)-specific CD8+ T cells in HTLV-1-associated neurologic disease.** *Blood* 2008, **112:**2411-2420.
295. Ndhlovu LC, Leal FE, Hasenkrug AM, Jha AR, Carvalho KI, Eccles-James IG, Bruno FR, Vieira RG, York VA, Chew GM, et al: **HTLV-1 tax specific CD8+ T cells express low levels of Tim-3 in HTLV-1 infection: implications for progression to neurological complications.** *PLoS Negl Trop Dis* 2011, **5:**e1030.
296. Fischer K, Voelkl S, Heymann J, Przybylski GK, Mondal K, Laumer M, Kunz-Schughart L, Schmidt CA, Andreesen R, Mackensen A: **Isolation and characterization of human antigen-specific TCRαβ+ CD4-CD8- double-negative regulatory T cells.** *Blood* 2005, **105:**2828-2835.
297. Paul S, Shilpi, Lal G: **Role of gamma-delta (γδ) T cells in autoimmunity.** *J Leukoc Biol* 2015, **97:**259-271.
298. Wu Z, Zheng Y, Sheng J, Han Y, Yang Y, Pan H, Yao J: **CD3(+)CD4(-)CD8(-) (Double-Negative) T Cells in Inflammation, Immune Disorders and Cancer.** *Front Immunol* 2022, **13:**816005.
299. Turner MD, Nedjai B, Hurst T, Pennington DJ: **Cytokines and chemokines: At the crossroads of cell signalling and inflammatory disease.** *Biochim Biophys Acta* 2014, **1843:**2563-2582.
300. Olbei M, Hautefort I, Modos D, Treveil A, Poletti M, Gul L, Shannon-Lowe CD, Korcsmaros T: **SARS-CoV-2 Causes a Different Cytokine Response Compared to Other Cytokine Storm-Causing Respiratory Viruses in Severely Ill Patients.** *Front Immunol* 2021, **12:**629193.
301. Saeidi A, Zandi K, Cheok YY, Saeidi H, Wong WF, Lee CYQ, Cheong HC, Yong YK, Larsson M, Shankar EM: **T-Cell Exhaustion in Chronic Infections: Reversing the State of Exhaustion and Reinvigorating Optimal Protective Immune Responses.** *Front Immunol* 2018, **9:**2569.
302. Muniz AL, Rodrigues W, Jr., Santos SB, de Jesus AR, Porto AF, Castro N, Oliveira-Filho J, Almeida JP, Moreno-Carvalho O, Carvalho EM: **Association of**

cytokines, neurological disability, and disease duration in HAM/TSP patients. *Arq Neuropsiquiatr* 2006, **64:**217-221.
303. Starling AL, Martins-Filho OA, Lambertucci JR, Labanca L, de Souza Pereira SR, Teixeira-Carvalho A, Martins ML, Ribas JG, Carneiro-Proietti AB, Gonçalves DU: **Proviral load and the balance of serum cytokines in HTLV-1-asymptomatic infection and in HTLV-1-associated myelopathy/tropical spastic paraparesis (HAM/TSP).** *Acta Trop* 2013, **125:**75-81.
304. Hanon E, Goon P, Taylor GP, Hasegawa H, Tanaka Y, Weber JN, Bangham CR: **High production of interferon gamma but not interleukin-2 by human T-lymphotropic virus type I-infected peripheral blood mononuclear cells.** *Blood* 2001, **98:**721-726.
305. Neco H, Teixeira V, da Trindade ACL, Magalhães PMR, de Lorena VMB, Castellano LRC, de Souza JR, Vasconcelos LR, de Moura P, de Morais CNL: **Mediators Go Together: High Production of CXCL9, CXCL10, IFN-γ, and TNF-α in HTLV-1-Associated Myelopathy/Tropical Spastic Paraparesis.** *AIDS Res Hum Retroviruses* 2017, **33:**1134-1139.
306. Guerreiro JB, Santos SB, Morgan DJ, Porto AF, Muniz AL, Ho JL, Teixeira AL, Jr., Teixeira MM, Carvalho EM: **Levels of serum chemokines discriminate clinical myelopathy associated with human T lymphotropic virus type 1 (HTLV-1)/tropical spastic paraparesis (HAM/TSP) disease from HTLV-1 carrier state.** *Clin Exp Immunol* 2006, **145:**296-301.
307. Sato T, Coler-Reilly A, Utsunomiya A, Araya N, Yagishita N, Ando H, Yamauchi J, Inoue E, Ueno T, Hasegawa Y, et al: **CSF CXCL10, CXCL9, and neopterin as candidate prognostic biomarkers for HTLV-1-associated myelopathy/tropical spastic paraparesis.** *PLoS Negl Trop Dis* 2013, **7:**e2479.
308. Yamamura M, Yamada Y, Momita S, Kamihira S, Tomonaga M: **Circulating interleukin-6 levels are elevated in adult T-cell leukaemia/lymphoma patients and correlate with adverse clinical features and survival.** *Br J Haematol* 1998, **100:**129-134.
309. Niitsu Y, Urushizaki Y, Koshida Y, Terui K, Mahara K, Kohgo Y, Urushizaki I: **Expression of TGF-beta gene in adult T cell leukemia.** *Blood* 1988, **71:**263-266.
310. Tendler CL, Greenberg SJ, Burton JD, Danielpour D, Kim SJ, Blattner WA, Manns A, Waldmann TA: **Cytokine induction in HTLV-I associated myelopathy and adult T-cell leukemia: alternate molecular mechanisms underlying retroviral pathogenesis.** *J Cell Biochem* 1991, **46:**302-311.
311. Wherry EJ, Ahmed R: **Memory CD8 T-cell differentiation during viral infection.** *J Virol* 2004, **78:**5535-5545.
312. Furman D, Campisi J, Verdin E, Carrera-Bastos P, Targ S, Franceschi C, Ferrucci L, Gilroy DW, Fasano A, Miller GW, et al: **Chronic inflammation in the etiology of disease across the life span.** *Nature Medicine* 2019, **25:**1822-1832.
313. Melamed A, Laydon DJ, Al Khatib H, Rowan AG, Taylor GP, Bangham CR: **HTLV-1 drives vigorous clonal expansion of infected CD8(+) T cells in natural infection.** *Retrovirology* 2015, **12:**91.
314. Petitjean G, Chevalier MF, Tibaoui F, Didier C, Manea ME, Liovat AS, Campa P, Müller-Trutwin M, Girard PM, Meyer L, et al: **Level of double negative T cells, which produce TGF-β and IL-10, predicts CD8 T-cell activation in primary HIV-1 infection.** *Aids* 2012, **26:**139-148.
