Use of Genome Wide Expression Profiles in Analysis of T Cell Dysfunction in Hepatitis C Virus Infection.

Prakash K Gupta MBBS BSc MRCP UK

Green-Templeton College, University of Oxford

Nuffield Department of Clinical Medicine

Thesis submitted to the Examination Schools

for the degree of Doctor of Philosophy

2014

Table of Contents

Figures

Supervisors

Professor Paul Klenerman

Professor of Immunology, University of Oxford

Peter Medawar Building

Nuffield Department of Medicine

Dr W. Nicholas Haining

Assistant Professor, Department of Pediatrics, Harvard Medical School

Dana-Farber Cancer Institute

Associate Member, Broad Institute of Harvard and MIT

Abstract

During the course of infection with chronic pathogens such as Hepatitis C virus (HCV), Hepatitis B virus (HBV) and HIV, virus-specific $CD8^+$ T cells differentiate into heterogeneous dysfunctional subpopulations. Advances in multi-parameter flow cytometry have allowed these subpopulations to be further classified into classes of exhausted T cells, primarily by their expression of multiple inhibitory receptors. However, the exact phenotype of $CD8^+$ T cells during exhaustion is an area of great interest as many inhibitory receptors are also expressed on functional $CD8^+$ T cells. Discovering novel and specific markers of T cell exhaustion is fundamental in developing strategies to restore $CD8^+$ T cell function in chronic viral infections. Recently, genome wide expression profiles have identified broad molecular phenotypes in exhausted T cells that could not have been discovered by flow cytometry alone. I show how similar genomic approaches identify and further characterise the ectonucleotidase CD39 as a novel marker of $CD8^+$ T cell exhaustion in chronic viral infection. I show that CD39 is highly expressed in HCV and HIV-specific $CD8^+$ T cells and that $CD39^+$ $CD8^+$ T cells are enriched with gene signatures of exhaustion. CD39 is highly co-expressed with multiple inhibitory receptors including PD-1, enzymatically active on $CD8^+$ T cells in HCV infection and positively correlated with viral load in both HCV and HIV. I also demonstrate the discovery of a novel $CD39^{High}$ population of cells in the mouse model of chronic Lymphocytic Choriomenigitis Virus (LCMV) infection, which express the highest degrees of PD-1, LAG3 and 2B4 in the $CD39^+$ fraction. Thus, CD39 is a novel and specific marker of severe $CD8^+$ T cell exhaustion in human and animal models of chronic viral infection.

Acknowledgements

This work was performed at the Peter Medawar Building for pathogen research and at the Dana-Farber Cancer Institute, Massachusetts General Hospital, Beth Israel Deaconess hospital, the Ragon Institute and Brigham and Women's Hospital in Boston, USA.

Firstly, I would like to thank my family particularly my wife, Taran, who has been extremely patient and supportive during many late evenings of research work this past year. Without her I could not have achieved this milestone.

To Nick and Paul, I cannot thank you enough. No two other people could have made better supervisors. Our regular meetings drove my work forward and your constant optimism, enthusiasm and support has given me a great professional experience that I will treasure lifelong.

I would like to thank Jernej Godec and Kathleen Yates for constantly teaching me about basic science and new techniques in the laboratory. I certainly could not have achieved any positive results without your guidance!

I would also like to thank Cormac Cosgrove, David Wolski and Emily Adland who were all great people to get to know and work with.

I am extremely grateful to my collaborators Professors Arlene Sharpe, Simon Robson, Wolfgang Junger and Drs. Galit Alter and Georg Lauer for engaging with my work and me so willingly and allowing access to crucial samples, mouse models and advanced equipment to perform the experiments that I present in my thesis.

Thanks to Professor Derek Jewell, who always made time to meet up with me in the UK and USA. You have been an exemplary educational supervisor.

Finally, I would like to thank all others that made my experience during the last few years so enjoyable. To Ray Chung for allowing me to dip into clinics and ward rounds and to Dick Bennet for taking time to chat whilst performing experiments.

Funding Body

The work presented in my thesis was generously funded by the following organisations:

1. The National Institute of Health grants AI082630 awarded to Drs W. N. Haining, G. Lauer, E.J. Wherry and Professor Paul Klenerman.

2. Wellcome Trust Fellowship Award WT096212AIA awarded to Dr P. K. Gupta.

Publications

Papers

Klenerman P, Gupta PK. **Hepatitis C virus: current concepts and future challenges.** *QJM.* 2012 Jan;105(1):29-32.

Posters

Gupta P K, Wolski D, Alexe G, Lauer G, Haining WN. **Identification of novel, biologically distinct T cell responses to chronic viral infection by unsupervised analysis of gene expression profiles of hepatitis C-specific CD8$^+$ T cells.** *J Immunol 2012 188:105.12*

Alexe G, Pablo Tamayo, Jill Mesirov, Gupta P K, Wolski D, Lauer G, Haining WN. **Metagene projection strategies for discovering mechanisms of T cell differentiation in gene expression profiling data**. *J Immunol 2012 188:58.13*

In preparation

Prakash K. Gupta, Jernej Godec, David Wolski, Emily Adland, Kathleen Yates, Cormac Cosgrove, Carola Ledderose, Wolfgang G. Junger, Simon C. Robson, E. John Wherry, Galit Alter, Philip J. R. Goulder, Paul Klenerman, Arlene H. Sharpe, Georg M. Lauer, W. Nicholas Haining. **CD39 expression identifies terminally exhausted CD8$^+$ T cells in chronic viral infection.**

Contributions

Processing of HCV-specific $CD8^+$ T cells for Figure 3.1 was performed by the Lauer lab and help with computational analysis using R for Figure 3.2 was performed by Gabriela Alexa from the Haining lab. All other work was performed solely by myself with contributions from those listed below.

I acknowledge work by those in the form of having to accompany me in biosafety laboratories, supervision, advice and providing information on samples. These contributions were necessary, as I was unable to enter labs alone to do work as a visitor.

Lauer lab, Massachussettes General Hospital, Boston, USA

- Obtaining HCV cohorts of patients and processing of PBMC's.
- Processing of HCV-specific $CD8^+$ T cells for transcriptional profiling.
- Providing clinical data.
- Providing multimers to identify antigen-specific $CD8^+$ T cells.
- Flow cytometry of HCV samples at Massachusetts General Hospital by David Wolski.

Alter Lab, Ragon Institute, Boston, USA

- Flow cytometry of HIV samples at the Ragon Institute by Cormac Cosgrove.

Goulder Lab, NDM, Oxford, UK

- Flow cytometry of HIV chronic samples by Emily Adland.

Sharpe Lab, NRB, Boston, USA

- Aid in mouse work by Jernej Godec
    - LCMV infections
    - Processing of mice and splenocytes
    - General teaching on GSEA and Principal Components Analysis

Abbreviations

A2A	Adenosine 2 A receptor
ACK	Ammonium-Chloride-Potassium
ADP	Adenosine diposhpate
AMP	Adenosine monophosphate
AP1	Activator Protein-1
ATP	Adenosine triphosphate
B2M	Beta-2-microglobulin
BATF	Basic leucine zipper transcription factor
cAMP	cyclic adenosine monophosphate
CD	Cluster Differentiation
CD127	Interleukin-7 receptor-α
CD4	Cluster of Differentiation 4
CD8	Cluster of Differentiation 8
cDNA	Complementary DNA
CMV	Cytomegalo virus
CT	Cycle threshold
CTL	Cytotoxic T Lymphocyte
CTLA-4	Cytotoxic T-lymphocyte-associated antigen 4
DAA	Direct acting antivirals
DC	Dendritic Cell
EBV	Epstein Barr Virus
EMRA	Effector Memory RA
ENTPD1(CD39)	Ectonucleoside triphosphate diphosphohydrolase 1
Eomes	Eomesodermin
FACS	Fluorescence-activated cell sorting
FBS	Fetal Bovine Serum
FDR	False Discovery Rate
FMO	Full Minus One
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GO	Gene Ontology
GSEA	Gene Set Enrichment Analysis
HBV	Hepatitis B Virus
HC	Hierarchical clustering
HCV	Hepatitis C VIRUS
HIV	Human Immun Deficiency virus
HLA	Human Leukocyte Antigen
ICA	Independent Components Analysis
IFC	Integrated Fluidics Chip
IHLs	Intrahepatic lymphocytes
IL-15	Interleukin 15
IL-2	Interleukin-2
IL-7	Interleukin 7
IVDU	Intra-venous Drug User
KEGG	Kyoto Encyclopedia of Genes and Genomes
KLRG-1	Killer-cell lectin like receptor G1
LAG-3	Lymphocyte activation gene-3
LCMV	Lymphocytic Choriomeningitis
MCODE	Molecular complex detection
MFI	Mean fluorescence intensity
MHC	Major Histocompatability Complex
mRNA	Messenger RNA
MSigDB	Molecular Signatures Database
NK	Natural killer
NMF	Non negative Matrix Factorisation
NS	Non Structural protein
p.f.u.	Plaque-forming unit
p.i.	Post infection
P2X	Purinergic 2 X receptor
P2Y	Purinergic 2 Y receptor
PBMC	Peripheral Blood Mononuclear Cells
PBS	Phosphate Buffered Saline
PCA	Principal Components Analysis
PCR	Polymerase chain reaction
PCR	Polymerase chain reaction
PD-1	Programmed Death-1
POM 1	Sodium polyoxotungstate
QC	quality control
qPCR	Quantitative polymerase chain reaction
RBC	Red blood cells
RMA	Robust multichip averaging
RNA	Ribonucleic acid
rpHPLC	Reversed phase high performance liquid chromatography
RPM	Revolutions per minute
RPMI	Roswell Park Memorial Institute Medium
RT	Reverse Transcriptase
SEM	Standard Error Mean
SHP	src homology 2 domain−containing phosphatase
SIV	Simian immunodeficiency virus
SVR	Sustained viral response
TCR	T cell receptor
Tfh	T follicular helper
Th	T helper
TIM-3	T cell immunoglobulin domain and mucin domain-3
Tregs	Regulatory T cells
WT	Wild type
YFV	Yellow Fever Virus

Introduction

Chapter 1

1.1 Hepatitis C infection

Chronic hepatitis C virus (HCV) is a global infection, with an estimated 170 million people affected worldwide and is a leading cause of hepatic failure requiring transplantation in the west (1). It was first proposed to be a distinct infectious organism in 1979 and formally identified as non-A, non-B in 1989 (2, 3). The majority of those infected develop chronic disease but interestingly, clearance of the virus may be seen in up to 30% of patients and is typically associated with a good T cell response over several weeks (4). This makes HCV infection unique among clinically relevant infections in humans and allows the study of immunologic response with hopes of producing a vaccine. However, chronic infection often leads to cirrhosis, end stage liver disease, hepatocellular carcinoma, and death for approximately 350,000 people per year (5).

HCV is a small, positive-sense, single stranded RNA virus with a 9.6kb genome. It circulates as a highly lipidated molecule that closely resembles host lipoproteins (6). After entering the cell, the viral genome is released and translated into a polyprotein of approximately 3000 amino acids. This is then cleaved by viral and cellular proteases into 10 viral proteins. The N-terminal region of the polyprotein contains the structural proteins that are incorporated into the virus particle, which consists of the capsid forming core protein and the glycoproteins E1 and E2. The C-terminus of the polypeptide contains the non-structural (NS) proteins: p7, NS2, NS3, NS4A, NS5A and NS5B. These NS proteins are expressed in virus-infected cells and aid in intracellular HCV replication via RNA synthesis, manipulation of host defense mechanisms and

virus assembly. The pathway of virus entry, assembly and release is an extremely complex process linked with lipoprotein secretion that is far from fully understood but is an area of ongoing research. The development of in vitro cell culture models have greatly aided in our understanding of host-virus interactions. These initially used drug selected sub-genomic replicons without the production of infectious virus (7) that required the use of specific hepatocyte cell lines (Huh-7) which supports HCV entry, replication and assembly. However, a major breakthrough came with the development of strains derived from the Japanese genotype 2 virus (JFH-1), which can complete a full replication cycle and produce infectious virus in vitro (8).

HCV has a large degree of genetic heterogeneity based on its ability to mutate through error-prone polymerase and its very long co-evolutionary history with humans. This is a major challenge for drug and vaccine development. There are 7 known major genotypes and more than 100 subtypes that have been identified. Geographical distribution is of interest as it gives clues to the history and transmission of the virus. For example endemic strains have persisted in specific locations for many centuries such as Genotype 2 in West Africa and Genotype 6 in South East Asia (9). In the last century there have been iatrogenic outbreaks leading to a massive spread of specific subtypes such as Genotype 4A in Egypt (10). There has also been International spread of Genotype 1 within Europe and North America and more recently Genotype 3 in European intravenous drug users (IVDU) (11).

1.2 Treatment of HCV

Traditional therapies for HCV with pegylated interferon (PEF-IFN) and ribavirin (RBV) have been superseded by novel Direct-Acting Antivirals (DAA's) that target several sites of HCV nonstructural proteins. Bocepravir and Telaprevir are the first NS3/4A protease inhibitors and have shown superior efficacy against genotype 1 HCV infection when combined with PEG-IFN/RBV compared with standard treatment alone (12, 13). Faldaprevir and Simeprevir are similar class agents close to approval and more second-generation drugs are in clinical trials currently (14, 15). Daclatasvir was the first approved class of NS5A replication complex inhibitor and Sofosbuvir the first approved NS5B polymerase inhibitor (16, 17). Both Daclatasvir and Sofosbuvir are highly potent and effective in combination therapies. The emergence of these new HCV-specific antiviral agents with even greater rates of sustained viral response (SVR) has suggested that HCV will be eradicated within a short period of time. This view is flawed for a number of reasons. Firstly, current drugs are demonstrating efficacy in specific genotypes and can be prone to resistance from the virus. Secondly, if only a small proportion of people are treated, there would still be a large burden of disease due to the total number infected with HCV worldwide. It is also very unlikely that these costly medications can be afforded by all those infected. Finally, IVDU's represent a large proportion of people who are recurrently infected in many countries. Therefore, the development of a vaccine and immunotherapies that can stimulate cellular immunity by targeting a sufficient range of viral proteins could limit escape mutations and be more broadly

effective. Barnes et al. have shown that recombinant adenoviral vectors expressing NS protein from HCV genotype 1B were able to prime T cell responses against multiple proteins (18). Both HCV-specific $CD4^+$ and $CD8^+$ T cells displayed responses of a quality that could provide protective immunity for up to 1 year. Such strategies demonstrate a more realistic attempt at eradicating HCV than using costly antiviral drugs alone.

1.3 Antigen recognition and liver immune-tolerance

Being able to identify the correct antigen and target cell is enabled through the interactions between CD4, CD8 molecules and T cell receptors (TCR) expressed on T cells as well as Major Histocompatability Complex (MHC) molecules on target cells. Foreign antigen is degraded within cells, their peptide fragments are captured by MHC molecules and this complex is displayed on the cellular surface. There are two main types of MHC molecules, known as MHC class I and MHC class II. These vary slightly in structure but both contain a cleft in their extracellular surface in which peptide can be held during its assembly. This peptide is transported and displayed on the cell surface, where the peptide is presented to T cells.

$CD8^+$ T cells recognise peptides bound to MHC class I molecules and and $CD4^+$ T cells recognise peptides presented by MHC class II. The source of peptide varies between MHC classes. MHC Class I molecules are expressed on most cells of the body and process peptides generated in the cytosol such as from viral proteins. Therefore, MHC class I are important in the defense against

viral infections. MHC class II molecules are expressed by antigen-presenting cells (APC) such as dendritic cells, macrophages and B cells. They bind peptides from proteins in intracellular vesicles. The T-cell receptor is specific for a combination of peptide and MHC molecule. Any given T cell receptor will recognise either an MHC class I molecule or an MHC class II molecule. T lymphocytes with receptors for recognising MHC class I must also express CD8 co-receptors and T cells recognizing MHC class II must express CD4 co-receptors.

After a T cell recognises antigen it becomes activated, proliferates and differentiates into one of several broad groups of effector T cells. These include cytotoxic T cells, Helper T cells and Regulatory T cells. Cytotoxic T cells kill cells infected with viruses or other intracellular pathogens. Helper T cells provide essential signals that influence the activity of other cells. Regulatory T cells suppress the activity of other lymphocytes and aid to control immune responses. This is mainly done through the secretion of IL-10, as it suppresses the T cell-production of IL-2, TNF- α and inhibits antitigen-presentaing cells by reducing the expression of MHC molecules and co-stimulatory molecules. TGF-β secretion by regulatory T cells can also block T-cell cytokine production, proliferation and cytotoxicity.