315. Yamano Y, Nagai M, Brennan M, Mora CA, Soldan SS, Tomaru U, Takenouchi N, Izumo S, Osame M, Jacobson S: **Correlation of human T-cell lymphotropic**

virus type 1 (HTLV-1) mRNA with proviral DNA load, virus-specific CD8+ T cells, and disease severity in HTLV-1–associated myelopathy (HAM/TSP). *Blood* 2002, **99:**88-94.
316. Marodon G, Warren D, Filomio MC, Posnett DN: **Productive infection of double-negative T cells with HIV in vivo.** *Proc Natl Acad Sci U S A* 1999, **96:**11958-11963.
317. Idriss HT, Naismith JH: **TNF alpha and the TNF receptor superfamily: structure-function relationship(s).** *Microsc Res Tech* 2000, **50:**184-195.
318. Booth V, Keizer DW, Kamphuis MB, Clark-Lewis I, Sykes BD: **The CXCR3 binding chemokine IP-10/CXCL10: structure and receptor interactions.** *Biochemistry* 2002, **41:**10418-10425.
319. Iyer SS, Cheng G: **Role of interleukin 10 transcriptional regulation in inflammation and autoimmune disease.** *Crit Rev Immunol* 2012, **32:**23-63.
320. Dooms H, Kahn E, Knoechel B, Abbas AK: **IL-2 induces a competitive survival advantage in T lymphocytes.** *J Immunol* 2004, **172:**5973-5979.
321. Taub DD, Anver M, Oppenheim JJ, Longo DL, Murphy WJ: **T lymphocyte recruitment by interleukin-8 (IL-8). IL-8-induced degranulation of neutrophils releases potent chemoattractants for human T lymphocytes both in vitro and in vivo.** *J Clin Invest* 1996, **97:**1931-1941.
322. Velazquez-Salinas L, Verdugo-Rodriguez A, Rodriguez LL, Borca MV: **The Role of Interleukin 6 During Viral Infections.** *Front Microbiol* 2019, **10:**1057.
323. Tanaka T, Narazaki M, Kishimoto T: **IL-6 in inflammation, immunity, and disease.** *Cold Spring Harb Perspect Biol* 2014, **6:**a016295.
324. Ayroldi E, Zollo O, Cannarile L, F DA, Grohmann U, Delfino DV, Riccardi C: **Interleukin-6 (IL-6) prevents activation-induced cell death: IL-2-independent inhibition of Fas/fasL expression and cell death.** *Blood* 1998, **92:**4212-4219.
325. Wu W, Dietze KK, Gibbert K, Lang KS, Trilling M, Yan H, Wu J, Yang D, Lu M, Roggendorf M, et al: **TLR ligand induced IL-6 counter-regulates the anti-viral CD8(+) T cell response during an acute retrovirus infection.** *Sci Rep* 2015, **5:**10501.
326. Longhi MP, Wright K, Lauder SN, Nowell MA, Jones GW, Godkin AJ, Jones SA, Gallimore AM: **Interleukin-6 is crucial for recall of influenza-specific memory CD4 T cells.** *PLoS Pathog* 2008, **4:**e1000006.
327. Rojas JM, Avia M, Martín V, Sevilla N: **IL-10: A Multifunctional Cytokine in Viral Infections.** *J Immunol Res* 2017, **2017:**6104054.
328. Brockman MA, Kwon DS, Tighe DP, Pavlik DF, Rosato PC, Sela J, Porichis F, Le Gall S, Waring MT, Moss K, et al: **IL-10 is up-regulated in multiple cell types during viremic HIV infection and reversibly inhibits virus-specific T cells.** *Blood* 2009, **114:**346-356.
329. Couper KN, Blount DG, Riley EM: **IL-10: the master regulator of immunity to infection.** *J Immunol* 2008, **180:**5771-5777.
330. Diehl S, Rincón M: **The two faces of IL-6 on Th1/Th2 differentiation.** *Mol Immunol* 2002, **39:**531-536.
331. Fiorentino DF, Bond MW, Mosmann TR: **Two types of mouse T helper cell. IV. Th2 clones secrete a factor that inhibits cytokine production by Th1 clones.** *J Exp Med* 1989, **170:**2081-2095.
332. Yamano Y, Cohen CJ, Takenouchi N, Yao K, Tomaru U, Li HC, Reiter Y, Jacobson S: **Increased expression of human T lymphocyte virus type I (HTLV-**

I) Tax11-19 peptide-human histocompatibility leukocyte antigen A*201 complexes on CD4+ CD25+ T Cells detected by peptide-specific, major histocompatibility complex-restricted antibodies in patients with HTLV-I-associated neurologic disease. *J Exp Med* 2004, **199:**1367-1377.

333. Fukudome K, Furuse M, Fukuhara N, Orita S, Imai T, Takagi S, Nagira M, Hinuma Y, Yoshie O: **Strong induction of ICAM-1 in human T cells transformed by human T-cell-leukemia virus type 1 and depression of ICAM-1 or LFA-1 in adult T-cell-leukemia-derived cell lines.** *Int J Cancer* 1992, **52:**418-427.
334. Tanaka Y, Fukudome K, Hayashi M, Takagi S, Yoshie O: **Induction of ICAM-1 and LFA-3 by Tax1 of human T-cell leukemia virus type 1 and mechanism of down-regulation of ICAM-1 or LFA-1 in adult-T-cell-leukemia cell lines.** *Int J Cancer* 1995, **60:**554-561.
335. Sasaki H, Nishikata I, Shiraga T, Akamatsu E, Fukami T, Hidaka T, Kubuki Y, Okayama A, Hamada K, Okabe H, et al: **Overexpression of a cell adhesion molecule, TSLC1, as a possible molecular marker for acute-type adult T-cell leukemia.** *Blood* 2005, **105:**1204-1213.
336. Nakahata S, Saito Y, Marutsuka K, Hidaka T, Maeda K, Hatakeyama K, Shiraga T, Goto A, Takamatsu N, Asada Y, et al: **Clinical significance of CADM1/TSLC1/IgSF4 expression in adult T-cell leukemia/lymphoma.** *Leukemia* 2012, **26:**1238-1246.
337. Manivannan K, Rowan AG, Tanaka Y, Taylor GP, Bangham CR: **CADM1/TSLC1 Identifies HTLV-1-Infected Cells and Determines Their Susceptibility to CTL-Mediated Lysis.** *PLoS Pathog* 2016, **12:**e1005560.
338. Kobayashi S, Nakano K, Watanabe E, Ishigaki T, Ohno N, Yuji K, Oyaizu N, Asanuma S, Yamagishi M, Yamochi T, et al: **CADM1 expression and stepwise downregulation of CD7 are closely associated with clonal expansion of HTLV-I-infected cells in adult T-cell leukemia/lymphoma.** *Clin Cancer Res* 2014, **20:**2851-2861.
339. Makiyama J, Kobayashi S, Watanabe E, Ishigaki T, Kawamata T, Nakashima M, Yamagishi M, Nakano K, Tojo A, Watanabe T, Uchimaru K: **CD4(+) CADM1(+) cell percentage predicts disease progression in HTLV-1 carriers and indolent adult T-cell leukemia/lymphoma.** *Cancer Sci* 2019, **110:**3746-3753.
340. Yoshie O, Fujisawa R, Nakayama T, Harasawa H, Tago H, Izawa D, Hieshima K, Tatsumi Y, Matsushima K, Hasegawa H, et al: **Frequent expression of CCR4 in adult T-cell leukemia and human T-cell leukemia virus type 1-transformed T cells.** *Blood* 2002, **99:**1505-1511.
341. Hieshima K, Nagakubo D, Nakayama T, Shirakawa AK, Jin Z, Yoshie O: **Tax-inducible production of CC chemokine ligand 22 by human T cell leukemia virus type 1 (HTLV-1)-infected T cells promotes preferential transmission of HTLV-1 to CCR4-expressing CD4+ T cells.** *J Immunol* 2008, **180:**931-939.
342. Yamano Y, Araya N, Sato T, Utsunomiya A, Azakami K, Hasegawa D, Izumi T, Fujita H, Aratani S, Yagishita N, et al: **Abnormally high levels of virus-infected IFN-gamma+ CCR4+ CD4+ CD25+ T cells in a retrovirus-associated neuroinflammatory disorder.** *PLoS One* 2009, **4:**e6517.
343. Araya N, Sato T, Ando H, Tomaru U, Yoshida M, Coler-Reilly A, Yagishita N, Yamauchi J, Hasegawa A, Kannagi M, et al: **HTLV-1 induces a Th1-like state in CD4+CCR4+ T cells.** *J Clin Invest* 2014, **124:**3431-3442.

344. Yamamoto K, Utsunomiya A, Tobinai K, Tsukasaki K, Uike N, Uozumi K, Yamaguchi K, Yamada Y, Hanada S, Tamura K, et al: **Phase I study of KW-0761, a defucosylated humanized anti-CCR4 antibody, in relapsed patients with adult T-cell leukemia-lymphoma and peripheral T-cell lymphoma.** *J Clin Oncol* 2010, **28:**1591-1598.
345. Sato T, Coler-Reilly ALG, Yagishita N, Araya N, Inoue E, Furuta R, Watanabe T, Uchimaru K, Matsuoka M, Matsumoto N, et al: **Mogamulizumab (Anti-CCR4) in HTLV-1-Associated Myelopathy.** *N Engl J Med* 2018, **378:**529-538.
346. Deeks SG, Tracy R, Douek DC: **Systemic effects of inflammation on health during chronic HIV infection.** *Immunity* 2013, **39:**633-645.
347. Matsuoka E, Takenouchi N, Hashimoto K, Kashio N, Moritoyo T, Higuchi I, Isashiki Y, Sato E, Osame M, Izumo S: **Perivascular T cells are infected with HTLV-I in the spinal cord lesions with HTLV-I-associated myelopathy/tropical spastic paraparesis: double staining of immunohistochemistry and polymerase chain reaction in situ hybridization.** *Acta Neuropathol* 1998, **96:**340-346.
348. Afonso PV, Ozden S, Cumont MC, Seilhean D, Cartier L, Rezaie P, Mason S, Lambert S, Huerre M, Gessain A, et al: **Alteration of blood-brain barrier integrity by retroviral infection.** *PLoS Pathog* 2008, **4:**e1000205.
349. Nakamura H, Eguchi K, Nakamura T, Mizokami A, Shirabe S, Kawakami A, Matsuoka N, Migita K, Kawabe Y, Nagataki S: **High prevalence of Sjögren's syndrome in patients with HTLV-I associated myelopathy.** *Ann Rheum Dis* 1997, **56:**167-172.
350. Tangy F, Ossondo M, Vernant JC, Smadja D, Blétry O, Baglin AC, Ozden S: **Human T cell leukemia virus type I expression in salivary glands of infected patients.** *J Infect Dis* 1999, **179:**497-502.
351. Setoguchi Y, Takahashi S, Nukiwa T, Kira S: **Detection of human T-cell lymphotropic virus type I-related antibodies in patients with lymphocytic interstitial pneumonia.** *Am Rev Respir Dis* 1991, **144:**1361-1365.
352. Mori S, Mizoguchi A, Kawabata M, Fukunaga H, Usuku K, Maruyama I, Osame M: **Bronchoalveolar lymphocytosis correlates with human T lymphotropic virus type I (HTLV-I) proviral DNA load in HTLV-I carriers.** *Thorax* 2005, **60:**138-143.
353. Ramirez E, Fernandez J, Cartier L, Villota C, Rios M: **Defective human T-cell lymphotropic virus type I (HTLV-I) provirus in seronegative tropical spastic paraparesis/HTLV-I-associated myelopathy (TSP/HAM) patients.** *Virus Res* 2003, **91:**231-239.
354. Kamihira S, Sugahara K, Tsuruda K, Minami S, Uemura A, Akamatsu N, Nagai H, Murata K, Hasegawa H, Hirakata Y, et al: **Proviral status of HTLV-1 integrated into the host genomic DNA of adult T-cell leukemia cells.** *Clin Lab Haematol* 2005, **27:**235-241.
355. Satou Y, Utsunomiya A, Tanabe J, Nakagawa M, Nosaka K, Matsuoka M: **HTLV-1 modulates the frequency and phenotype of FoxP3+CD4+ T cells in virus-infected individuals.** *Retrovirology* 2012, **9:**46.
356. Miyazato P, Matsuoka M: **Human T-cell leukemia virus type 1 and Foxp3 expression: viral strategy in vivo.** *International Immunology* 2014, **26:**419-425.
357. Bangham CR, Toulza F: **Adult T cell leukemia/lymphoma: FoxP3(+) cells and the cell-mediated immune response to HTLV-1.** *Adv Cancer Res* 2011, **111:**163-182.

358. Fontenot JD, Gavin MA, Rudensky AY: **Foxp3 programs the development and function of CD4+CD25+ regulatory T cells.** *Nat Immunol* 2003, **4:**330-336.