The T cell receptor is a highly variable structure owing to its variable amino-terminal region within its α and β chains. Each T cell will mature bearing a unique variant of a prototype antigen receptor. Therefore, hundreds of millions of

lymphocytes circulating the human body at any one time creates a huge repertoire of highly diverse antigen binding sites to recognise foreign pathogens.

Despite the vast array of foreign material it processes, the human liver displays immense tolerogenic properties whilst still allowing innate immunity to eliminate pathogens. It does this by employing the use of parenchymal cells such as hepatocytes and non parenchymal cells such as liver sinusoidal endothelial cells (LSECs), Kupffer cells (KCs), hepatic stellate cells (HSCs) and resident lymphocytes (19). The liver has a large lymphocyte population normally resident in the portal tract but also throughout the parenchyma which consist of T-cells, B-cells, natural killer cells, natural killer T cells and dendritic cells (DCs) (20). The liver is unique in that it can recruit distinct cell types to become antigen presenting cells (APCs) such as LSECs, KCs and hepatic DCs. These cells in turn can prime the resident T cells in the liver. Immunological tolerance may be achieved by liver-resident DCs containing distinct properties that promote tolerance rather than an immune response (21). Another explanation is the innate tolerogenic ability of liver-recruited APCs such as LSECs and KCs. Studies in animal orthotopic liver transplantation models have suggested that donor cell chimerism could also account for this capacity for tolereance (22). In addition, the immunotolerance of the liver has been attributed to the production of alospecific T regulatory cells which are $CD4^{+}CD25^{+}FoxP3^{+}$ (23).

1.4 T cell responses in HCV

There is strong evidence that clearance of HCV infection requires a robust $CD4^+$ and $CD8^+$ T cells response demonstrating adequate number and quality as characterised by proliferation and cytolytic activity over many weeks. In chimpanzee models where there was self-limiting infection, a subsequent re-exposure to the same virus accompanied with antibody mediated depletion of $CD4^+$ or $CD8^+$ T cells resulted in poor viral control (24, 25). In addition the ability to provide a broad T cell response is important since the variation in sequence homology among viruses requires T cells to recognise multiple strains of HCV virus (26).

Also, the role of T lymphocytes in clearance of virus is supported by HLA association with favourable outcome such as in HLA B27, HLA B57, which are interestingly also protective in HIV (27, 28). Both human and chimpanzee models demonstrate that control of HCV viraemia only appears to occur after the development of functional T cell responses (29). Thus the importance of T cell responses in control of HCV appears to be paramount above all other immune defenses.

Other genetic evidence includes data from recent genome-wide association studies (GWAS) that have identified single-nucleotide polymorphisms (SNPs) with single nucleotide base substitutions. This has helped to identify genetic factors that explain disparities in sustained virologic response (SVR) in HCV infected patients that have or have not received Interferon-based therapies (30, 31). SNPs have been discovered near or within the IL-28B gene on

chromosome 19, which code for IFN- λ3 (30, 32). IFN- λ binds to receptors on cell surfaces and activate interferon-signaling genes that contribute to viral suppression. Racial differences have been found such as Caucasians with genotype 1 virus, C/C IL-28B type are associated with SVR rates of upto 70% whereas those with non-C/C, genotype 1 phenotypes have SVR rates between 27% and 33% whereas Africans demonstrating a lower frequency for the C allele (33). These discoveries can aid our prediction of treatment outcomes.

Over the last 10 years we have begun to understand more about the immunological features of HCV infection and drawn more conclusions on the sequence of events that lead to spontaneous clearance or chronic infection. This has come about from increased screening for the virus. On the basis of clinical and immunological features, 4 phases have been postulated after the first year from onset of infection (34) (Figure 1.1).

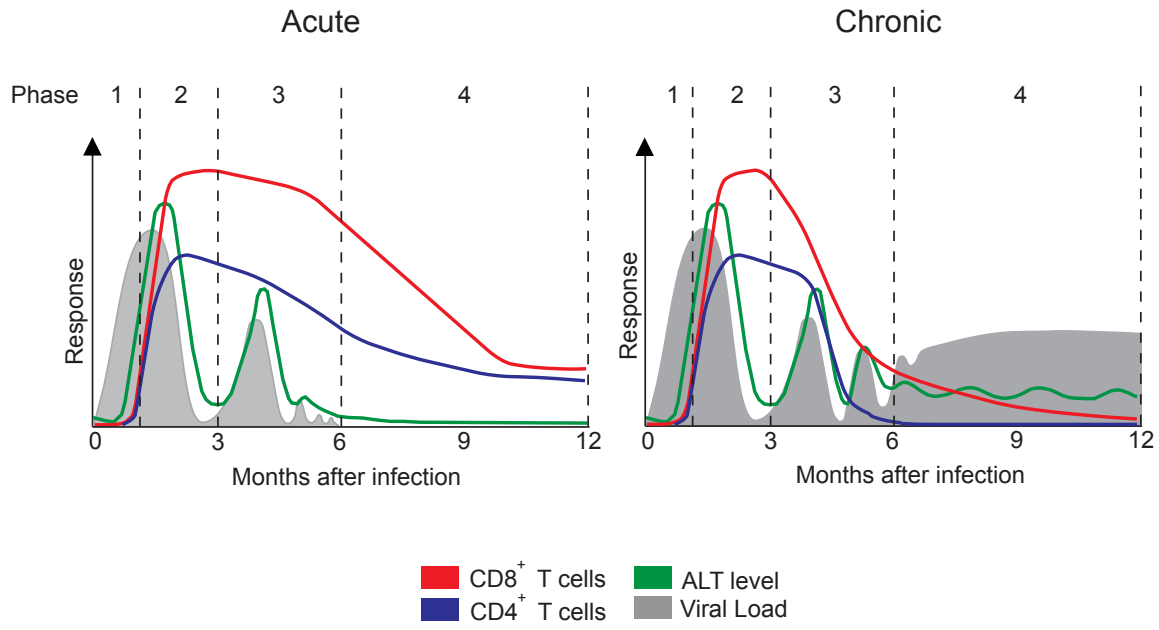


Figure 1.1. Course of T cell response in patients that are infected with Hepatitis C Virus.

The first phase (0 to 6 weeks) begins with infection and high viral titers, but with no initial change in liver function or presence of adaptive immune response by T or B cells (29). Both human studies on hospital workers after needle stick injury and experiments on chimpanzees confirm these findings (35). The large delay in the adaptive immune response is an important scientific question that still needs answering, as it may be during this period of infection that the fate of a successful immunological response has already been decided.

The second phase begins with adaptive immunity and generation of virus-specific antibodies and expansion of crucial HCV-specific $CD8^+$ T cells. Following this is a peak in liver enzymes indicating hepatocellular damage (29). This period involves the majority of people developing some control of viraemia, but it is impossible to elucidate based on viral load, liver function or immunological response, which people are more likely to clear virus in the long-term. Certainly a small proportion will be unable to control virus at all.

The third phase (3 to 6 months post infection) demonstrates wide variations in viral load counts as well as liver enzymes(36, 37). In addition, the variation in $CD4^+$ and $CD8^+$ T cell response becomes much more apparent. In acute infection both subsets remain substantial whereas in chronic infection both $CD4^+$ and CD8's become diminished and lose functional capacity.

The fourth phase (6 to 12 months post infection) sees those infected either clear HCV RNA completely or remain viraemic. Those who have cleared virus display an increased number of more functional $CD8^+$ and $CD4^+$ T cells than those who have persistent infection. In contrast to HIV where T cell responses

grow, in HCV there is a decline during this period in both chronic and resolved infection (38). Following phase 4, $CD8^+$ T cell responses to HCV antigen are less detectable in peripheral blood than in the liver and contribute to immune-mediated pathology (39).

1.5 $CD8^+$ T cell immunity to viral infection

During a prototypical viral infection, circulating virus-specific naive T cells engage with antigen-presenting dendritic cells (DCs) via the formation of immunological synapse (40). $CD8^+$ T cells encountering antigen in the setting of appropriate environmental signals from cytokines and co-stimulatory ligands undergo activation and rapid clonal proliferation. During this time they also differentiate, acquire effector function and the ability to migrate to sites of infection (41). These functional cytotoxic T lymphocytes (CTLs) then clear the pathogen by expressing cytokines and granzymes that destroy pathogen infected cells. Following pathogen clearance, effector $CD8^+$ T cells undergo an equally massive contraction phase, resulting in apoptosis of all but approximately 5-10% of antigen-specific effector $CD8^+$ T cells, which further mature into a long-lived protective memory population (41). The number of these memory $CD8^+$ T cells remains quite stable overtime maintaining the ability to self-renew through homeostatic turnover (42-44). They also exhibit a stem cell like ability to rapidly proliferate and differentiate into secondary effector $CD8^+$ T cells upon re-infection (44). This constitutes the basis of long term T cell mediated immunity to viral

infections. However, it is clear that antigen-specific $CD8^+$ T cells do not all acquire the properties described above (43, 45). As effector $CD8^+$ T cells expand and differentiate during a primary response, most cells terminally differentiate into end-stage effectors that have a shortened life span and die following infection. But a small subset of cells differentiate into memory cell precursors that have increased memory potential (43, 46). Studies have shown that naive $CD8^+$ T cells are not pre-programmed to develop into effector or memory cells and that these decisions occur after T cell activation (47). However, the diversity of the T cell population after antigen exposure is not limited to these 2 groups. The existence of effector $CD8^+$ T cells that exhibit a mixture of terminal and memory properties with varying life spans and proliferative potential suggests an alternative model for differentiation states (48). A more accurate description would be that differentiation of effector $CD8^+$ T cells lies along a spectrum with cells harboring greater longevity and memory potential at one end with end-stage terminally differentiated effector cells at the other and intermediate differentiation states in between (49) (Figure 1.2).

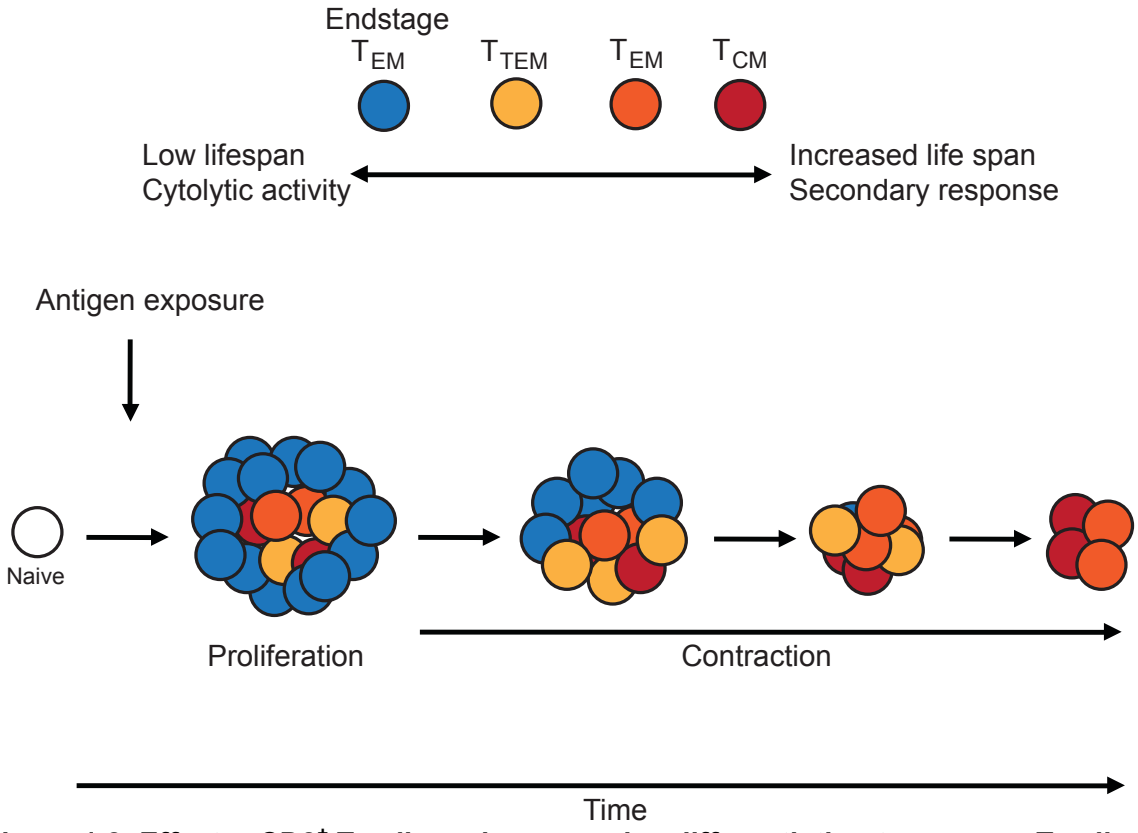


Figure 1.2. Effector $CD8^+$ T cells and progressive differentiation to memory T cells. (End-stage T_{EM}, End-stage effector memory; T_{EM}, effector memory T cells; T_{TEM}, Transitional T_{EM}; T_{CM}, central memory T cells)

This model helps account for the numerous subsets of effector $CD8^+$ T cells that can develop during antigen T cell responses and allows for cells to alter their differentiation state as per signals received. Therefore, a combination of inherent programming together with environmental factors such as antigen load, duration of exposure and location of cell priming results in a diverse range of phenotypes in the memory T cell pool, which has yet to be fully characterised. (49).

Until recently the main tools for characterising T cells have been monoclonal antibodies and flow cytometric analysis. Within the $CD4^+$ memory T cell compartment, heterogeneity is based on selective expression of particular

cytokines and is driven by lineage-specific transcription factors identifying a range of cellular states such as Th1, Th2, Th17 and Tfh cells. Within the memory $CD8^+$ T cell pool classically 2 subsets have been identified, central memory (T_{CM}) and effector memory (T_{EM}) cells, which can be divided by the markers CCR7 and CD62L (50, 51). T_{CM} cells express high levels of CD62L, CCR7 and efficiently home to lymph nodes, whereas T_{EM} cells lack the expression of these molecules and reside mainly in non-lymphoid peripheral tissues. However, these differences in localisation are not absolute, and both subsets can be found to some degree in the different tissues (52). In addition, when circulating memory cells enter specific organs they can upregulate tissue-specific markers that acquire effector function and migratory abilities (53). T_{EM} cells are thought to provide immediate effector function on pathogen entry by expressing perforin and granyme B, which can kill ex vivo, but demonstrate reduced proliferative capacity (52). However, T_{CM} cells undergo homeostatic turnover and can rapidly proliferate with effector response on secondary exposure to the same pathogen (54). Together these two broad phenotypes greatly boost host defenses, but by no means is it well established how distinct these populations actually are. Studies have suggested that T_{CM} and T_{EM} exist as 2 well defined subpopulations from birth and that they do not interchange at all (55). Other reports display contrasting evidence that T_{CM} can convert to T_{EM} upon arrival to non-lymphoid tissues (56). Therefore, In the $CD8^+$ memory compartment considerable heterogeneity exists although the extent to which different phenotypes represent durable, heritable cell fates or more plastic

activation states is unknown (57). Moreover, each compartment can contain antigen-experienced cells of very different functional potential (58).

The development of the first reagents adopting a method by which peptide-Major Histocompatibility Complexes (MHC) could be multimerised with fluorescent labeled streptavidin molecules provided an instrumental tool to allow analysis of the antigen-specific pool regardless of surface phenotype (59, 60) (Figure 1.3). In combination with multi-parameter flow cytometry, this has allowed the phenotype of T cells responding to viruses (61) and tumors (62) to be better mapped.

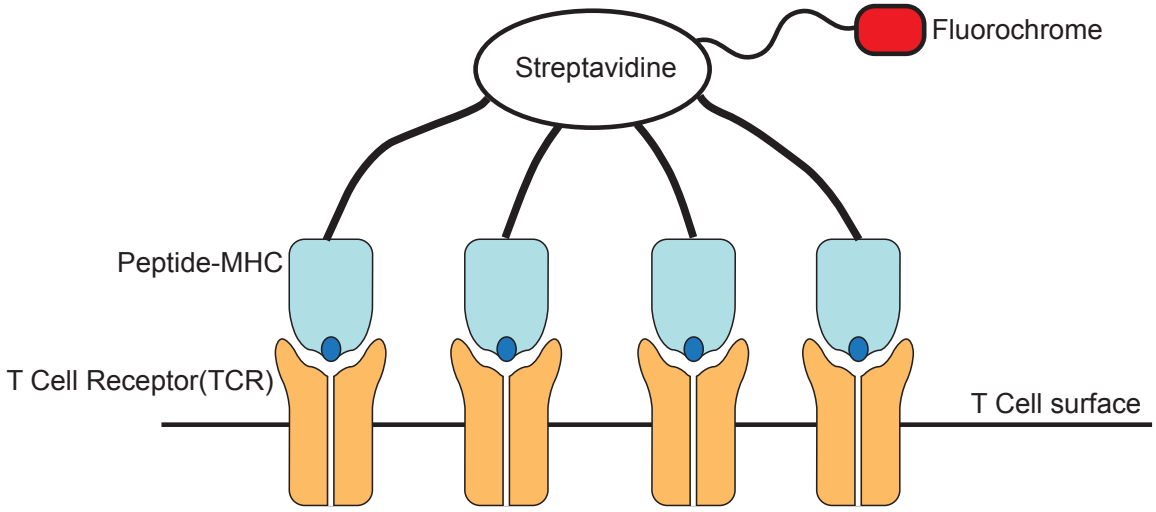


Figure 1.3. Peptide-MHC multimers used to identify antigen-specific T cells.