359. Hori S, Nomura T, Sakaguchi S: **Control of regulatory T cell development by the transcription factor Foxp3.** *Science* 2003, **299:**1057-1061.
360. Kleinewietfeld M, Starke M, Di Mitri D, Borsellino G, Battistini L, Rötzschke O, Falk K: **CD49d provides access to "untouched" human Foxp3+ Treg free of contaminating effector cells.** *Blood* 2009, **113:**827-836.
361. Berlin C, Berg EL, Briskin MJ, Andrew DP, Kilshaw PJ, Holzmann B, Weissman IL, Hamann A, Butcher EC: **Alpha 4 beta 7 integrin mediates lymphocyte binding to the mucosal vascular addressin MAdCAM-1.** *Cell* 1993, **74:**185-195.
362. Cepek KL, Shaw SK, Parker CM, Russell GJ, Morrow JS, Rimm DL, Brenner MB: **Adhesion between epithelial cells and T lymphocytes mediated by E-cadherin and the alpha E beta 7 integrin.** *Nature* 1994, **372:**190-193.
363. Slight SR, Khader SA: **Chemokines shape the immune responses to tuberculosis.** *Cytokine Growth Factor Rev* 2013, **24:**105-113.
364. Mikhak Z, Strassner JP, Luster AD: **Lung dendritic cells imprint T cell lung homing and promote lung immunity through the chemokine receptor CCR4.** *J Exp Med* 2013, **210:**1855-1869.
365. Bordon Y: **Mucosal immunology: air miles for T cells.** *Nat Rev Immunol* 2013, **13:**705.
366. Barakat TS, Halbritter F, Zhang M, Rendeiro AF, Perenthaler E, Bock C, Chambers I: **Functional Dissection of the Enhancer Repertoire in Human Embryonic Stem Cells.** *Cell Stem Cell* 2018, **23:**276-288.e278.
367. Gomes JAN, da Silva Dias GA, Fujihara S, Yoshikawa GT, Koyama RVL, Sousa RCM, Quaresma JAS, Fuzii HT: **Decrease in naïve T cell production due to HTLV-1-associated myelopathy/tropical spastic paraparesis (HAM/TSP) development.** *Immunobiology* 2021, **226:**152050.
368. van Dijk EL, Jaszczyszyn Y, Naquin D, Thermes C: **The Third Revolution in Sequencing Technology.** *Trends Genet* 2018, **34:**666-681.
369. Morozov VA, Lagaye S, Taylor GP, Matutes E, Weiss RA: **Chimeric matrix proteins encoded by defective proviruses with large internal deletions in human T-cell leukemia virus type 1-infected humans.** *J Virol* 2000, **74:**3933-3940.
370. Sugata K, Yasunaga J, Kinosada H, Mitobe Y, Furuta R, Mahgoub M, Onishi C, Nakashima K, Ohshima K, Matsuoka M: **HTLV-1 Viral Factor HBZ Induces CCR4 to Promote T-cell Migration and Proliferation.** *Cancer Res* 2016, **76:**5068-5079.
371. Matsuo M, Ueno T, Monde K, Sugata K, Tan BJY, Rahman A, Miyazato P, Uchiyama K, Islam S, Katsuya H, et al: **Identification and characterization of a novel enhancer in the HTLV-1 proviral genome.** *Nat Commun* 2022, **13:**2405.
372. Vandermeulen C, O'Grady T, Wayet J, Galvan B, Maseko S, Cherkaoui M, Desbuleux A, Coppin G, Olivet J, Ben Ameur L, et al: **The HTLV-1 viral oncoproteins Tax and HBZ reprogram the cellular mRNA splicing landscape.** *PLoS Pathog* 2021, **17:**e1009919.
373. Paces J, Pavlícek A, Paces V: **HERVd: database of human endogenous retroviruses.** *Nucleic Acids Res* 2002, **30:**205-206.

374. Mager DL, Goodchild NL: **Homologous recombination between the LTRs of a human retrovirus-like element causes a 5-kb deletion in two siblings.** *Am J Hum Genet* 1989, **45:**848-854.
375. Pavelitz T, Rusché L, Matera AG, Scharf JM, Weiner AM: **Concerted evolution of the tandem array encoding primate U2 snRNA occurs in situ, without changing the cytological context of the RNU2 locus.** *Embo j* 1995, **14:**169-177.
376. Löwer R, Löwer J, Kurth R: **The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences.** *Proc Natl Acad Sci U S A* 1996, **93:**5177-5184.
377. Belshaw R, Watson J, Katzourakis A, Howe A, Woolven-Allen J, Burt A, Tristem M: **Rate of recombinational deletion among human endogenous retroviruses.** *J Virol* 2007, **81:**9437-9442.
378. Mortreux F, Leclercq I, Gabet AS, Leroy A, Westhof E, Gessain A, Wain-Hobson S, Wattel E: **Somatic mutation in human T-cell leukemia virus type 1 provirus and flanking cellular sequences during clonal expansion in vivo.** *J Natl Cancer Inst* 2001, **93:**367-377.
379. Aldrup-Macdonald ME, Sullivan BA: **The past, present, and future of human centromere genomics.** *Genes (Basel)* 2014, **5:**33-50.
380. McKinley KL, Cheeseman IM: **The molecular basis for centromere identity and function.** *Nat Rev Mol Cell Biol* 2016, **17:**16-29.
381. Altemose N, Logsdon GA, Bzikadze AV, Sidhwani P, Langley SA, Caldas GV, Hoyt SJ, Uralsky L, Ryabov FD, Shew CJ, et al: **Complete genomic and epigenetic maps of human centromeres.** *Science* 2022, **376:**eabl4178.
382. Hartley G, O'Neill RJ: **Centromere Repeats: Hidden Gems of the Genome.** *Genes (Basel)* 2019, **10**.
383. Zafar F, Okita AK, Onaka AT, Su J, Katahira Y, Nakayama JI, Takahashi TS, Masukata H, Nakagawa T: **Regulation of mitotic recombination between DNA repeats in centromeres.** *Nucleic Acids Res* 2017, **45:**11222-11235.
384. Melamed A, Fitzgerald TW, Wang Y, Ma J, Birney E, Bangham CRM: **Selective clonal persistence of human retroviruses in vivo: Radial chromatin organization, integration site, and host transcription.** *Science Advances* 2022, **8:**eabm6210.
385. Bury L, Moodie B, Ly J, McKay LS, Miga KH, Cheeseman IM: **Alpha-satellite RNA transcripts are repressed by centromere-nucleolus associations.** *Elife* 2020, **9**.
386. Carnell AN, Goodman JI: **The long (LINEs) and the short (SINEs) of it: altered methylation as a precursor to toxicity.** *Toxicol Sci* 2003, **75:**229-235.
387. Dewannieux M, Esnault C, Heidmann T: **LINE-mediated retrotransposition of marked Alu sequences.** *Nat Genet* 2003, **35:**41-48.
388. Deininger P: **Alu elements: know the SINEs.** *Genome Biol* 2011, **12:**236.
389. Dunker W, Zhao Y, Song Y, Karijolich J: **Recognizing the SINEs of Infection: Regulation of Retrotransposon Expression and Modulation of Host Cell Processes.** *Viruses* 2017, **9**.
390. Jones RB, Song H, Xu Y, Garrison KE, Buzdin AA, Anwar N, Hunter DV, Mujib S, Mihajlovic V, Martin E, et al: **LINE-1 retrotransposable element DNA accumulates in HIV-1-infected cells.** *J Virol* 2013, **87:**13307-13320.

391. Hirons A, Khoury G, Purcell DFJ: **Human T-cell lymphotropic virus type-1: a lifelong persistent infection, yet never truly silent.** *Lancet Infect Dis* 2021, **21:**e2-e10.
392. Lee MY, Khoury G, Olshansky M, Sonza S, Carter GP, McMahon J, Stinear TP, Turner SJ, Lewin SR, Purcell DFJ: **Detection of Chimeric Cellular: HIV mRNAs Generated Through Aberrant Splicing in HIV-1 Latently Infected Resting CD4+ T Cells.** *Front Cell Infect Microbiol* 2022, **12:**855290.
393. Bamford RN, Battiata AP, Burton JD, Sharma H, Waldmann TA: **Interleukin (IL) 15/IL-T production by the adult T-cell leukemia cell line HuT-102 is associated with a human T-cell lymphotrophic virus type I region /IL-15 fusion message that lacks many upstream AUGs that normally attenuates IL-15 mRNA translation.** *Proc Natl Acad Sci U S A* 1996, **93:**2897-2902.
394. Chi KD, McPhee RA, Wagner AS, Dietz JJ, Pantazis P, Goustin AS: **Integration of proviral DNA into the PDGF beta-receptor gene in HTLV-I-infected T-cells results in a novel tyrosine kinase product with transforming activity.** *Oncogene* 1997, **15:**1051-1057.
395. Li G, Reinberg D: **Chromatin higher-order structures and gene regulation.** *Curr Opin Genet Dev* 2011, **21:**175-186.
396. Singh R, Bassett E, Chakravarti A, Parthun MR: **Replication-dependent histone isoforms: a new source of complexity in chromatin structure and function.** *Nucleic Acids Res* 2018, **46:**8665-8678.
397. Rattray Alexander MJ, Müller B: **The control of histone gene expression.** *Biochemical Society Transactions* 2012, **40:**880-885.
398. Yang Y, Zhang M, Wang Y: **The roles of histone modifications in tumorigenesis and associated inhibitors in cancer therapy.** *Journal of the National Cancer Center* 2022, **2:**277-290.
399. Amatori S, Tavolaro S, Gambardella S, Fanelli M: **The dark side of histones: genomic organization and role of oncohistones in cancer.** *Clin Epigenetics* 2021, **13:**71.
400. Panfil AR, Al-Saleem JJ, Green PL: **Animal Models Utilized in HTLV-1 Research.** *Virology (Auckl)* 2013, **4:**49-59.
401. Haruyama N, Cho A, Kulkarni AB: **Overview: engineering transgenic constructs and mice.** *Curr Protoc Cell Biol* 2009, **Chapter 19:**Unit 19.10.
402. Nerenberg M, Hinrichs SH, Reynolds RK, Khoury G, Jay G: **The tat gene of human T-lymphotropic virus type 1 induces mesenchymal tumors in transgenic mice.** *Science* 1987, **237:**1324-1329.
403. Grossman WJ, Kimata JT, Wong FH, Zutter M, Ley TJ, Ratner L: **Development of leukemia in mice transgenic for the tax gene of human T-cell leukemia virus type I.** *Proc Natl Acad Sci U S A* 1995, **92:**1057-1061.
404. Hall AP, Irvine J, Blyth K, Cameron ER, Onions DE, Campbell ME: **Tumours derived from HTLV-I tax transgenic mice are characterized by enhanced levels of apoptosis and oncogene expression.** *J Pathol* 1998, **186:**209-214.
405. Watters KM, Dean J, Hasegawa H, Sawa H, Hall W, Sheehy N: **Cytokine and growth factor expression by HTLV-1 Lck-tax transgenic cells in SCID mice.** *AIDS Res Hum Retroviruses* 2010, **26:**593-603.
406. Iwakura Y, Saijo S, Kioka Y, Nakayama-Yamada J, Itagaki K, Tosu M, Asano M, Kanai Y, Kakimoto K: **Autoimmunity induction by human T cell leukemia**

virus type 1 in transgenic mice that develop chronic inflammatory arthropathy resembling rheumatoid arthritis in humans. *J Immunol* 1995, **155:**1588-1598.
407. Rauch D, Gross S, Harding J, Niewiesk S, Lairmore M, Piwnica-Worms D, Ratner L: **Imaging spontaneous tumorigenesis: inflammation precedes development of peripheral NK tumors.** *Blood* 2009, **113:**1493-1500.
408. Rauch D, Gross S, Harding J, Bokhari S, Niewiesk S, Lairmore M, Piwnica-Worms D, Ratner L: **T-cell activation promotes tumorigenesis in inflammation-associated cancer.** *Retrovirology* 2009, **6:**116.
409. Satou Y, Yasunaga J, Zhao T, Yoshida M, Miyazato P, Takai K, Shimizu K, Ohshima K, Green PL, Ohkura N, et al: **HTLV-1 bZIP factor induces T-cell lymphoma and systemic inflammation in vivo.** *PLoS Pathog* 2011, **7:**e1001274.
410. Mitra-Kaushik S, Harding JC, Hess JL, Ratner L: **Effects of the proteasome inhibitor PS-341 on tumor growth in HTLV-1 Tax transgenic mice and Tax tumor transplants.** *Blood* 2004, **104:**802-809.
411. Kawaguchi A, Orba Y, Kimura T, Iha H, Ogata M, Tsuji T, Ainai A, Sata T, Okamoto T, Hall WW, et al: **Inhibition of the SDF-1alpha-CXCR4 axis by the CXCR4 antagonist AMD3100 suppresses the migration of cultured cells from ATL patients and murine lymphoblastoid cells from HTLV-I Tax transgenic mice.** *Blood* 2009, **114:**2961-2968.