1.6 T cell exhaustion

As touched upon earlier, in contrast to acute infection where upon exposure to virus, antigen-specific $CD8^+$ T cells acquire effector function and give

rise to memory $CD8^+$ T cells, during chronic infections severe defects in $CD8^+$ T cell responses develop. The term T cell exhaustion was described more than a decade ago as dysfunction and subsequent physical deletion of antigen-specific T cells during chronic viral infection in mice (63, 64). Since then, T cell exhaustion has been demonstrated in a wide variety of animal models and in humans with chronic viral and bacterial infections as well as malignancy (65). Characteristics of exhaustion for pathogens differ in general, but phenotypic and functional properties are broadly similar. It is clear that extrinsic negative regulation by cytokines and intrinsic inhibitory pathways such as the receptor Programmed Death-1 (PD-1) play major roles. In addition, the molecular phenotype of these pathways is becoming clearer with novel tools such as genome-wide transcription profiling and suggests a distinct state of T cell differentiation that can provide targeting opportunities clinically.

During early and intermediate stages of exhaustion, $CD8^+$ T cells typically lose functions such as production of IL-2, TNF α and proliferative capacity (66). Severe exhaustion results in virus-specific cells with impaired IFNγ secretion and subsequently die (66, 67). More severe $CD8^+$ T cell exhaustion correlates with viral load counts (68). In addition, loss of help from $CD4^+$ T cells can lead to more severe exhaustion (55, 69). Importantly, exhausted $CD8^+$ T cells are often characterised by their expression profile of inhibitory receptors such as PD-1, 2B4, LAG-3 or CD160 (70). These markers increase in number and frequency as exhaustion progresses but it is not clear how many actually exist. Therefore, identifying novel markers of exhaustion and elucidating the pattern and frequency

of inhibitory receptors expressed during chronic viral infection is crucial to understanding how T cell function can be restored (71) (Figure 1.4).

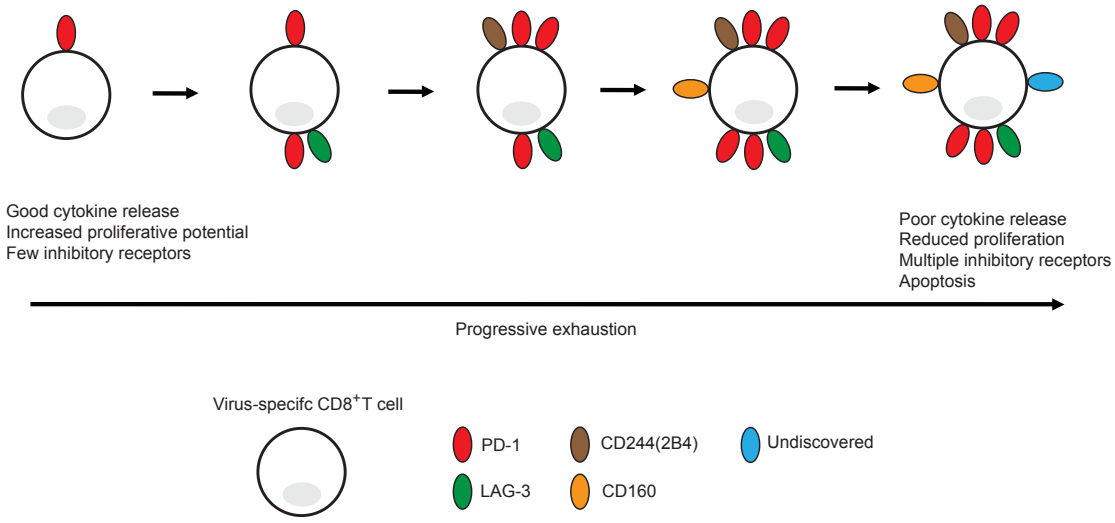


Figure 1.4. Progressive T cell exhaustion in chronic infection

After the initial discovery of T cell exhaustion in mice, similar responses were found in human infections such as HCV, HIV and HBV (55). Similar to mouse studies, hallmark features correlate with exhaustion including high viral load and low $CD4^+$ T cell help. As mentioned earlier the characteristics between infections do differ, for example most HIV-specific $CD8^+$ T cells will produce IFN γ whereas lack of IFN γ production in HCV-specific $CD8^+$ T cells is well described (72). In HCV, peripheral responses deteriorate over time whereas in HIV these remain strong.

A key property of memory $CD8^+$ T cells in acute infection is the ability to persist long term without antigen via IL-7 and IL-15 mediated homeostatic

renewal. During chronic infection in mice, virus-specific $CD8^+$ T cells never develop this property. Exhausted $CD8^+$ T cells have low expression of CD122 (the B-chain of the IL-2 and IL-15 receptor) and CD127 (the IL-7 receptor α chain) and respond poorly to IL-7 and IL-15 and require antigen-specific signals for long term maintenance (73). Transfer of these exhausted $CD8^+$ T cells from mice with chronic infection to mice with no antigen results in very little recovery of normal memory differentiation, suggesting that exhausted $CD8^+$ T cells are committed to exhaustion and clearance of virus does not induce recovery. CD127 is also low in human infections such as HCV and HIV and similar antigen-dependent maintenance is thought to be required for $CD8^+$ T cell survival (74-77). This expression pattern can be useful when assessing HCV-specific $CD8^+$ T cells in chronic HCV infection. This is because HCV-specific T cells identified using tetramers without information from viral sequencing could be demonstrating responses to epitopes where viral escape mutations have already taken place. This is a well identified mechanism of immune evasion by HCV and consequently a reliable phenotype may not be identified by tetramers alone (78, 79).

1.7 T cell exhaustion in HCV

In patients with chronic hepatitis C infection, cytotoxic T lymphocyte responses are crucial to eliminating virus-infected hepatocytes that are functionally impaired in both the liver and peripheral blood (80). Recently, in patients with chronic HCV infection, upregulation of the inhibitory receptors PD-1,

T cell immunoglobulin-and mucin-domain-containing molecule-3 (TIM-3), cytotoxic T lymphocyte-associated antigen 4 (CTLA-4), and CD244 (2B4) have all been associated with exhaustion of HCV-specific $CD8^+$ T cells (77, 81-83).

Continuous antigenic stimulation has been associated with a particularly high level of PD-1 expression on HCV-specific $CD8^+$ T cells compared with other inhibitory receptors (84). Emerging data further suggests that co-expression of multiple inhibitory receptors on HCV-specific $CD8^+$ T cells in the blood leads to more severely exhausted phenotype than expression of PD-1 alone. However, the location of viral-specific $CD8^+$ T cells may play an important role also. A study by Kroy et al. studied a large number of inhibitory receptors including PD-1, 2B4, CD160, TIM-3, LAG-3 and phenotyped both peripheral blood and intrahepatic HCV-specific $CD8^+$ T cells. They discovered that the markers PD-1 and 2B4 were more highly expressed in intrahepatic lymphocytes (IHLs) than peripheral blood irrespective of disease. Thus the liver may represent a microenvironment for T cell regulation as of yet understudied.

In vivo antibody-mediated blockade of these receptors can restore effector functions of HCV-specific $CD8^+$ T cells (85). Several studies in both LCMV mouse models of chronic infection and in patients with chronic HCV have suggested that interruption of a single co-regulatory pathway alone is insufficient to restore T cell function (70). However, what combinations of blockade required to restore T cell function is unclear. The first in vivo animal studies and clinical trials on anti-PD-1 (86) and anti-CTLA-4 (84) antibodies as immunotherapeutic interventions against HCV demonstrated individualized responses. Similar

therapies in chronic infections affecting the liver are being investigated. Maini et al. demonstrated that blocking PD-1 in combination with Tim-3 in chronic HBV infection improved and even recovered functional antiviral responses. This provides supporting evidence that reversal of exhaustion must be targeted by multiple co-regulatory pathways (87). Therefore, further understanding the repertoire for inhibitory receptor-mediated regulation of viral-specific $CD8^+$ T cells may help create more personalized and effective therapies to chronic HCV.

1.8 T cell inhibition

The negative pathways that influence T cell exhaustion can be grouped into 3 main categories: cell surface inhibitory receptors (such as PD-1), soluble factors (such as IL-10) and immunoregulatory cell types (such as $CD4^+$ regulatory T cells) (88). Here we concentrate on inhibitory receptors, which have a key role in many aspects of adaptive immunity that provide self-tolerance and prevention of autoimmunity. Although functional effector T cells can transiently express inhibitory receptors during activation, prolonged and/or high expression of multiple inhibitory receptors is a key feature of the exhaustion of $CD8^+$ T cells in both animal and humans (65). PD-1 and its ligand appear to be a major inhibitory receptor pathway involved in T cell exhaustion, and blocking this pathway during chronic LCMV infection invigorates virus-specific $CD8^+$ T cell responses and results in lower viral-load (42). T cell exhaustion is thus an active process at least partly under the control of inhibitory receptors. Therefore, the appropriate pathway could be targeted and severe exhaustion of $CD8^+$ T cells be

potentially reversed. PD-1 was the first inhibitory receptor discovered to play a central role in T cell dysfunction in HIV (89). In vivo blockade of PD-1 during infection with simian immunodeficiency virus (SIV) leads to a substantial improvement in virus-specific $CD8^+$ T cell responses but more importantly improves survival (90). The PD-1 pathway also has an important role in limiting the effectiveness of antigen-specific T cells during several other persisting viral and non-viral infections as well as in cancer (88). Early clinical trials suggest that blocking the PD-1 pathway could be a major immunotherapeutic strategy for achieving immunological control of tumors in humans (91). In addition to PD-1, many other cell surface inhibitory receptors co-regulate T cell exhaustion (70). Virus-specific $CD8^+$ T cells during chronic infection in animal models and in humans can also co-express LAG-3, CD244, CD160, TIM-3 and CTLA-4. The degree and combination of inhibitory receptor co-expression on individual $CD8^+$ T cells greatly alters the degree of T cell exhaustion (70). Understanding which combination contributes most in the precise biological context has been investigated via immunological blockade of these pathways. For example, at least partial restoration of $CD8^+$ T cell function has been showed by simultaneous blockade of PD-1 and CTLA-4 as well as PD-1 and LAG-3 (70).

Understanding the intracellular machinery downstream of inhibitory receptors is of paramount importance if we are to be able to manipulate them in order to modulate immune response to pathogen. The PD-1 pathway seems to strongly affect survival and proliferation of exhausted $CD8^+$ T cells (92). In contrast, LAG-3 affects cell cycle progression but has less influence on cell

survival or apoptosis (93). Although the molecular basis for inhibitory receptor regulation has not been defined, studies suggest that inhibitory receptors might attenuate T cell responses in multiple ways. CTLA-4 competes with CD28 for co-stimulatory ligands (94) and PD-1 recruits phosphatases (such as SHP-1, SHP-2) to TCR proximal signaling complexes and attenuates signaling (95). In addition, PD-1 ligation can induce genes such as BATF that encode molecules actively involved in inhibiting T cell function suggesting a potential new mechanism by which these receptors operate (96).

1.9 Purinergic pathway in T cell inhibition

The purinergic system has evolved to aid in the modulation of immune functions such as cellular proliferation, interaction and cytokine secretions (97). The purinergic mediators include ATP and its degradation products 5' cAMP, ADP and adenosine. ATP is released by cell destruction or actively through exocytosis of ATP-containing vesicles, nucleotide permeable channels (such as pannexin hemi-channels) or via transport vesicles and lysozomes. CD39 (Ectonucleotidase triphosphate diphosphohydrolase 1, ENTPDase1) hydrolyses extracellular ATP into AMP, which is further hydrolysed by CD73 (ecto-5'-nucleotidase) into adenosine. A number of other less well known enzymes are also involved in the metabolism of extracellular nucleotides including alkaline phosphatases, pyrophosphatases and phosphodiesterases as well as adenylate kinase which counteracts ATP degradation. Adenosine and ATP activate the P1 and P2 group of receptors respectively. P1 receptors include the adenosine

receptors A1, A2A, A2B and A3, whereas the P2 receptors are split into 2 subgroups, the ionotropic P2X and metabotropic P2Y receptors (98). CD39 is being increasingly used as a marker for $CD4^+$ Tregs, which are recognised as a primary regulator of immune tolerance preventing autoimmunity and inflammation in conditions such as asthma and inflammatory bowel disease. However, they can also prevent sterilising immunity to viruses and suppress antitumor effects (99, 100). The activity of CD39 is in tune with cell activity and therefore increased ATP metabolism occurs with stimulation of the TCR complex. Deaglio et al. demonstrated that CD39 activity together with CD73 used the purinergic system to induce suppression on effector T cells (Teff) via the production of adenosine from T regs (101). Adenosine causes immunosuppression by way of engaging with A_{2A} receptors on T_{EM} cells and down regulates nuclear factor-κB (NF-κB) (Figure 1.5) (101). This mechanism of T cell inhibition in Tregs may not be isolated in the $CD4^+$ compartment. Growing evidence suggests that $CD8^+$ T cells can also express CD39 and may act as a regulatory cell mediating inhibition of effector cells in the context of viral infection (102).

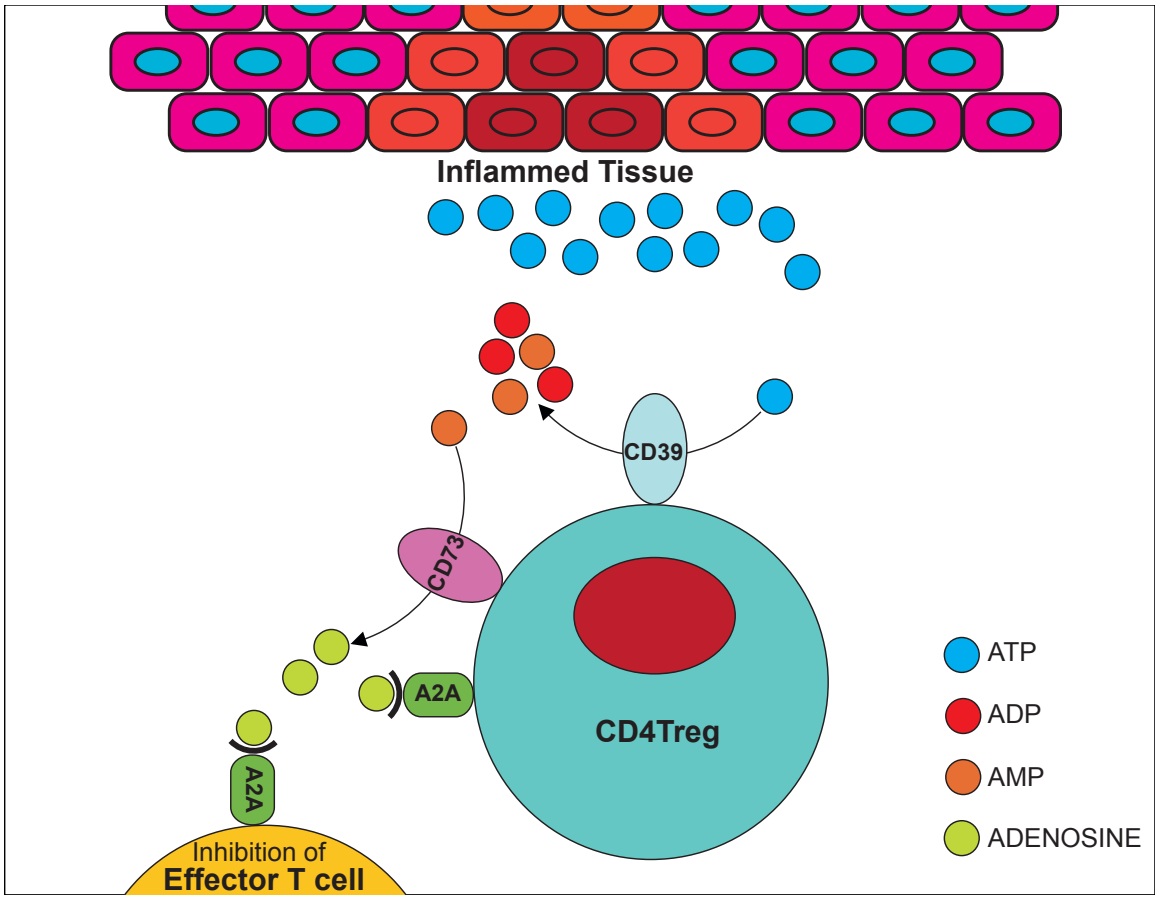


Figure 1.5. Action of CD39 on CD4 T regulatory cells.

1.10 Genome-wide expression profiles in chronic infection

A major advance in immunology has been the ability to define states of cellular differentiation not only by phenotype and function through flow cytometry but also through the use of comprehensive gene transcriptional profiles. Over the last 20 years large-scale high throughput gene expression studies have become feasible. This ability to profile small numbers of cells genome-wide has changed

the way adaptive immunity can be studied. High-density microarrays were introduced in 1995 and by the turn of the century applied to immunological questions such as gene expression patterns of activated T and B cells (103, 104). Confirming what was already discovered and defining genes associated with lymphocyte activation was key in validating microarrays for use in biological discovery.

Initial gene expression profiling studies of differentially expressed phenotypic markers between viral-specific $CD8^+$ T cells were conducted in mouse models of acute infection (105). Comparison of memory $CD8^+$ T cells at day 30 with naive and effector cells at day 8 demonstrated a broad and distinct difference in gene expression patterns. Out of this came several striking findings, firstly the difference in gene patterns suggested that memory $CD8^+$ T cells were a distinct lineage of cells rather than quiescent effectors. Secondly, memory development occurred gradually long after virus had been cleared suggesting signals independent of the virus take part in memory development. Thirdly, the number of genes required to regulate these changes was broad and not just 1 or 2 individual genes. As genome-wide expression profiles became more adopted discovery of key molecules came about such as CD127 (106), PD1 (107), Blimp-1 (108) to name but a few.

This analysis was then applied to human cohorts in which the expression of candidate $CD8^+$ T cell markers in patients with defined clinical phenotypes, such as those that do or do not control viral load in HIV or, HBV infection (109, 110). For example, Lopes et al. examined signatures of HBV-specific $CD8^+$ T

cells between patients that developed chronic HBV infection and those that resolved virus. Signatures enriched with apoptotic genes revealed that the proapoptotic protein Bcl2-interacting mediator (BIM) was upregulated in patients with chronic HBV infection. Subsequent blockade of Bim resulted in recovery of HBV-specific $CD8^+$ T cells.

In chronic infections such as these, the antigen-specific $CD8^+$ T cell population contain cells with longevity and good proliferative potential as well as those with an exhaustion phenotype that has lost cardinal functions such as interleukin 2 (IL2) secretion, proliferation potential and cytotoxicity (111). Similar approaches can be used to analyse HCV-specific $CD8^+$ T cells, which also represent a diverse collection of T cell states reflected by complex patterns of gene expression. Identifying those patterns that represent protective T cell responses poses a particular challenge to immunologists since flow cytometry is limited in the number of candidate phenotypic markers that can be simultaneously surveyed in antigen-specific T cells (49). It is likely that important biological phenotypes have been undiscovered and lie within what we perceive as homogenous populations of cells.

In addition, the comparison of gene expression profiles from 2 cell types can lead to hundreds or thousands of potentially interesting genes. Without a biological context for these genes, choosing relevant findings may be based on the researcher's area of expertise. It is practically impossible to pick up subtle coordinate changes in gene expression across many genes by focusing on individual genes alone. Therefore, successful techniques have been developed

to look at coordinated patterns of gene expression that have already been predefined with a biological context in pre-existing databases.

1.11 Gene Set Enrichment Analysis

Several groups have developed methods to test the enrichment of biologically meaningful gene sets and this has now become an essential tool to integrate genomic data from different experiments and platforms (112). Enrichment analysis starts with obtaining a profile of differentially expressed genes from 2 cellular contexts such as $CD39^+$ and $CD39^-$ T cells. By using a predetermined gene set made of genes that are biologically related, one can test whether that set of genes ranks towards the top (or bottom) of the rank ordered set of all genes differentially expresses between $CD39^+$ versus $CD39^-$ T cell groups. Sets of genes that are highly related to one cell type should be distributed at the top of the rank ordered list; gene sets that are not related would be expected to be randomly distributed throughout the list. Gene set enrichment analysis provides a score to measure the degree of enrichment of a given gene set at the top (highly correlated) or bottom (anti-correlated) of the rank ordered data set as a statistical value for the confidence of the enrichment compared to chance alone. Enrichment analysis is only as useful as is the availability of annotated catalogues of biologically meaningful gene sets there are. Available and growing databases include Gene Ontology (GO) (113), Kegg (114) and TRANSFAC (115). However, limitations to these gene sets do exist. Firstly, the fraction of genes that are incorrectly annotated is not known. Secondly, signals

from uncoordinated sets of genes may overshadow coordinated sets of genes in any given pathway. To counter these limitations, a growing library of experimentally derived data sets exists (116). Unlike GO and KEGG databases these have been curated from analyses of previous gene expression profiling experiments and require no researcher-dependent annotation. However, the biological context is inferred from the experiment that created the data. But since the ability to integrate and compare microarray data is based on enrichment scored by relative ranks of gene sets rather than absolute values, comparison can be made across platforms and species such as comparing viral-specific $CD8^+$ T cells in humans with mice.

These tools have now allowed an alternative approach to identify broad molecular phenotypes of and HIV-specific $CD8^+$ T cells using cDNA microarrays (96, 110). Quigley et al. found that the global pattern of gene expression in HIV-specific $CD8^+$ T cells from individuals with chronic elevation of viral load (progressors) compared with those that spontaneously control viral replication (controllers) was highly similar to that seen in mouse LCMV-specific $CD8^+$ T cells responding to chronic versus acute viral infection. Moreover, analysis of the genes that distinguish between functional and exhausted $CD8^+$ T cells identified a new role for the transcription factor BATF in $CD8^+$ T cell inhibition. However, such approaches rely on the supervised analysis of T cells from distinct clinical states such as those that do or do not control viral load in HIV infection.

Genome-wide expression profiling of multimer-specific T cells also offers the ability to discover novel molecular phenotypes that are responding to the

same viral antigen. This unsupervised approach takes advantage of the fact that samples that belong to a distinct biological class or state will share similar patterns of expression distributed across large numbers of genes. This assumption is supported by increasing data suggesting that differentiation within the hematopoietic system is governed by well circumscribed, often redundantly used patterns of gene expression (117). Based on robust similarities in the pattern of expressed genes, samples can be grouped together into classes of samples that are putatively similar in important biological characteristics. This approach has been widely used in cancer genomics to identify biologically distinct subtypes of samples associated with discrete clinical outcomes within a single tumor type (118).

1.12 Single cell heterogeneity

It is clear that gene expression data, initially from microarrays and more recently from next generation sequencing have provided invaluable insight in transcription factor networks associated with cellular differentiation states (119). However, most of these analyses on viral-specific $CD8^+$ T cells have been performed on groups of cells. Individual cells make cell fate decisions, thus the averaged expression of transcription factors from a pool of cells may mask important single cell dynamics. Thus a key issue remains as to whether these phenotypes are a result of gene expression of all or in a distinct subset of cells (Figure 1.6).

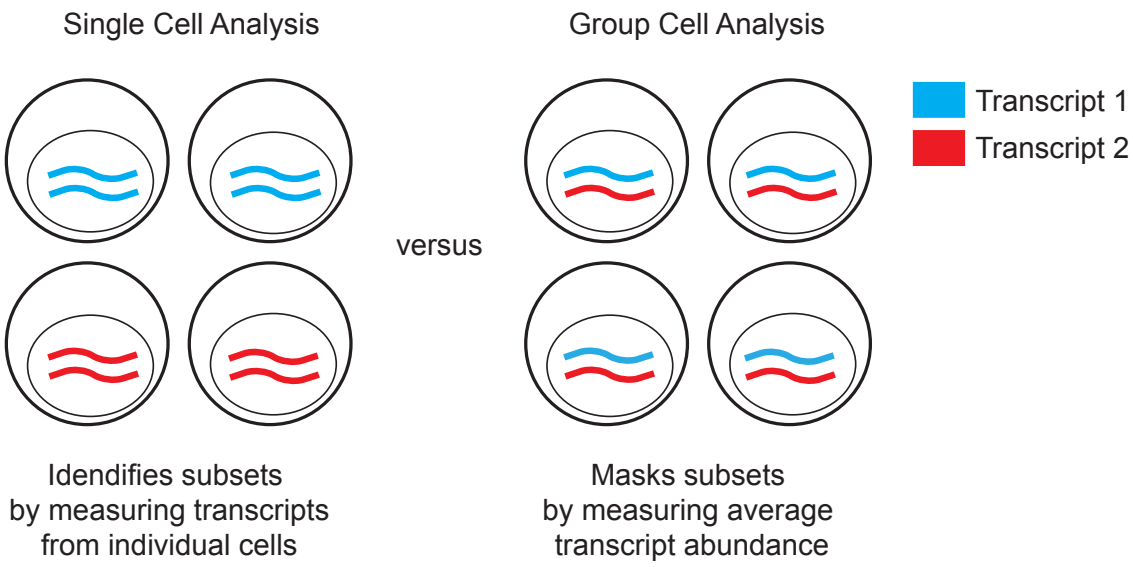


Figure 1.6. Single cell gene expression analysis can reveal differences among cells even in populations that were thought to be homogeneous.

The ability to profile individual cells enables not only the identification of previously unrecognized subpopulations, but also the dissection of regulatory networks. Such approaches have become possible by recent technologies allowing analysis of single antigen-specific T cells with many more parameters than previously possible. Flatz et al. analyzed the gene expression profiles of single HIV-specific $CD8^+$ T cells elicited by vaccine regimens in a mouse model (120). Using a nanofluidic device, the expression of 96 genes was assayed from cDNA generated from individual cells, revealing a complex picture of gene expression in individual cells generated by vaccination. Novel phenotypes of T cells within each vaccine response could be defined based on the expression of previous uncharacterised combinations of marker genes. These techniques can

be used to identify previously unrecognised T cell subsets within a population of HCV-specific $CD8^+$ T cells.

1.13 Aims of thesis

The aims of my thesis are to use gene transcriptional profiling and systems biology approaches to analyse $CD8^+$ T cell dysfunction in HCV infection. I will begin by using gene transcription profiles of HCV-specific $CD8^+$ T cells to identify novel subclass-specific gene markers and analyse these at the single cell level. Next I will provide an in-depth phenotypic analysis of one subclass-specific marker (CD39) in human and mouse models of chronic infection by flow cytometry and further gene transcription profiling.

Aims:

1. Phenotype HCV-specific $CD8^+$ T cells by single cell gene expression analysis using candidate gene markers identified from genome-wide expression profiles.

2. Phenotypic analysis of CD39 on virus-specific $CD8^+$ T cells in HCV, HIV and mouse models of chronic viral infection.

3. Use genome-wide transcription signatures to profile the molecular characteristics of CD39 in HCV infection.

CHAPTER - 2

METHODS

2.1 Human Subjects

Healthy human donors were recruited at the Kraft Family Blood Donor Center, Dana-Farber Cancer Institute, Boston, Massachusetts, USA (DFCI) with written informed consent following approval by Partners Institutional Review Board. Yellow fever vaccine samples were kindly donated by the Ahmed laboratory (Emory University, Atlanta, Georgia, USA).

All human subjects with HCV infection were recruited at the Massachusetts General Hospital (MGH) Gastrointestinal Unit and the Department of Surgery with written consent in accordance with the IRB approved study: “Cell mediated immunity in Hepatitis C virus infection”; Protocol # 1999- P-004983/54; MGH Legacy #: 90-7246 (Appendix Figure 9.1). HCV chronics (n = 27) were defined by positive anti-HCV antibody and detectable viral load. HCV resolvers (n = 14) were defined as having an undetectable viral load for at least 6 months (Appendix Table 9.1). All HCV patients were treatment naive and obtained between 5.9 and 237.3 weeks post infection. HCV RNA levels were determined using the VERSANT HCV RNA 3.0 (bDNA 3.0) assay (Bayer Diagnostics). Additional information on HCV donors not included above (15 with chronic infection and 12 that resolved infection) can be obtained in appendices (Appendix Table 9.3). Gene expression profiling results for this cohort are in chapter 3 (Figure 3. 1, 3.2) where work was performed by Lauer and Haining laboratories (complete list of enriched genes not shown).

All HIV infected cohorts were recruited after written informed consent from

the Ragon Institute, Massachusetts General Hospital, Boston USA and the Peter Medawar Building for Pathogen Research, Oxford, UK where ethics approval was given by the Oxford Research Committee. HIV controllers included elite controllers (n = 5) defined as having HIV RNA below the level of detection (<75 copies per ml); viremic controllers (n = 7) with HIV RNA levels < 2,000 copies per ml. HIV chronic progressors (n = 28) were defined as having > 2,000 copies per ml. All subjects except one were off therapy (Appendix Table 9.2). Viral load in chronic infection was measured using the Roche Amplicor version 1.5 assay.

2.2 Leukopak cone processing

Peripheral blood mononuclear cells (PBMCs) from healthy donors were derived from leukopaks. Blood was extracted from the leukopak after sterilisation of instruments and tissue culture hood with 70% ethanol. The leukopak cone was placed into a 50 ml falcon and blood allowed to slowly drain into the tube. A needle was placed through the hole at the top of the cone and using the 50 ml syringe, 40 ml of sterile phosphate buffered solution (PBS) was washed through the cone. 15 ml of Ficoll-Plaque (GE Healthcare) was carefully pipetted to the bottom of the falcon tube underneath the blood layer and then spun without accelerator or break at 22^{o}C, 2220 revolutions per minute (rpm), for 22 minutes. Thereafter, the PBMC layer was removed with a pipette, washed with 50 ml of Roswell Park Memorial Institute Medium **(**RPMI) (Life technologies) media and spun down again for 10 minutes at 1800 rpm. Repeat washing was carried out and re-centrifuged for 5 minutes at 1500 rpm. Cells were re-suspended with a further 50 ml of RPMI and counted, prior to being frozen for storage or used

directly.

2.3 Fresh blood processing

Blood preparation was performed in a standard defined protocol used in the Peter Medawar Building for Pathogen Research laboratories. All infectious blood (HCV and HIV) or blood products isolated were stored and processed in a designated Category III laboratory either at US (MGH, RAGON, DFCI) or UK (Peter Medawar) institutes. Venous blood was collected in heparinised tubes and then diluted 1:1 in R10 media (RPMI/10% fetal bovine serum/100 U/mL penicillin and 100 µg/mL streptavidin) in 50 ml Falcon tubes. 10 ml of Ficoll-Plaque was carefully pipetted into the bottom of the Falcon tube underneath the blood. This tube was centrifuged for 25 minutes at 2200 rpm, 25°C with a slow start and stop (i.e. brake off). The lymphocyte layer was then removed with a pipette and re-suspended into a total of 50 ml of R10 media. The cells were then centrifuged at 1700 rpm for 10 minutes at 20°C. Red blood cells (RBC) then lysed using 10 ml of RBC lysis buffer (ebioscience) for 10 minutes at room temperature and washed with 10 ml PBS prior to cryopreservation in 4% formaldehyde and 10% Fetal Bovine Serum or further processed. All freshly isolated PBMCs were used the day processed or rested overnight.

2.4 PBMC thawing

PBMC samples from patients with HIV or HCV were previously cryopreserved. Thawing media was prepared in the following dilutions:

-1:10 the 10X stock CTL-Wash (supplement medium solution in RPMI)
-1:100 L-glutamine
-30U/ml DNAse 1
-1:100 L-glutamine Penicillin Streptomycin

The thawing media and the R10 media were pre-warmed in 37^{o}C (+/- 1^{o}C) in the water bath. The required PBMC vial(s) were removed from liquid nitrogen and placed on dry ice until ready to thaw. The cells were thawed with gentle agitation in a 37^{o}C (+/- 1^{o}C) water bath. Using a sterile Pasteur pipette each thawed cell suspension was transferred to an appropriately labeled 15 ml tube containing 5ml of thawing media. The total volume was made up per falcon to 10 ml with R10 media then centrifuged at room temperature (18-20^{o}C) at 16000 rpm for 10 minutes. The supernatant was aspirated and the cells re-suspended in R10 media and counted.