412. Ohsugi T, Ishida T, Shimasaki T, Okada S, Umezawa K: **p53 dysfunction precedes the activation of nuclear factor-κB during disease progression in mice expressing Tax, a human T-cell leukemia virus type 1 oncoprotein.** *Carcinogenesis* 2013, **34:**2129-2136.
413. Ikebe E, Kawaguchi A, Tezuka K, Taguchi S, Hirose S, Matsumoto T, Mitsui T, Senba K, Nishizono A, Hori M, et al: **Oral administration of an HSP90 inhibitor, 17-DMAG, intervenes tumor-cell infiltration into multiple organs and improves survival period for ATL model mice.** *Blood Cancer J* 2013, **3:**e132.
414. Kushida S, Maeda N, Fang J, Uchida K, Miwa M: **Establishment of HTLV-1 carrier mice by injection with HTLV-1-producing T cells.** *Leukemia* 1997, **11 Suppl 3:**260-262.
415. Fang J, Kushida S, Feng R, Tanaka M, Kawamura T, Abe H, Maeda N, Onobori M, Hori M, Uchida K, Miwa M: **Transmission of human T-cell leukemia virus type 1 to mice.** *J Virol* 1998, **72:**3952-3957.
416. Feng R, Tanaka M, Abe H, Arashi N, Sun B, Uchida K, Miwa M: **Human T-cell leukemia virus type 1 can infect a wide variety of cells in mice.** *Jpn J Cancer Res* 1999, **90:**48-54.
417. Tanaka M, Sun B, Fang J, Nitta T, Yoshida T, Kohtoh S, Kikukawa H, Hanai S, Uchida K, Miwa M: **Human T-cell leukemia virus type 1 (HTLV-1) infection of mice: proliferation of cell clones with integrated HTLV-1 provirus in lymphoid organs.** *J Virol* 2001, **75:**4420-4423.
418. Delebecque F, Pramberger K, Prévost MC, Brahic M, Tangy F: **A chimeric human T-cell lymphotropic virus type 1 with the envelope glycoprotein of Moloney murine leukemia virus is infectious for murine cells.** *J Virol* 2002, **76:**7883-7889.
419. Delebecque F, Combredet C, Gabet AS, Wattel E, Brahic M, Tangy F: **A chimeric human T cell leukemia virus type I bearing a deltaR Moloney-murine leukemia virus envelope infects mice persistently and induces humoral and cellular immune responses.** *J Infect Dis* 2005, **191:**255-263.

420. Rahman S, Manuel SL, Khan ZK, Wigdahl B, Acheampong E, Tangy F, Jain P: **Depletion of dendritic cells enhances susceptibility to cell-free infection of human T cell leukemia virus type 1 in CD11c-diphtheria toxin receptor transgenic mice.** *J Immunol* 2010, **184:**5553-5561.
421. Rahman S, Khan ZK, Wigdahl B, Jennings SR, Tangy F, Jain P: **Murine FLT3 ligand-derived dendritic cell-mediated early immune responses are critical to controlling cell-free human T cell leukemia virus type 1 infection.** *J Immunol* 2011, **186:**390-402.
422. Chen J, Liao S, Xiao Z, Pan Q, Wang X, Shen K, Wang S, Yang L, Guo F, Liu H-f, Pan Q: **The development and improvement of immunodeficient mice and humanized immune system mouse models.** *Frontiers in Immunology* 2022, **13**.
423. Priestley A, Beamish HJ, Gell D, Amatucci AG, Muhlmann-Diaz MC, Singleton BK, Smith GC, Blunt T, Schalkwyk LC, Bedford JS, et al: **Molecular and biochemical characterisation of DNA-dependent protein kinase-defective rodent mutant irs-20.** *Nucleic Acids Res* 1998, **26:**1965-1973.
424. Ishihara S, Tachibana N, Okayama A, Murai K, Tsuda K, Mueller N: **Successful graft of HTLV-I-transformed human T-cells (MT-2) in severe combined immunodeficiency mice treated with anti-asialo GM-1 antibody.** *Jpn J Cancer Res* 1992, **83:**320-323.
425. Ohsugi T, Ishibashi K, Shingu M, Nomura T: **Engraftment of HTLV-I-transformed human T-cell line into SCID mice with NK cell function.** *J Vet Med Sci* 1994, **56:**601-603.
426. Feuer G, Zack JA, Harrington WJ, Jr., Valderama R, Rosenblatt JD, Wachsman W, Baird SM, Chen IS: **Establishment of human T-cell leukemia virus type I T-cell lymphomas in severe combined immunodeficient mice.** *Blood* 1993, **82:**722-731.
427. Imada K, Takaori-Kondo A, Akagi T, Shimotohno K, Sugamura K, Hattori T, Yamabe H, Okuma M, Uchiyama T: **Tumorigenicity of human T-cell leukemia virus type I-infected cell lines in severe combined immunodeficient mice and characterization of the cells proliferating in vivo.** *Blood* 1995, **86:**2350-2357.
428. Imada K, Takaori-Kondo A, Sawada H, Imura A, Kawamata S, Okuma M, Uchiyama T: **Serial transplantation of adult T cell leukemia cells into severe combined immunodeficient mice.** *Jpn J Cancer Res* 1996, **87:**887-892.
429. Ishitsuka K, Kunami N, Katsuya H, Nogami R, Ishikawa C, Yotsumoto F, Tanji H, Mori N, Takeshita M, Miyamoto S, Tamura K: **Targeting Bcl-2 family proteins in adult T-cell leukemia/lymphoma: in vitro and in vivo effects of the novel Bcl-2 family inhibitor ABT-737.** *Cancer Lett* 2012, **317:**218-225.
430. Makino S, Kunimoto K, Muraoka Y, Mizushima Y, Katagiri K, Tochino Y: **Breeding of a Non-Obese, Diabetic Strain of Mice.** *Experimental Animals* 1980, **29:**1-13.
431. Shultz LD, Schweitzer PA, Christianson SW, Gott B, Schweitzer IB, Tennent B, McKenna S, Mobraaten L, Rajan TV, Greiner DL: **Multiple defects in innate and adaptive immunologic function in NOD/LtSz-scid mice.** *The Journal of Immunology* 1995, **154:**180-191.
432. Tan C, Waldmann TA: **Proteasome inhibitor PS-341, a potential therapeutic agent for adult T-cell leukemia.** *Cancer Res* 2002, **62:**1083-1086.