2.5 HLA Class I Tetramers

Major histocompatibility complex (MHC) class I HIV Gag-specific tetramers were generated by the Goulder lab as previously described (One-pot, mix-and-read peptide-MHC tetramers) and Alter lab as described (121)**.** HCV-specific MHC class I multimers (tetramers and pentamers) were provided by the Lauer lab (Proimmune). CMV- and EBV-specific MHC class I dextramers conjugated with FITC and APC were purchased from Immudex. YFV-specific tetramers were conjugated with APC and donated by the Ahmed laboratory (Emory University, Atlanta, Georgia, USA). Mouse MHC class I tetramers of H-2D^{b} complexed with LCMV $GP_{276-286}$ and GP_{33-41} as previously described (66, 122). Biotinylated complexes were tetramerised using allophycocyanin-

conjugated (APC) streptavidin (Molecular Probes). The complete list of multimers can be found in appendices (Appendix Table 9.4).

2.6 Antibodies and flow cytometry

The following fluorochrome-conjugated antibodies were used: anti-huCD8α, anti-huCD4, anti-huCD3, anti-huCD39, anti-huPD-1, anti-huCD25, anti-huCCR7, anti-huCD45RA, anti-huCD160, anti-huCD161, anti-huCD69, anti-huCD38, anti-huHLA DR, anti-mCD8, anti-mCD4, anti-mCD3, anti-mCD39 anti-mCD244, anti-mPD-1, anti-mLag3, anti-mCD44 (IM7), anti-mCD127 (all from Biolegend), anti-mT-Bet (04-46; BD Pharmingen), anti-mEomes(Dan11mag; eBioscience). Peripheral blood mononuclear cells (PBMCs) were thawed in complete RPMI (10% fetal calf serum [FCS], 1% penicillin/streptomycin, 1% glutamine); rested for 1 hour at 37°C in 5% CO2. 5 million PBMC cells were re-suspended in 50ul of staining buffer (PBS, 0.1% FBS) and stained with the LIVE⁄DEAD Fixable Aqua Dead Cell Stain Kit (Invitrogen) for exclusion of dead cells prior to staining with MHC Class I multimer for 10 minutes at room temperature to identify antigen-specific cells. Surface staining was performed for 25 minutes in the dark at 4°C and cells were fixed using 4% paraformaldehyde for 20 minutes prior to analysis. Intracellular staining was performed following surface stains and fixed and permeabilized using the FoxP3/Transcription Factor Staining Buffer Set (eBioscience). Cells were sorted by BD FACS ARIA II and all other analyses were performed on BD LSR II and BD LSR Fortessa flow

cytometers equipped with FACSDiva v6.1. Gates were set using Full Minus One (FMO) controls. Data were analyzed using Flowjo software (Treestar).

2.7 Cell culture and stimulation of primary T lymphocytes

Primary human $CD8^+$ T cells were isolated by CD8 negative selection MACS kits (Miltenyi) and were cultured in RPMI-1640 medium supplemented with 10% FBS, 10 mM HEPES, 50 U/ml of penicillin, 50 µg/ml of streptomycin and 50 µM β-mercaptoethanol. Naive cells were stimulated for 5 days with soluble plate-bound anti-CD3 (4 µg/ml; 2C11; BD Pharmingen) and anti-CD28 (4 µg/ml; 37.51; BD Pharmingen) for the generation of $CD8^+CD39^+$ T cells that were used for optimization of ATP hydrolysis assay experiment.

2.8 Cell culture of CEM and E61 Jurkat cells

CEM and E61 Jurkat cells were cultured in RPMI-1640 medium supplemented with 10% FBS, 10 mM HEPES, 50 U/ml of penicillin, 50 µg/ml of streptomycin and 50 µM β-mercaptoethanol at 37°C in 5% CO2 prior to use in ATP hydrolysis experiments.

2.9 Luciferase assay for ATP analysis

Measurement of ATP hydrolysis was performed on healthy primary $CD3^+$ T cells by addition of exogenous ATP (sigma) at a concentration of 100 uM to cell cultures enriched and depleted for CD39 expression. Cells were incubated in cell culture media (RPMI, 10% FBS, 100 iu/ml penicillin, 100 ug/ml streptomycin) for

1 hour after addition of a CD39 inhibitor (Sodium polyoxotungstate 1). Supernatant was removed and mixed at a 1:1 ratio with cell-titer glo (sigma) and incubated for 10 minutes. Luminescence of cell titer glo bound to ATP was measured at 520 nm on fluostar Omega luminometer.

2.10 HPLC analysis of ATP levels

The concentrations of ATP hydrolyzed by effector $CD39^+CD8^+$ from samples of chronic HCV infected subjects (n=6) were assessed by high performance liquid chromatography (HPLC) as previously described (123). Briefly, 10,000 $CD39^+$ $CD8^+$ T cells were sorted and placed on ice to minimize ATP production by cells. 20 mM of ATP was added and incubated for 1 hour at 37°C in 5% CO2 to allow for cellular activity to increase and CD39-mediated ATP hydrolysis to occur. Samples were then placed in an ice bath for 10 minutes to halt enzymatic activity, collected, and centrifuged for 10 minutes at 2,000 rpm and 0°C. Cells were discarded and supernatant centrifuged again to remove remaining cells (5,000 rpm, 5 min, 0°C). The resulting RPMI samples (160 ml) were treated with 10 ml of an 8 M perchloric acid solution (Sigma-Aldrich, St. Louis, MO, USA) and centrifuged at 13,000 rpm for 10 minutes at 0°C to precipitate proteins. In order to neutralize the pH of the resulting solutions and to remove lipids, supernatants (80 ml) were treated with 4 M K_2HPO_4 (8 ml) and tri-N-octylamine (50 ml). These samples were mixed with 50 ml of 1,1,2-trichloro-trifluoroethane and centrifuged (13,000 rpm, 10 minutes, 0°C) and this last lipid extraction step was repeated once. The resulting supernatants were subjected to

the following procedure to generate fluorescent etheno-adenine products: 150 ml supernatant (or nucleotide standard solution) was incubated at 72°C for 30 minutes with 250 mM Na_2HPO_4 (20 ml) and 1 M chloroacetaldehyde (30 ml; Sigma-Aldrich) in a final reaction volume of 200 ml resulting in the formation of 1,N6-etheno derivatives as previously described (123). Samples were placed on ice, alkalinized with 0.5 M NH_4HCO_3 (50 ml), filtered with 1 ml syringe and 0.45 mM filter and analyzed using a Waters HPLC system (Waters, Milford, MA, USA) and Supelcosil 3 μM LC-18T reverse phase column (Sigma) consisting of a gradient system described previously, a Waters autosampler, and a Waters 474 fluorescence detector (124). Empower2 software was used for the analysis of data and all samples were compared with water and ATP standard controls as well as a sample with no cells to determine background degradation of ATP.

2.11 Real-Time PCR

First strand cDNA synthesis was performed in 0.2 ml PCR tubes. First, RNA was added to 0.1 ul random primers (Invitrogen), 1ul oligo (dT) primers (50uM) (Invitrogen, 0.5ul dNTPs (10mM) (Invitrogen) and RNase free water to make the reaction up to 13ul. The reaction was placed into a thermal cycler (Biorad) and heated to 65°C for 10 minutes. After this the Superscript III reverse transcriptase kit (Invitrogen) was used to make the following reaction mix; 4ul 5X first strand buffer, 1ul 0.1M DTT, 1ul RNAseOUT recombinant RNAase inhibitor. This was added to reactions and placed in the thermal cycler at 50°C for 60 minutes. 1ul

cDNA was added to: 6.7ul RNase free water, 1 ul forward primer, 1 ul reverse primer, 0.3 ul of gene probe, 10ul Fastart universal probe (Roche). Samples were loaded onto a 96 well block and set at: 95°C for 15 minutes, 55 cycles of 94°C for 15 seconds, 54°C for 30 seconds and 72°C for 30 seconds. Light cycler 480 software (Roche) was used for analysis of Real Time Polymerase Chain Reaction data (R-T PCR).

2.12 RNA extraction

RNA extraction from $CD8^+$ T cells was performed using the RNeasy mini kit. After obtaining cell pellets 350 ul of RLT was added. Samples were then homogenized on QIAShedder columns (Qiagen) and centrifuge for 2 minutes at maximum speed. 350 ul of 70% ethanol was added to the homogenized sample then spun at 10,000 x g for 15 seconds. 700 ul of RW1 buffer added to columns and mixed by inverting. Columns were centrifuged at 10,000 x g for 15 seconds and flow through discarded. This step was repeated and then a further 500 ul added to columns, which were centrifuged at 10,000g for 2 minutes. Collection tubes were changed and centrifuged dry. 30ul of RNAse free water was used to elute RNA by centrifugation at 10,000 x g for 60 seconds. 1.25ul of each sample was analysed using nanodrop 2000 spectrophotometer.

2.13 Microarray data acquisition and analysis

Effector $CD39^+$ $CD8^+$ T cells from HCV samples were sorted and pelleted and re-suspended in TRIzol (Invitrogen). RNA extraction was performed using

the RNAdvance Tissue Isolation kit (Agencourt). Purification of RNA begins with homogenization and lysis of cells with lysis buffer prepared by manufacturers' instructions. Following lysis, 400 ul homogenized lysate is transferred to 1.7 ml microcentrifuge tube and incubated at 37 °C. Next 400 ul binding buffer slowly pipetted and placed in magnet for 6minutes. Supernantent is removed and and washed with 800ul of wash buffer. Sample is put back on magent for 5minutes and supernatant removed again, followed by wash with 800 ul 70% ethanol. Sample is placed back on magnet for 5minutes and as much ethanol removed. 100 ul of DNase solution added, sample is sealed and incubated at 37 °C for 15 minutes to facilitate digestion of DNA. Sample washed with 550 ul wash buffer and 70% Ethanol 3 times and sample elutes with 40 ul nuclease free water and transferred into a fresh tube.

Concentrations of total RNA were determined with a Nanodrop spectrophotometer or Ribogreen RNA quantification kits (Molecular Probes/Invitrogen). RNA purity was determined by Bioanalyzer 2100 traces (Agilent Technologies). Total RNA was amplified with the WT-Ovation Pico RNA Amplification system (NuGEN) according to the manufacturer's instructions. Firstly, first strand cDNA is prepared from total RNA using a unique first strand DNA/RNA chimeric primer mix and reverse transcriptase (RT). The primers have a DNA portion that hybridizes either to the 5' portion of the poly (A) sequence or randomly across the transcript. RT extends the 3' DNA end of each primer generating first strand cDNA. The resulting cDNA/mRNA hybrid molecule contains a unique RNA sequence at the 5' end of the cDNA strand. Secondly, is

the generation of a DNA/RNA Heteroduplex Double Strand cDNA. Fragmentation of the mRNA within the cDNA/mRNA complex creates priming sites for DNA polymerase to synthesize a second strand, which includes DNA complementary to the 5' unique sequence from the first strand chimeric primers. The result is a double stranded cDNA with a unique DNA/RNA heteroduplex at one end. Thirdly, apply the NuGEN SPIA DNA/RNA chimeric primer, DNA polymerase and RNase H in a homogeneous isothermal assay that provides highly efficient amplification of DNA sequences. RNase H is used to degrade RNA in the DNA/RNA heteroduplex at the 5' end of the first cDNA strand. This results in the exposure of a DNA sequence that is available for binding a second SPIA DNA/RNA chimeric primer. DNA polymerase then initiates replication at the 3' end of the primer, displacing the existing forward strand. The RNA portion at the 5' end of the newly synthesized strand is again removed by RNase H, exposing part of the unique priming site for initiation of the next round of cDNA synthesis. The process of SPIA DNA/RNA primer binding, DNA replication, strand displacement and RNA cleavage is repeated, resulting in rapid accumulation of cDNA with sequence complementary to the original mRNA. An average mRNA amplification of 15,000-fold is observed with 500 pg starting total RNA. After fragmentation and biotinylation, cDNA was hybridized to Affymetrix HG-U133A2.0 microarrays.

Prior to analysis, microarray data were pre-processed and normalized using robust multichip averaging, as previously described (125). Differentially gene expression and consensus clustering was performed using Gene-E software (www.broadinstitute.org/cancer/software/GENE-E/), and gene set

enrichment analysis was performed as described previously using gene sets from MSigDB (116) or published resources (126, 127).

2.14 Clinical data collection

Clinical data was acquired from hospital databases in collaboration with the Lauer laboratory (MGH, Gastrointestinal Unit), Alter laboratory (Ragon Institute), Boston, USA as well as the Goulder laboratory (Peter Medawar Building) Oxford, UK (Appendix Tables 9.1-9.3). Data on estimated time from infection, age, sex, HLA type, viral load and whether patients resolved HCV infection or developed chronic disease was correlated with identified molecular subclasses of HCV-specific T cell response.

2.15 Single cell sort optimisation

Optimisation of single cell sorting was performed on the BD FACS ARIA II on E61 Jurkat cells stained with Hoechst blue (ATCC) and on FITC conjugated fluorescent beads (BD bioscoences). A range of flow rates and nozzle sizes were tested to sort single cells onto flat bottom 96 well plates (Costar). Fluorescent cells and beads were detected and counted using the Nikon Eclipse inverted microscope and sorting efficiency calculated.

2.16 Gene expression profiling of $CD8^+$ T cells by fluidigm platform

Healthy human PBMCs obtained by density centrifugation were stained with antibodies anti-huCD4-FITC, anti-huCCR7-PE, anti-huCD8-PE Cy7, anti-huCD3-Pac Blue, anti-huCD45ra-APC (Biolegend), aqua fluorescent HV500 live dead marker (Invitrogen) and a range of multimers specific to HIV, HCV, CMV, EBV and YFV tagged with APC fluorochrome (Proimmune, Immudex, Beckmann). $CD8^+$ T cells were separated into memory populations based on CD45ra and CCR7 expression (effector memory RA positive, effector memory RA negative, central memory and naive) using the BD FACS ARIA II. Memory populations were sorted into a 96 well PCR plate (Eppendorf) containing 10.1ul reverse transcriptase specific target amplification (RT-STA) master mix solution (Appendix Table 9.5). Individual antigen-specific cells were sorted by flow cytometry into a 96 well PCR plate with the mentioned reaction mix above. Samples were placed on dry ice and reverse transcribed using Bio rad C1000 touch thermal cycler (Appendix Table 9.6) 96 genes (Taqman) were tested including 4 control genes and 5 genes of interest (Appendix Table 9.9). After cDNA has been formed, the 96 x 96 dynamic integrated fluidics chip was loaded with control line fluid and placed in the IFC controller to prime the control fluid into the chip. After priming was complete the chip was loaded with Taqman probe mix (Appendix Table 9.7) and cDNA sample mix (Appendix Table 9.8) into inlets either side. The chip is then run allowing 96 x 96 Taqman PCR reactions to be read in parallel using the Biomark real time PCR reader (Fluidigm).

2.17 Mice and infections

All mice were used according to the Harvard Medical School Standing Committee on Animals and National Institutes of Animal Healthcare Guidelines. Wildtype C57BL/6J mice were purchased from The Jackson Laboratory and CD39 null mice kindly donated from the Robson lab, Beth Israel Hospital, Boston, USA. 6-8 week old female mice were infected with 2 x 10^5 p.f.u. LCMV-Armstrong intraperitonealy or 4 x 10^6 p.f.u. LCMV-Clone 13 intravenously and analyzed at indicated time points by homogenizing the spleen into a single-cell suspension, ACK lysis of red blood cells, followed by antibody staining. Viruses were propagated by triple-plaque purification on Vero cells and grown in BHK-21 cells (128). Virus stocks at passage 1 or 2 were harvested for experiments. Virus was obtained from the Wherry lab, University of Pennsylvania, PA, USA. In vivo inhibition of CD39 with POM 1 Wild type mice were administered 10mg/kg/day of the CD39 inhibitor sodium polyoxotungstate (POM 1) (n=5) or 1 ml of placebo (n=5) water intra-veneously at Day 0 after infection with chronic clone 13 strain of LCMV.

2.18 Statistical analysis

All statistical analysis was performed on PRISM software packaging. Data was expressed as mean +/- Standard Error Mean (SEM) where applicable. Analysis of paired data was conducted using paired t-test. However, if sample were not paired, an unpaired t-test was used (with non-assumption of normal distribution). A p-value of < 0.05 was considered significant, except in the gene

array study where a raw p-value of < 0.01 was used. For multiple sample sets one-way ANOVA test was applied as appropriate.

2.19 GSEA

Prior to analysis, microarray data were pre-processed and normalised using robust multichip averaging (RMA), as previously described (125). Differentially gene expression and consensus clustering was performed using the web based server Gene-E software (www.broadinstitute.org/cancer/software/GENE-E/), and gene set enrichment analysis was performed as described previously using either gene sets from web based GSEA server Molecular Signatures Database (MSigDB) (116) or published resources (126, 127).

2.20 Single cell analysis

Samples of single cell gene expression data were first examined for quality control analysis by removal of cells that did not express the control gene Beta- 2 microglobulin (B2M), as this was deemed as an unsuccessful reaction. Next samples were analyzed on AABLE as 1 vs 100 cell analysis to determine a linear relationship between the expression of a single cell and gene expression of 100 cells. Finally once appropriate samples were excluded, gene expression data was analysed by the online program Gene e to identify novel subsets of T cells within the viral-specific pool.

CHAPTER 3

Single cell analysis of HCV-specific CD8$^+$ T cells

3.1 Introduction

During the course of viral infection, $CD8^+$ T cells that respond to antigen become activated, rapidly proliferate and gain a range of effector functions (49). The resulting large number of effector cells then contract leaving a small subset of highly heterogeneous antigen-specific memory $CD8^+$ T cells (129). In acute infections this group of cells display a high degree of proliferative potential but in chronic infections such as HCV, HBV and HIV a loss in functional capacity marked by reduced cytokine secretion, proliferation and cytolytic activity is observed. The latter results in a poor ability of virus-specific $CD8^+$ T cells to control virus and is a process known as T cell exhaustion (71). Furthermore, Infection with HCV can result in chronic infection or complete resolution of virus in approximately 30% of individuals. Thus, HCV represents a paradigm of human infection where virus-specific $CD8^+$ T cells display a wide variability of states and functionality. Being able to identify and characterise this heterogeneity has been a major task for immunologists in order to understand how effective T cell immunity is maintained. A number of novel technologies and analytical tools have developed over the last 15 years allowing for a more detailed analysis of the antigen-specific pool (130). Much of this has come in the form of genomic studies and systems biology approaches that are able to deconvolve complex patterns of gene expression and identify novel phenotypes (131).

Traditional methods of analysing HCV-specific $CD8^+$ T cells have involved the use of MHC tetramers that allow antigen-specific $CD8^+$ T cells to be identified agnostic of other surface markers. When combined with other fluorochrome-

conjugated markers, multi-parameter flow cytometery allows analysis of up to 20 markers on a single cell (132). More recent technologies have expanded on this by using transition element isotope chelated tags for atomic mass spectrometric analysis of up to 34 parameters simultaneously (133, 134). This combined with powerful software applications such as SPADE and flowjo as well as clustering algorithms such as Principal Components Analysis (PCA) have helped to identify much greater complexity in the virus-specific $CD8^+$ T cell compartment than previously appreciated. However, analysis by these techniques is limited by the number of antibodies available as well as the number of parameters able to be analysed.

cDNA microarrays have been increasingly used to generate genome-wide expression profiles of antigen-specific T cells in order to increase the depth of phenotypic analysis beyond that of flow cytometry. (96, 135) Quigley et al. used gene expression profiles to identify similar signatures from HIV-specific T cells of HIV progressors versus controllers to that in mouse LCMV-specific T cells responding to chronic versus acute viral infection (96). In addition, analysis of genes enriched in $CD8^+$ T cell exhaustion identified the transcription factor BATF's role in T cell inhibition through modulation of AP-1 target genes. Thus, analysis of genome-wide expression profiles can be used to identify novel genes that are mediators of T cell dysfunction in chronic viral infection which would be far more difficult to achieve by antibody based techniques alone.

An important consideration in the analysis of gene expression profiles is whether clustering techniques are supervised or unsupervised. Analyses of

genome-wide expression profiles of HIV-specific CD8$^+$ T cells above was performed through a supervised approach. This means that samples were segregated into subclasses based on known clinical phenotypes such as developing chronic infection (progressor) or resolution of virus (controller). In contrast an unsupervised analysis makes no a priori assumption on the number of subclasses that exist within the dataset. Thus, by applying clustering approaches to samples and genes of genome-wide expression profiles, novel subclasses of samples can be identified based on similar patterns of gene expression. This has already been applied to cancer genomics and correlated to malignancies with differing prognoses (136). I demonstrate how novel subclasses of CD8$^+$ T cell response to HCV infection can be identified by unsupervised analysis of HCV-specific CD8$^+$ T cells.

However, the identification of novel phenotypes of samples through analysis of genome-wide expression profiles of cell population assumes that all cells are equivalent as this method averages population data. This may vastly oversimplify cell phenotype and functionality as well as mask subgroups of cells that are predominately responsible for the novel phenotypes discovered (137). Recent technological advances have allowed the coupling of microfluidic devices with high throughput qRT-PCR enabling a detailed analysis of CD8$^+$ T cells (120, 138). Using such techniques Flatz et al. were able to identify gene expression profiles of single HIV-specific CD8$^+$ T cells elicited by vaccine regimens in a mouse model (120). Novel subgroups of T cells within each vaccine response were identified based on the expression of combinations of marker genes that

were previously not described. I show how these techniques can be applied to the analysis of HCV-specific $CD8^+$ T cells to potentially identify novel subgroups of virus-specific $CD8^+$ T cell response.

3.2 Aims/Hypotheses

Specific Aim 1 (SA1): Use genome-wide expression profiles to discover novel subclasses of $CD8^+$ T cell response to HCV infection and identify subclass-specific gene markers.

Hypothesis 1: an unsupervised analysis of genome-wide expression profiles of antigen specific $CD8^+$ T cells can be used to identify novel subclasses of $CD8^+$ T cell responses in HCV infection.

Specific Aim 2 (SA2): Optimise the Biomark single cell platform for analysis of a range of viral-specific $CD8^+$ T cells using subclass-specific gene markers.

Hypothesis 2: Subclass specific markers associated with novel HCV-specific $CD8^+$ T cell responses will be differentially expressed in viral-specific $CD8^+$ T cells from acute and chronic infections.

Specific Aim 3 (SA3): Use single cell gene expression profiling to identify novel T cell subsets within HCV-specific $CD8^+$ T cells.

Hypothesis 3: The Biomark single cell gene expression platform can be used to identify novel subsets within HCV-specific $CD8^+$ T cells.

3.3 Methods

Sample and multimer acquisition, antibodies, flow cytomtery staining, RT PCR and RNA extraction protocols are outlined in detail in chapter 2 methods.

3.3.1 Bioinformatics (ICA/NMF/GSEA)

Data was analyzed primarily by Gabriela Alexe (Haining laboratory, Broad Institute). Independent component analysis (ICA) using R and non-negative matrix factorization (NMF) were applied to gene expression profiles from HCV-specific $CD8^+$ T cells to cluster samples into biologically distinct molecular phenotypes. Gene set enrichment analysis (GSEA) was then used to associate these phenotypes with known biological pathways from pre-existing databases such as the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO).

3.3.2 Clinical data collection and correlation with subclass.

Clinical data was acquired from hospital databases in collaboration with the Lauer laboratory (MGH, Gastrointestinal Unit), (Appendix Tables 9.1-9.3). Data on estimated time from infection, age, sex, HLA type, IL28b polymorphisms, viral load and whether patients resolved HCV infection or, developed chronic disease was correlated with identified molecular subclasses of HCV-specific T cell response.

3.3.3 Functional network analysis

Functional network analysis using gene expression profiles from HCV-specific CD8$^+$ T cells was performed Gabriela Alexe (Haining laboratory, Broad Institute). GSEA was run on the Molecular Signatures Database (MSigDB) to identify leading edge genes i.e differentially expressed genes associated with identified HCV subclasses. Next, functional interactions associated with the leading edge genes were identified by projecting the leading edge onto a comprehensive collection of interaction networks available through GeneMANIA network integration platform. Further clustering of genes by applying molecular detection software (MCODE) and selection of candidate gene subsets was obtained through highest enrichment scores (data not included) with biological processes from GO database (Appendix Table 9.9).

3.3.4 Single cell sorting optimisation

Optimisation of single cell sorting was performed on the BD FACS ARIA II on E61 Jurkat cells stained with Hoechst blue (ATCC). A range of flow rates were used depending on cell concentration and a nozzle size of 100um was used to sort single cells onto flat bottom 60 well Terasaki plates (Bioexpress). Fluorescent cells (Hoechst blue dye) were detected and counted using the Nikon Eclipse inverted microscope and sorting efficiency calculated.

3.3.5 Optimisation of Biomark platform (Fluidigm)

Healthy human PBMC's obtained by density centrifugation were stained with antibodies CD4-FITC, CCR7-PE, CD8-PE Cy7, CD3-Pac Blue, CD45ra-APC (Biolegend) and aqua fluorescent HV500 live dead marker (Invitrogen). $CD8^+$ T cells were separated into memory populations based on CD45ra and CCR7 expression (effector memory RA positive, effector memory RA negative, central memory and naive) using the BD FACS ARIA II. Cells were sorted into a 96 well PCR plate (Eppendorf) containing 10.1 ul reverse transcriptase specific target amplification (RT-STA) master mix solution (Appendix Table 9.5). Samples were placed on dry ice and reverse transcribed using Bio rad C1000 touch thermal cycler (Appendix Table 9.6) 96 genes (Taqman) were tested including 4 control genes (Appendix Table 9.9). After cDNA has been formed, the 96 x 96 dynamic integrated fluidics chip was loaded with control line fluid and placed in the IFC controller to prime the control fluid into the chip. After priming was complete the chip was loaded with Taqman probe mix (Appendix Table 9.7) and cDNA sample mix (Appendix Table 9.8) into inlets either side. The chip is then run allowing 96 x 96 Taqman PCR reactions to be read in parallel using the Biomark real time PCR reader (Fluidigm).

3.4 Results

3.4.1 Identification of 2 novel subclasses of HCV-specific CD8$^+$ T cell response

Unpublished data from collaborative work performed by the Lauer and Haining labs on HCV-specific CD8$^+$ T cells is described here in figures 3.1 A, B and 3.2 A, B in order to accurately describe how my subsequent project work was initiated. HLA tetramers were used to isolate HCV-specific CD8$^+$ T cells (Figure 3.1 A) from 27 patients infected with HCV, 15 of which developed chronic and 12 who spontaneously cleared the virus. Clinical information including when samples were taken in relation to time of infection and specific multimers used can be found in Appendix Table 9.3. HCV samples obtained from Massachusetts General Hospital were screened with a range of multimers to identify HCV-specific responses (Appendix Table 9.4). Cell inputs were variable ranging from as few as 2000 up to as many as 20,000 antigen-specific cells. Good amounts of cDNA were yielded for microarray analysis ranging from 5.62ug to 9.64 which correlated with increasing cell input number (Appendix Figure 9.2).

In order to identify novel subtypes of T cell response, an algorithm by which unsupervised clustering of human gene expression samples from HCV-specific CD8$^+$ T cells was created. Independent Component analysis (ICA) is a computational method for separating multivariate signal into its additive components based in the mutual statistical independence of the non-Gaussian source signals. This approach assumes that the “signal” (i.e. gene expression profiles of different types of CD8$^+$ T cells) is composed of components (i.e.

groups of genes that we might consider representative of different biological processes) that themselves are different from one another. The algorithm detects modules of genes based on their non-Gaussianity (i.e. outlier status) compared to the distribution of all genes in the sample set as a whole. ICA returns a series of individual components (or modules of genes) that together comprise some or all of the overall "signal" in a complex dataset. Applying ICA to gene expression profiles from 27 samples of HCV-specific CD8$^+$ T cells identified components, or groups of genes creating a greater degree of substructure than would be expected by chance alone. A second clustering algorithm, non-negative matrix factorization (NMF) was applied to identify homogenous groups of samples. Therefore, ICA identified components that represented biological features that were heterogeneous across the dataset and NMF clustered samples into subtypes that were robustly similar in their expression level of component genes. This process identified 2 main subtypes of samples that can be visualized with a correlation matrix (Figure 3.1 B). Successful clustering of samples can also be achieved by hierarchical clustering techniques.

3.4.2 Clinical correlation with HCV subclasses

I next asked whether groups defined on the basis of distinct gene expression profiles corresponded to clinical characteristics of the subjects. I found no association between HCV subclass 1 or 2 with age, sex, HLA type and IL28b polymorphism. Viral load was difficult to correlate due to lack of accurate data. However, subclass 1 demonstrates a statistically significant difference in

the number of patients that resolve HCV infection when compared with subclass 2 (Figure 3.1 C). Thus these novel molecular subtypes of HCV-specific CD8+ T cell response correlate with clinical outcome; HCV subclass 1 more likely to resolve disease spontaneously and HCV 2 more likely develop chronic infection.

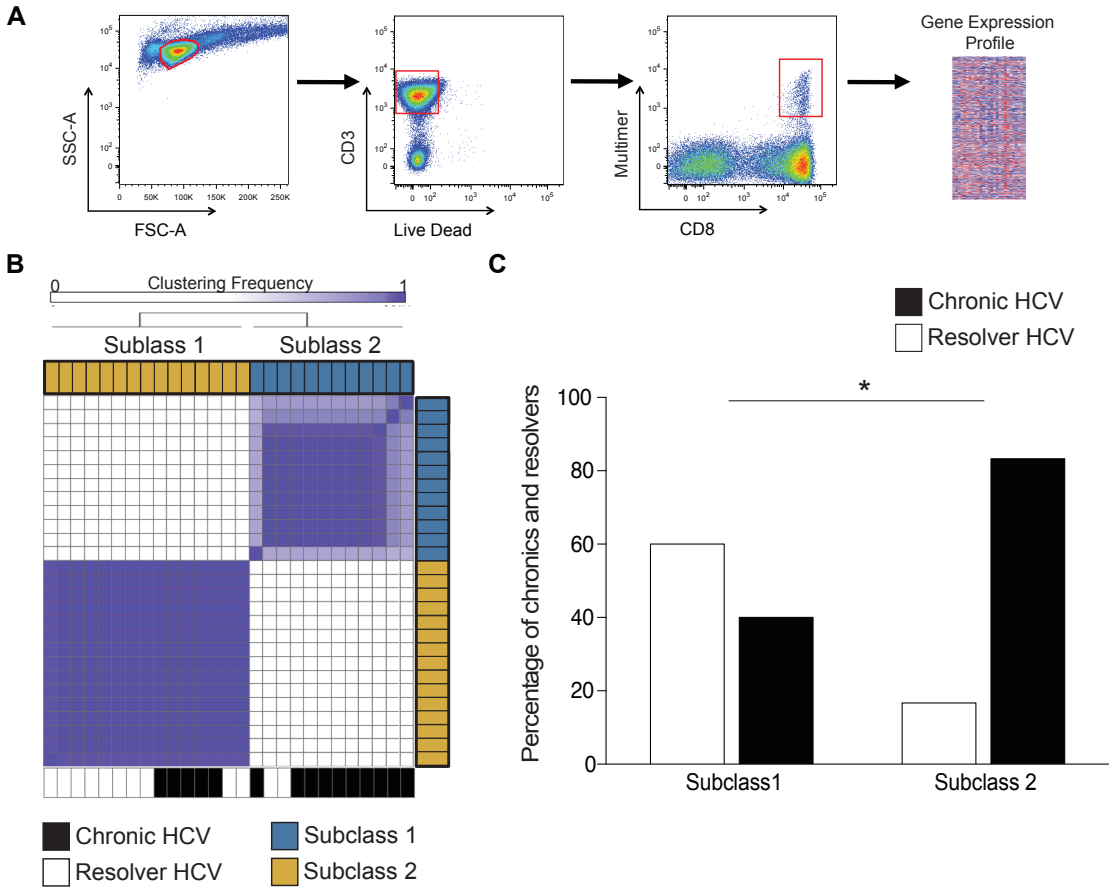


Figure 3.1. Unsupervised analysis of genome-wide transcriptional profiles of viral-specific CD8+ T cells identifies 2 biologically distinct subclasses in HCV infection. **(A)** Representative gating strategy for FACS cell sort to obtain HCV-specific CD8+ T cells for gene expression profiling using Affymetrix microarray platform. **(B)** Correlation matrix of expression profiles from HCV-specific CD8+ T cells from 27 HCV infected patients. Sample similarity is annotated with color from low (white) to high (purple). **(C)** Analysis of clinical parameters associated with sample donors revealed a statistically significant

difference in chronic or resolver status. Statistical significance was assessed by Fisher's test (C). *P<0.05.