433. Zhang Z, Zhang M, Ravetch JV, Goldman C, Waldmann TA: **Effective therapy for a murine model of adult T-cell leukemia with the humanized anti-CD2 monoclonal antibody, MEDI-507.** *Blood* 2003, **102:**284-288.

434. Shu ST, Nadella MV, Dirksen WP, Fernandez SA, Thudi NK, Werbeck JL, Lairmore MD, Rosol TJ: **A novel bioluminescent mouse model and effective therapy for adult T-cell leukemia/lymphoma.** *Cancer Res* 2007, **67:**11859-11866.
435. Miyazato P, Yasunaga J, Taniguchi Y, Koyanagi Y, Mitsuya H, Matsuoka M: **De novo human T-cell leukemia virus type 1 infection of human lymphocytes in NOD-SCID, common gamma-chain knockout mice.** *J Virol* 2006, **80:**10683-10691.
436. Cao X, Shores EW, Hu-Li J, Anver MR, Kelsall BL, Russell SM, Drago J, Noguchi M, Grinberg A, Bloom ET, et al.: **Defective lymphoid development in mice lacking expression of the common cytokine receptor gamma chain.** *Immunity* 1995, **2:**223-238.
437. Shultz LD, Lyons BL, Burzenski LM, Gott B, Chen X, Chaleff S, Kotb M, Gillies SD, King M, Mangada J, et al: **Human lymphoid and myeloid cell development in NOD/LtSz-scid IL2R gamma null mice engrafted with mobilized human hemopoietic stem cells.** *J Immunol* 2005, **174:**6477-6489.
438. Ishikawa F, Yasukawa M, Lyons B, Yoshida S, Miyamoto T, Yoshimoto G, Watanabe T, Akashi K, Shultz LD, Harada M: **Development of functional human blood and immune systems in NOD/SCID/IL2 receptor {gamma} chain(null) mice.** *Blood* 2005, **106:**1565-1573.
439. Ito M, Hiramatsu H, Kobayashi K, Suzue K, Kawahata M, Hioki K, Ueyama Y, Koyanagi Y, Sugamura K, Tsuji K, et al: **NOD/SCID/gamma(c)(null) mouse: an excellent recipient mouse model for engraftment of human cells.** *Blood* 2002, **100:**3175-3182.
440. Goldman JP, Blundell MP, Lopes L, Kinnon C, Di Santo JP, Thrasher AJ: **Enhanced human cell engraftment in mice deficient in RAG2 and the common cytokine receptor gamma chain.** *Br J Haematol* 1998, **103:**335-342.
441. Traggiai E, Chicha L, Mazzucchelli L, Bronz L, Piffaretti JC, Lanzavecchia A, Manz MG: **Development of a human adaptive immune system in cord blood cell-transplanted mice.** *Science* 2004, **304:**104-107.
442. Dewan MZ, Terashima K, Taruishi M, Hasegawa H, Ito M, Tanaka Y, Mori N, Sata T, Koyanagi Y, Maeda M, et al: **Rapid tumor formation of human T-cell leukemia virus type 1-infected cell lines in novel NOD-SCID/gammac(null) mice: suppression by an inhibitor against NF-kappaB.** *J Virol* 2003, **77:**5286-5294.
443. Takajo I, Umeki K, Morishita K, Yamamoto I, Kubuki Y, Hatakeyama K, Kataoka H, Okayama A: **Engraftment of peripheral blood mononuclear cells from human T-lymphotropic virus type 1 carriers in NOD/SCID/gammac(null) (NOG) mice.** *Int J Cancer* 2007, **121:**2205-2211.
444. Yamamoto I, Takajo I, Umeki K, Morishita K, Hatakeyama K, Kataoka H, Nomura H, Okayama A: **Multiple integrations of human T-lymphotropic virus type 1 proviruses in the engrafted cells from the asymptomatic carriers in NOD/SCID/gammacnull mice.** *Intervirology* 2010, **53:**229-239.
445. Arnold J, Zimmerman B, Li M, Lairmore MD, Green PL: **Human T-cell leukemia virus type-1 antisense-encoded gene, Hbz, promotes T-lymphocyte proliferation.** *Blood* 2008, **112:**3788-3797.
446. Shultz LD, Ishikawa F, Greiner DL: **Humanized mice in translational biomedical research.** *Nat Rev Immunol* 2007, **7:**118-130.

447. Shultz LD, Brehm MA, Garcia-Martinez JV, Greiner DL: **Humanized mice for immune system investigation: progress, promise and challenges.** *Nat Rev Immunol* 2012, **12:**786-798.
448. Pérès E, Bagdassarian E, This S, Villaudy J, Rigal D, Gazzolo L, Duc Dodon M: **From Immunodeficiency to Humanization: The Contribution of Mouse Models to Explore HTLV-1 Leukemogenesis.** *Viruses* 2015, **7:**6371-6386.
449. Feuer G, Fraser JK, Zack JA, Lee F, Feuer R, Chen IS: **Human T-cell leukemia virus infection of human hematopoietic progenitor cells: maintenance of virus infection during differentiation in vitro and in vivo.** *J Virol* 1996, **70:**4038-4044.
450. Villaudy J, Wencker M, Gadot N, Gillet NA, Scoazec JY, Gazzolo L, Manz MG, Bangham CR, Dodon MD: **HTLV-1 propels thymic human T cell development in "human immune system" Rag2(-)/(-) gamma c(-)/(-) mice.** *PLoS Pathog* 2011, **7:**e1002231.
451. Tezuka K, Xun R, Tei M, Ueno T, Tanaka M, Takenouchi N, Fujisawa J: **An animal model of adult T-cell leukemia: humanized mice with HTLV-1-specific immunity.** *Blood* 2014, **123:**346-355.
452. Springer LE, Patton JB, Zhan T, Rabson AB, Lin HC, Manser T, Lok JB, Hess JA, Abraham D: **Strongyloides stercoralis and HTLV-1 coinfection in CD34+ cord blood stem cell humanized mice: Alteration of cytokine responses and enhancement of larval growth.** *PLoS Negl Trop Dis* 2021, **15:**e0009559.
453. Fulop GM, Phillips RA: **The scid mutation in mice causes a general defect in DNA repair.** *Nature* 1990, **347:**479-482.
454. DiSanto JP, Müller W, Guy-Grand D, Fischer A, Rajewsky K: **Lymphoid development in mice with a targeted deletion of the interleukin 2 receptor gamma chain.** *Proc Natl Acad Sci U S A* 1995, **92:**377-381.
455. Crozat A, Aman P, Mandahl N, Ron D: **Fusion of CHOP to a novel RNA-binding protein in human myxoid liposarcoma.** *Nature* 1993, **363:**640-644.