3.4.3 GSEA and identification of gene clusters

If HCV-specific CD8$^+$ T cells can be clustered into biologically distinct subclasses of T cell response, a Gene Set Enrichment Analysis (GSEA) could be used to see if these subclasses demonstrated broad differences in the relative activity of genes corresponding to known biological processes. Interrogation of genes differentially expressed between subclasses HCV 1 and HCV 2 was performed with GSEA using a library of gene expression signatures from canonical pathway datasets such as Gene Ontology (GO) and Kyoto Encyclopaedia of Genes and Genomes (KEGG). Next, the leading edge GSEA genes differentially expressed between HCV 1 and HCV 2, which corresponded most with biological processes identified through KEGG and GO databases were isolated. Projection of the leading edge genes onto interaction networks (using GeneMANIA network integration platform) and further clustering (by Molecular Complex Detection algorithms - MCODE) identifies clusters of genes (4 clusters associated with HCV subclass 1 and 2 clusters associated with subclass 2) with functional interactions (Steps summarized in Figure 3.2 A). A total of 87 representative genes were chosen from clusters based on the highest enrichment scores with biological processes from the GO database (Appendix Table 9.9). These biological processes included metabolism, proliferation, DNA repair, Immune response, Inflammation and AP1 transcription factors (Figure 3.2 B). Projection of gene expression data of the 27 samples of HCV-specific CD8$^+$ T

cells within the space of these 87 genes (Appendix Table 9.9) is represented by a heat map demonstrating an enrichment of gene clusters A to D with subclass 1 samples (left) and enrichment of gene clusters E and F with subclass 2 samples (right) (Figure 3.2 B).

Using single cell analysis I test whether the expression of these gene clusters are driven by an over/underrepresentation of specific subpopulations of HCV-specific $CD8^+$ T cells or represent aggregate expression profiles of groups of similar HCV-specific $CD8^+$ T cells.

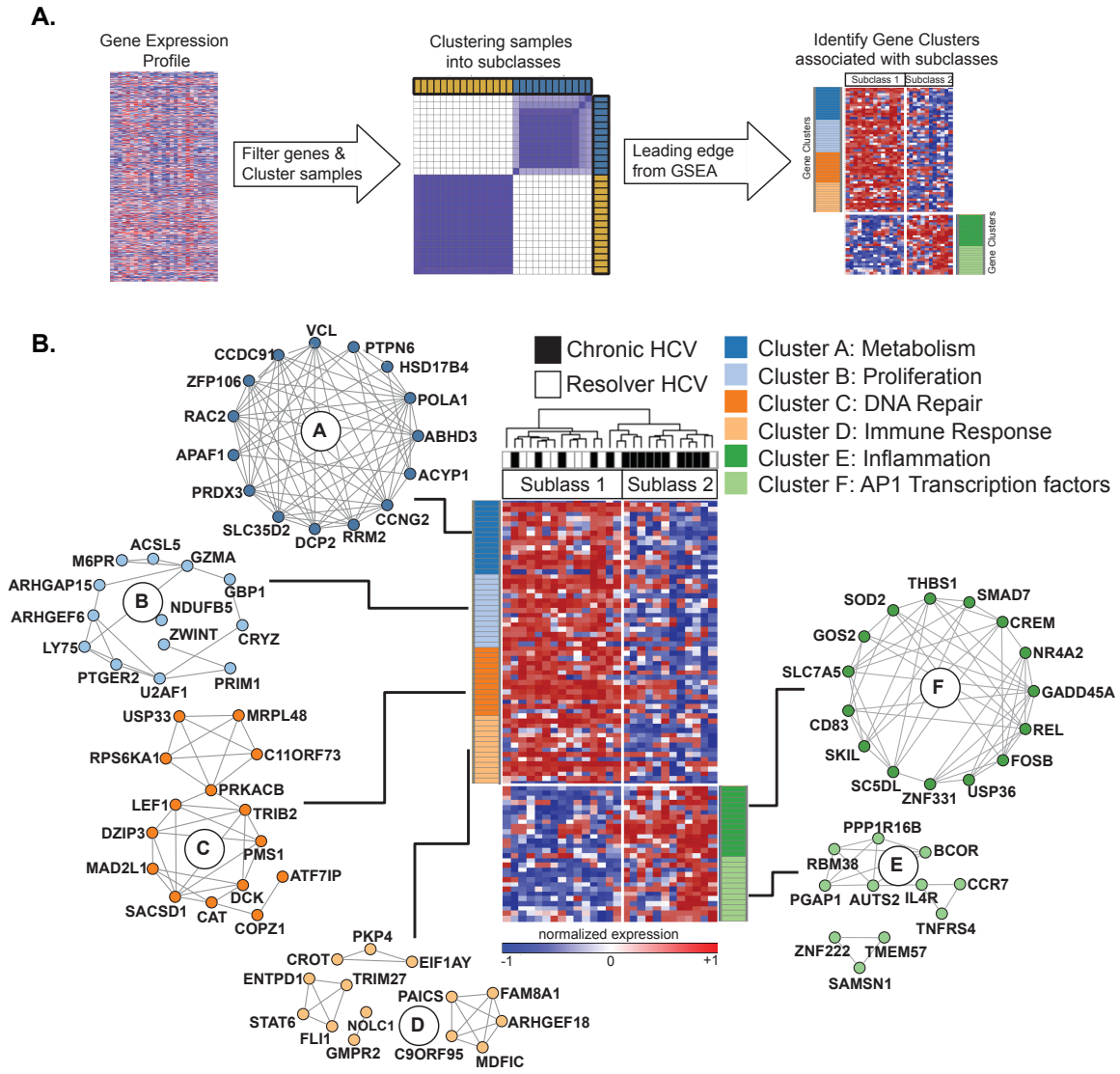


Figure 3.2. Identification of gene network clusters associated with 2 novel subclasses of CD8+ T cell response to HCV. (A) Flow diagram demonstrating how gene network clusters were identified from gene expression profiling of HCV-specific CD8+ T cells. Genes were initially filtered using Independent Components Analysis (ICA) followed by clustering of samples using Non-negative Matrix Factorization (NMF). Next Gene Set Enrichment Analysis (GSEA) of genes differentially expressed between the 2 subclasses was performed. The leading edge of the identified signatures were projected onto Genemania network interaction database allowing identification of gene clusters associated with novel subclasses of HCV-specific CD8+ T cell response. **(B)** Heat map

and representative diagram of 87 genes segregated into 6 clusters (4 associated with subclass 1 and 2 associated with subclass 2 samples) and their associated biological pathways identified through GSEA analysis.

3.4.4 Single cell sort optimisation

In order to identify single cell heterogeneity I first sought to develop a reliable, robust technique to sort individual primary CD8$^+$ T cells, which could be analysed on a single cell gene expression platform. I first tested my ability to sort individual cells by using the ARIA II flow cytometer at the Dana Farber flow cytometry core facility (Boston, USA) to sort E61 Jurkat T cells into a Terasaki microplate with a flat bottom. The Terasaki microplate has a conical well with a flat bottom, which allows for cells that hit the side of the well during the sorting process to be 'funneled' down to the bottom. The flat bottom also allows for good visualisation of the cell under microscopy. Therefore, I processed frozen E61 Jurkats and cultured in media for 3 days in order to allow cells to acclimatise and proliferate after being frozen. I then stained them with Hoechst blue nuclear stain (as described in methods page) and used flow cytometry (100um nozzle) to sort individual cells into each well of 3 Terasaki microplates (Figure 3.3 A). Within 30 minutes I counted cells that I could visually detect under fluorescent and light microscopy resulting in a sort efficiency of 96% (cells were counted only if seen in both fluorescent and light microscopy) (Figure 3.3 B).

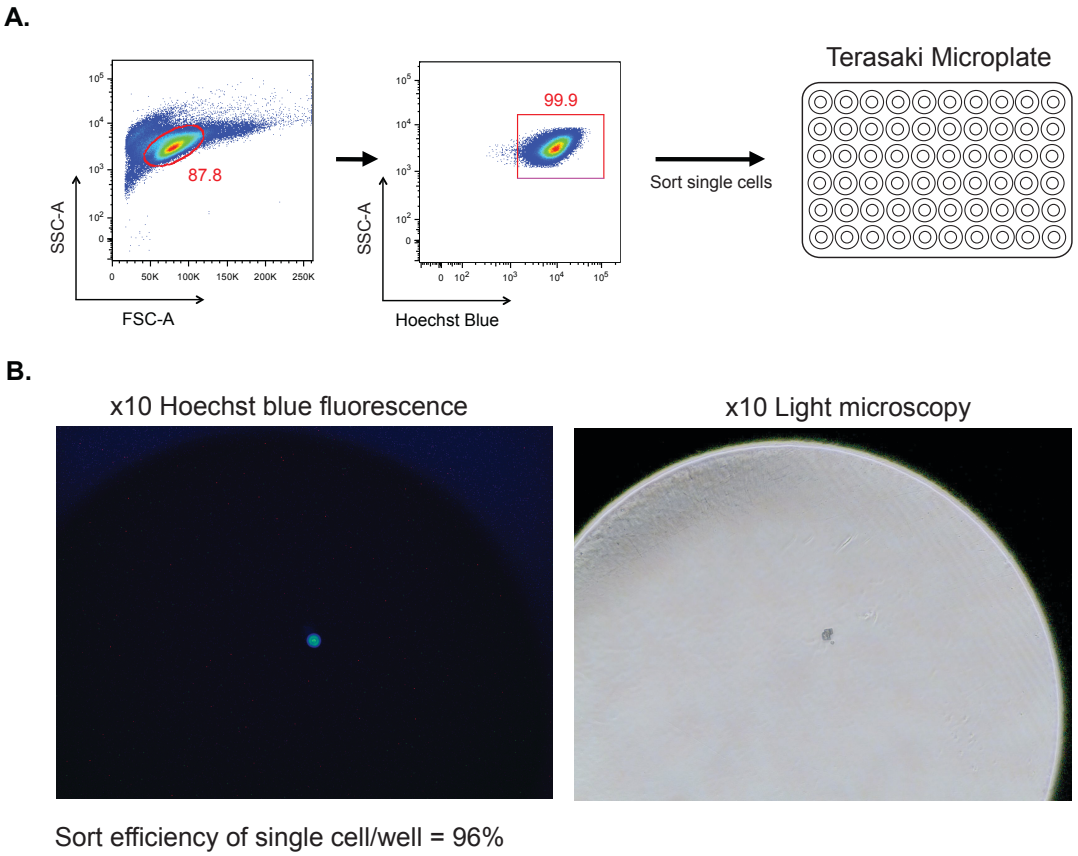


Figure 3.3. Single cell sorting optimisation with Hoechst blue stained E61 Jurkat T cell line using ARIA II flow cytometer. **(A)** Flow diagram demonstrating method for optimizing single cell sorting. E61 Jurkat cells stained with Hoechst blue and sorted using ARIA II flow cytometer (100um nozzle size) into single wells of a 60 well Terasaki microplate. **(B)** Demonstrates 1 representative well with single hoechst blue stained E61 Jurkat cell detected under fluorescent (left) and light microscopy (right). Sorting efficiency based on experiment done on 3 plates (total of 180 wells).

3.4.5 Testing Taqman gene probes by RNA dilution

Having identified the 87 genes associated with HCV subclass 1 and 2 (appendix table 9.9) I wished to look for single cell heterogeneity within HCV-

specific $CD8^+$ T cells and optimised a method to sort individual T cells, I next purchased the relevant gene probes and sought to test their efficacy on the Biomark single cell platform (designed by Fluidigm). Since I had 87 genes and 96 samples that can be used on the Biomark platform, I chose 4 control genes and 5 further genes of interest, which were mostly those associated with T cell inhibition including BATF and JunB as well as those of the purinergic pathway such as ADA, ADORA2A and NT5E (Appendix Table 9.9). I sorted 100,000 $CD8^+$ T cells from a healthy human donor by flow cytometry and processed in order to extract RNA and form cDNA (Figure 3.4 A). In order to validate the gene probes I performed a serial dilution of RNA samples prior to RT reaction, formation of cDNA and placement of samples into the IFC chip for gene expression analysis. This was done to determine what quantity of RNA is required for genes to be detected. Previous data suggests that a single cell contains about 1pg of RNA (139). After analysing samples I found that 52% of genes were detected in 50% of replicates at the 1pg RNA level (Figure 3.4 B). Further analysis of expression in control genes demonstrated a good linear relationship between cycle threshold expression value (CT) and RNA quantity of B2M and CD8A (Figure 3.4 C). In addition there was a poor degree of expression in the negative control gene CD4, which indicated a clean population of $CD8^+$ T cells sorted (Figure 3.4 C).

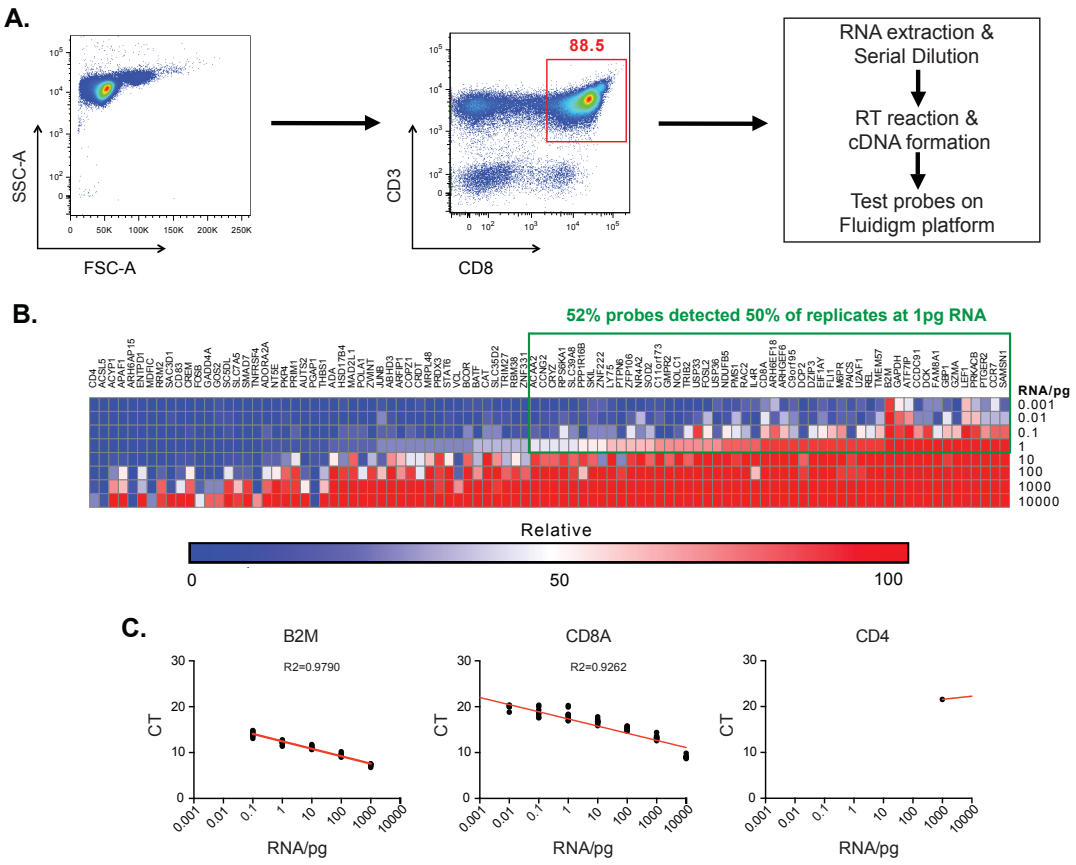


Figure 3.4. Use of healthy CD8+ T cells to test Taqman gene probes of 87 identified genes from network clusters. (A) Flow diagram and gating strategy showing sort of 100,000 healthy CD8+ T cells by flow cytometry and subsequent RNA extraction and serial dilution with RNAase free water into 8 concentrations and 11 technical replicates on an 96 well PCR plate prior to cDNA amplification and testing on fluidigm Integrated Fluidics Chip (IFC). **(B)** Analysis of 96 genes (87 genes from cluster, 4 control genes and 5 genes of interest) by gene e online program demonstrates 52% of probes detected 50% of replicates at 1pg level. **(C)** Analysis of positive control genes (B2M, CD8A) demonstrates a good correlation with concentration and cycle threshold value (CT). Negative control gene (CD4) demonstrated poor expression and poor correlation with concentration of RNA. R^2 values in (C) were calculated by linear regression analysis.