456. Rabbitts TH, Forster A, Larson R, Nathan P: **Fusion of the dominant negative transcription regulator CHOP with a novel gene FUS by translocation t(12;16) in malignant liposarcoma.** *Nat Genet* 1993, **4:**175-180.
457. Pérez-Losada J, Pintado B, Gutiérrez-Adán A, Flores T, Bañares-González B, del Campo JC, Martín-Martín JF, Battaner E, Sánchez-García I: **The chimeric FUS/TLS-CHOP fusion protein specifically induces liposarcomas in transgenic mice.** *Oncogene* 2000, **19:**2413-2422.
458. Patil N, Ahmed Kabeer Rasheed S, Abba M, Hendrik Leupold J, Schwarzbach M, Allgayer H: **A mechanistic study on the metastasis inducing function of FUS-CHOP fusion protein in liposarcoma.** *Int J Cancer* 2014, **134:**2808-2819.
459. Shah A, Vaidya NK, Bhat HK, Kumar A: **HIV-1 gp120 induces type-1 programmed cell death through ER stress employing IRE1α, JNK and AP-1 pathway.** *Sci Rep* 2016, **6:**18929.
460. Ma R, Yang L, Niu F, Buch S: **HIV Tat-Mediated Induction of Human Brain Microvascular Endothelial Cell Apoptosis Involves Endoplasmic Reticulum Stress and Mitochondrial Dysfunction.** *Mol Neurobiol* 2016, **53:**132-142.
461. Chipitsyna G, Sawaya BE, Khalili K, Amini S: **Cooperativity between Rad51 and C/EBP family transcription factors modulates basal and Tat-induced activation of the HIV-1 LTR in astrocytes.** *J Cell Physiol* 2006, **207:**605-613.
462. Derse D, Crise B, Li Y, Princler G, Lum N, Stewart C, McGrath CF, Hughes SH, Munroe DJ, Wu X: **Human T-cell leukemia virus type 1 integration target**

sites in the human genome: comparison with those of other retroviruses. *J Virol* 2007, **81:**6731-6741.
463. Cook LB, Rowan AG, Melamed A, Taylor GP, Bangham CR: **HTLV-1-infected T cells contain a single integrated provirus in natural infection.** *Blood* 2012, **120:**3488-3490.
464. Niederer HA, Laydon DJ, Melamed A, Elemans M, Asquith B, Matsuoka M, Bangham CR: **HTLV-1 proviral integration sites differ between asymptomatic carriers and patients with HAM/TSP.** *Virol J* 2014, **11:**172.
465. Allen TM, Brehm MA, Bridges S, Ferguson S, Kumar P, Mirochnitchenko O, Palucka K, Pelanda R, Sanders-Beer B, Shultz LD, et al: **Humanized immune system mouse models: progress, challenges and opportunities.** *Nat Immunol* 2019, **20:**770-774.
466. Curran M, Mairesse M, Matas-Céspedes A, Bareham B, Pellegrini G, Liaunardy A, Powell E, Sargeant R, Cuomo E, Stebbings R, et al: **Recent Advancements and Applications of Human Immune System Mice in Preclinical Immuno-Oncology.** *Toxicol Pathol* 2020, **48:**302-316.
467. Zeng Y, Liu B, Rubio MT, Wang X, Ojcius DM, Tang R, Durrbach A, Ru Z, Zhou Y, Lone YC: **Creation of an immunodeficient HLA-transgenic mouse (HUMAMICE) and functional validation of human immunity after transfer of HLA-matched human cells.** *PLoS One* 2017, **12:**e0173754.
468. Zhao Y, Wright J, Begg S, Guthridge S: **Decomposing Indigenous life expectancy gap by risk factors: a life table analysis.** *Population Health Metrics* 2013, **11:**1.
469. Matloubian M, Concepcion RJ, Ahmed R: **CD4+ T cells are required to sustain CD8+ cytotoxic T-cell responses during chronic viral infection.** *J Virol* 1994, **68:**8056-8063.
470. Williams ES, Martins TB, Shah KS, Hill HR, Coiras M, Spivak AM, Planelles V: **Cytokine Deficiencies in Patients with Long-COVID.** *J Clin Cell Immunol* 2022, **13**.
471. Kim YJ, Park SJ, Broxmeyer HE: **Phagocytosis, a potential mechanism for myeloid-derived suppressor cell regulation of CD8+ T cell function mediated through programmed cell death-1 and programmed cell death-1 ligand interaction.** *J Immunol* 2011, **187:**2291-2301.
472. Jin YH, Hou W, Kang HS, Koh CS, Kim BS: **The role of interleukin-6 in the expression of PD-1 and PDL-1 on central nervous system cells following infection with Theiler's murine encephalomyelitis virus.** *J Virol* 2013, **87:**11538-11551.
473. Bradley LM, Dalton DK, Croft M: **A direct role for IFN-gamma in regulation of Th1 cell development.** *J Immunol* 1996, **157:**1350-1358.
474. Tanaka-Nakanishi A, Yasunaga J, Takai K, Matsuoka M: **HTLV-1 bZIP factor suppresses apoptosis by attenuating the function of FoxO3a and altering its localization.** *Cancer Res* 2014, **74:**188-200.
475. Mitra-Kaushik S, Harding J, Hess J, Schreiber R, Ratner L: **Enhanced tumorigenesis in HTLV-1 tax-transgenic mice deficient in interferon-gamma.** *Blood* 2004, **104:**3305-3311.
476. MacNamara A, Rowan A, Hilburn S, Kadolsky U, Fujiwara H, Suemori K, Yasukawa M, Taylor G, Bangham CR, Asquith B: **HLA class I binding of HBZ determines outcome in HTLV-1 infection.** *PLoS Pathog* 2010, **6**.

477. Satake M, Iwanaga M, Sagara Y, Watanabe T, Okuma K, Hamaguchi I: **Incidence of human T-lymphotropic virus 1 infection in adolescent and adult blood donors in Japan: a nationwide retrospective cohort analysis.** *Lancet Infect Dis* 2016, **16:**1246-1254.
478. Nishijima T, Shimada S, Noda H, Miyake K: **Towards the elimination of HTLV-1 infection in Japan.** *Lancet Infect Dis* 2019, **19:**15-16.
479. Rosadas C, Menezes MLB, Galvão-Castro B, Assone T, Miranda AE, Aragón MG, Caterino-de-Araujo A, Taylor GP, Ishak R: **Blocking HTLV-1/2 silent transmission in Brazil: Current public health policies and proposal for additional strategies.** *PLoS Negl Trop Dis* 2021, **15:**e0009717.