3.4.6 Optimisation of Polymerase Chain Reaction steps

Although a reasonable number of genes were able to be detected at the level of 1pg RNA, I sought to further optimise experimental conditions prior to analysing viral-specific $CD8^+$ T cells. I therefore tested single naive $CD8^+$ T cells from healthy donors to optimise steps during thermocycler incubation, namely cDNA annealing and extension times as well as the number of polymerase chain reaction (PCR) cycles performed. The rational for this was to obtain as much cDNA as possible from small, degraded quantities of RNA from cells that are of poor quality or have taken a long time to sort and have been out of culture media for a prolonged period of time. I therefore sorted naive $CD8^+$ T cells, using the antibody markers CCR7 and CD45RA, into wells of a 96 well PCR plate containing the RT STA mastermix, with only GAPDH and CCR7 gene probes (Figure 3.5 A). I then processed samples at various thermocycler settings and then performed conventional quantitative PCR to detect GAPDH and CCR7 levels to identify optimal conditions. My results demonstrate that elongation time of 4 minutes only makes a significant difference at 24 PCR cycles (Figure 3.5 B, Figure 3.5 C). Therefore, I chose to perform future experiments at these settings.

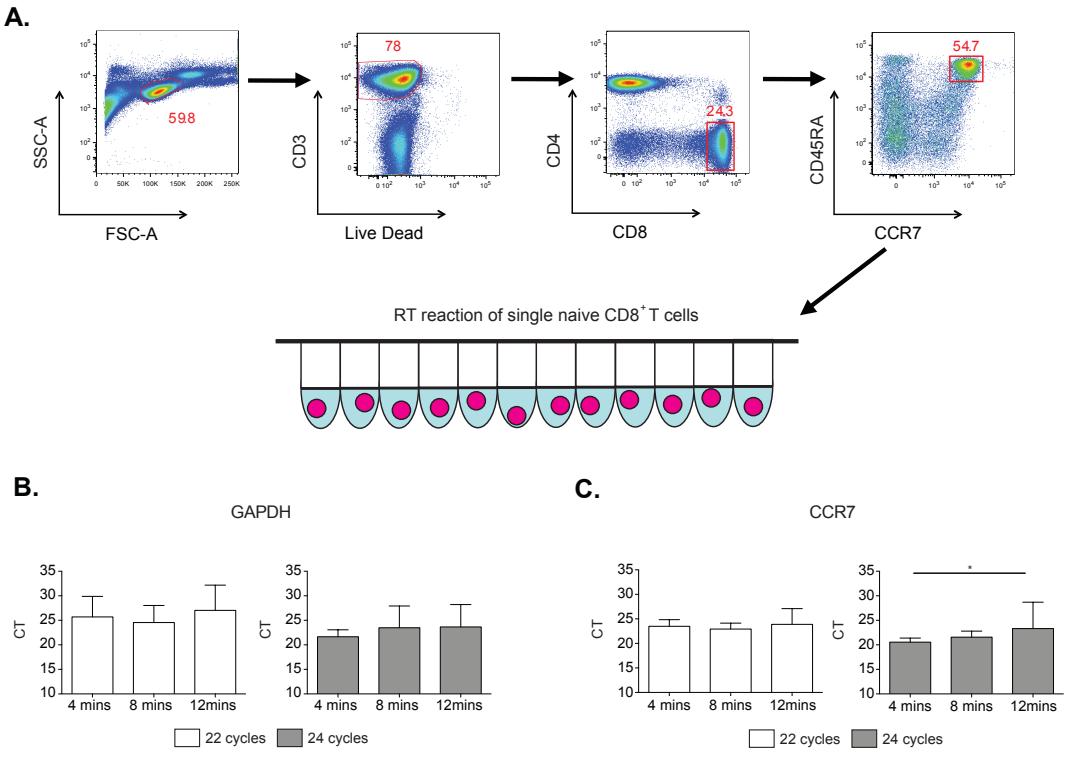


Figure 3.5. Optimisation of Polymerase Chain Reaction steps. (A) Gating strategy demonstrating sort of single naïve $CD8^+$ T cells from a healthy donor into a 96 well PCR plate containing RT-STA master mix containing GAPDH and CCR7 probes ready for placing into thermocycler for RT reaction and cDNA amplification of GAPDH and CCR7 genes only. **(B)** Conventional qPCR results of GAPDH expression showing no significant difference in expression between different elongation times but lower CT values at 24 cycles for elongation step. (C) Conventional qPCR results of CCR7 expression showing significant difference in expression between different 4 and 12 minutes elongation time only and lower CT values at 24 cycles for elongation step. Statistical significance was assessed by paired Student's t-test (B, C). **P* <0.05.

3.4.7 Validation of single cell gene expression data

Having optimised many of the parameters required to perform single cell analysis successfully, I next tested naive $CD8^+$ T cells from a healthy human donor on the Biomark platform rather than viral-specific $CD8^+$ T cells so as to not waste valuable samples in case further optimisation steps were necessary. In addition, this would provide data that could help identify gene signatures that are different between naive and effector cell pool and virus-specific responses. I sorted individual naive and $EMRA^-$ $CD8^+$ T cells from a healthy donor into a 96 well PCR plate containing the appropriate master mix as previously described. After performing the RT reaction, annealing and extension steps I then placed aliquots of cDNA sample mix and assay mix containing gene probes into the Fluidigm IFC, which was inserted into the Biomark platform (Figure 3.6 A). My results demonstrate that of the replicates that were detected, the majority followed the expected expression patterns of control genes B2M, CD4, CCR7 and Granzyme A (Figures 3.6 B – E). B2M was highly expressed in all detectable replicates (97%), which is to be expected since B2M is routinely used as a control gene that is well recognized to be highly expressed in lymphocytes (Figure 3.6 B). CD4 was not expressed in any replicates indicating this was a pure population of $CD8^+$ T cells (Figure 3.6 C). CCR7 was highly expressed in the majority of naive cells (78%) but in none of the $EMRA^-$ cells, which was also expected as $EMRA^-$ $CD8^+$ T cells were gated on $CCR7^-$ (Figure 3.6 D). Finally, Granzyme A was expressed in 46% of $EMRA^-$ cells but in 0% of naive cells, which would supports Granzyme A as a cytolytic enzyme recognised to be

expressed in effector $CD8^+$ T cells (Figure 3.6 E). Furthermore, analysis of wells containing 100 cells compared with 1 cell demonstrate that the aggregate expression of 100, individual wells correlated highly with the expression of a well containing a pool of 100 cells. This was an important control to test for the possibility of a false negative detection of transcript. This was calculated by converting the CT values of 100 cells into linear format then dividing by 100 before reconverting to Log_2 CT expression values. All Log_2 CT values are expressed as 40-CT so that high expression values are expressed as increasing rather than decreasing values.

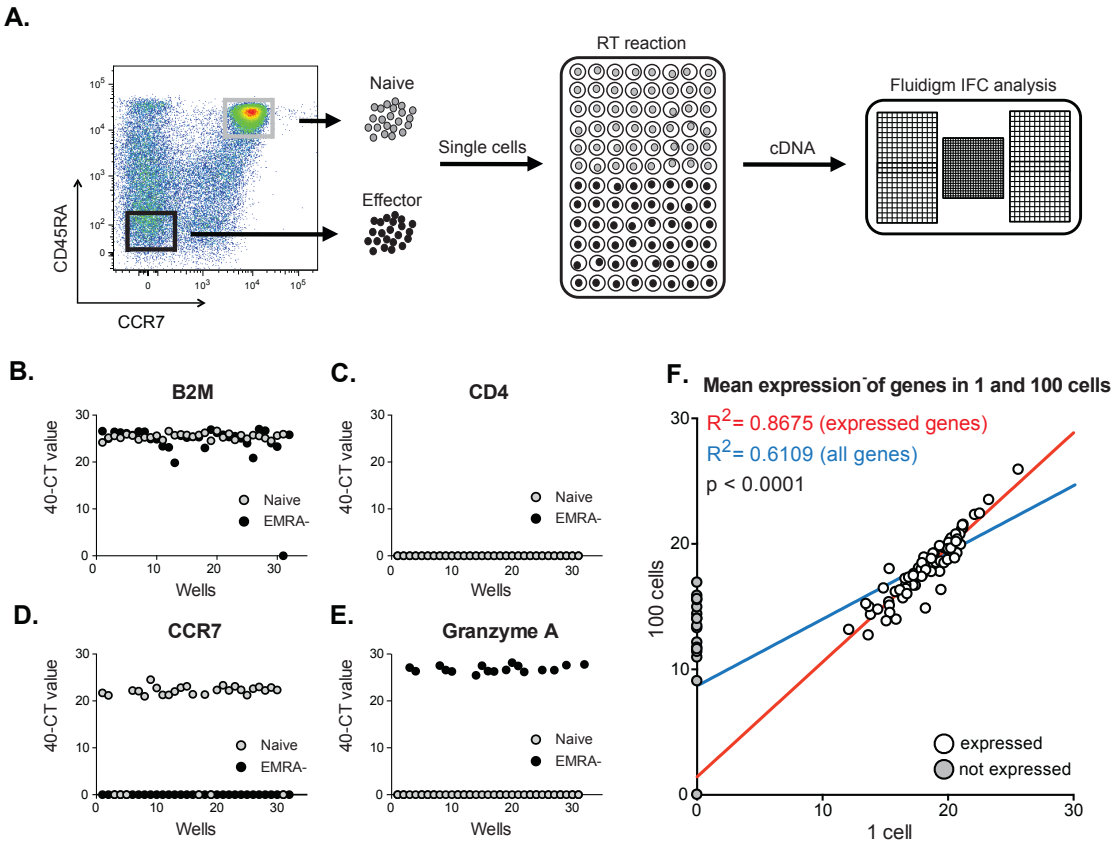


Figure 3.6. Testing genes on fluidigm panel using healthy effector memory and naïve CD8+ T cells. (A) Diagram representing major steps involved in single cell analysis of CD8+ T cells using Biomark platform (Fluidigm). Individual Naive and Effector Memory CD8+ T cells were sorted using flow cytometry into a 96 well PCR plate containing 10.1ul of RT-STA master mix with 96 gene probes. After RT reaction and 24 elongation cycles, samples were pipetted into the Fluidigm IFC chip and analysed on the Biomark platform. **(B-E)** Majority of replicates expressed high levels of the positive control B2M (B) and no replicates expressed the negative control CD4 (C). CCR7 was expressed in the majority of Naive cells (D) and Granzyme A in the majority of effector cells (E). **(F)** Analysis of mean expression values of genes in single cell replicates compared with 100 cell replicates demonstrates a statistically significant positive correlation. Statistical significance was assessed by linear regression (F).

3.4.8 Single cell analysis of viral-specific CD8$^+$ T cells

After successfully obtaining gene expression data on individual naive and effector CD8$^+$ T cells, I next sorted a range of viral-specific CD8$^+$ T cells for processing and gene expression analysis within the space of the described 96 gene panel (Figure 3.7 A). Of the 424 single cells sorted (187 HCV, 40 CMV, 40 EBV, 93 YFV-specific, 32 EMRA$^-$, 32 naive CD8$^+$ T cells) and with a potential 36,888 data points in the space of 87 subclass-specific genes, only 11% (4071 data points) were detected (Figure 3.8 A). Only 3.5% of data points were expressed in the 5 additional genes of interest chosen (ADA, ADORA2A, BATF, JUNB, NT5E) and therefore these genes were excluded from analysis with subclasses (Appendix figure 9.3). 42 cells did not express the B2M control and were therefore also excluded from analysis. However, after filtering genes and samples, a supervised analysis of data according to cell type and subclass-specific genes demonstrates a possible increased expression in subclass 1 genes in the CMV, EBV, YFV-specific as well as EMRA$^-$ and naive CD8$^+$ T cell populations. Whereas subclass 2 genes appear enriched in the HCV-specific CD8$^+$ T cells regardless of chronic or resolver status of infection (Figure 3.8 A).

Unsupervised analysis by hierarchical clustering of samples and genes highlights 3 gene signatures that appear enriched in specific cell populations (Figure 3.9 A, B). Signature 1 appears enriched in HCV-specific CD8$^+$ T cells and consists of 7 genes (SLC39A8, SKIL, SMASN1, CREM, NR4A2, FOSL2, THBS1) predominately from clusters E (inflammation) and F (AP1 transcription factors). Signature 2 appears enriched in CMV, EBV, YFV-specific as well as EMRA$^-$

populations and is made of 8 genes (CAT, U2AF1, ATF7IP, GZMA, PTGER2, PRKACB, M6PR, DZIP3) from clusters B (proliferation) and C (DNA repair). Signature 3 appears enriched in naive $CD8^+$ T cells and contains 10 genes (DCK, USP33, SOD2, CCDC91, ARHGEF6, EIF1AY, ARH6EF18, LEF1, CCR7, FLI1) from clusters B, C and D (Immune response). Unsupervised analysis of single cells by hierarchical clustering segregates samples into 2 groups, HCV-specific $CD8^+$ T cells and all other (non-HCV infected) samples (Figure 3.10 A). Clustering of samples and/or genes did not identify novel subsets of HCV-specific $CD8^+$ T cell response within the space of subclass specific gene clusters. However, upon further literature research on the 87 subclass-specific genes, one gene was found to be worth exploring further owing to its involvement in modulating T cell responses in cancer and chronic infection models. ENTPD1 (CD39) an ectoenzyme, which hydrolyses extracellular ATP to adenosine via the CD73 pathway has recently been discovered as being highly expressed on CD4 T regulatory cells and is upregulated during infection and malignancy (140, 141). It is being increasingly implicated as a modulator of T cell inhibition but has not been investigated in the context of chronic HCV infection on $CD8^+$ T cells. Therefore, in Chapter 4 I demonstrate how I next tested the phenotype of CD39 in chronic HCV and HIV infection to determine if it is a potential marker of T cell exhaustion.

A.

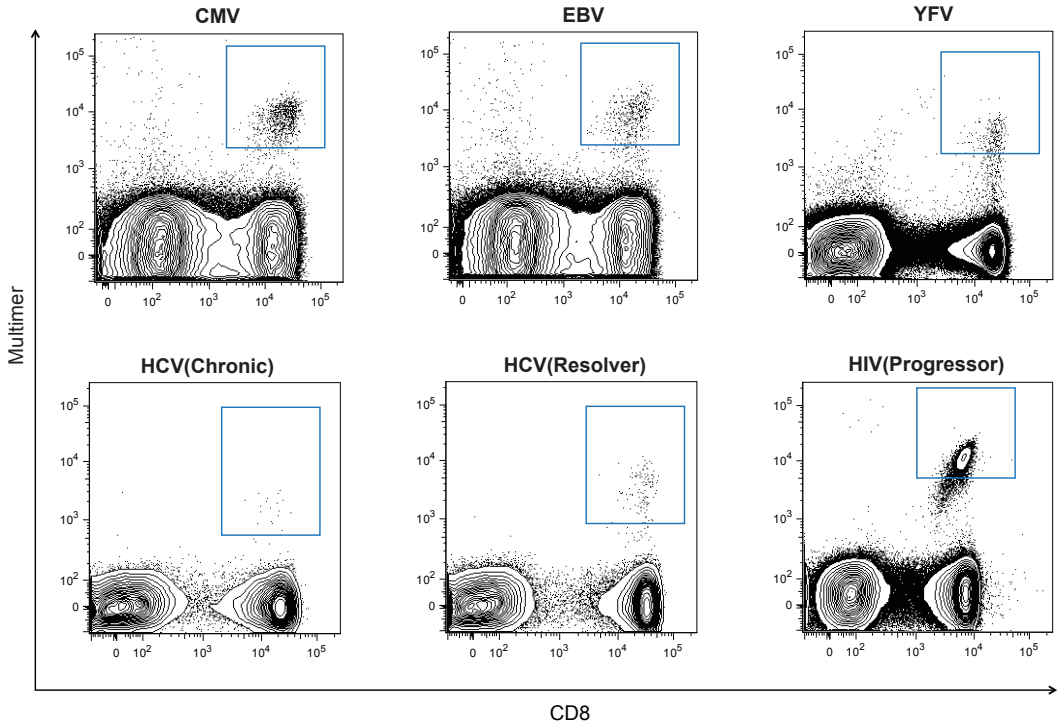


Figure 3.7. Analysis of viral-specific CD8+ T cells from chronic and latent infections. (A) Flow cytometry analysis of single cell sorts of a range of viral-specific CD8+ T cells (CMV, EBV, YFV, HCV, HIV) by ARIA II flow cytometer. Samples were processed and analyzed on the Biomark platform as previously described (Figure 3.6 A).

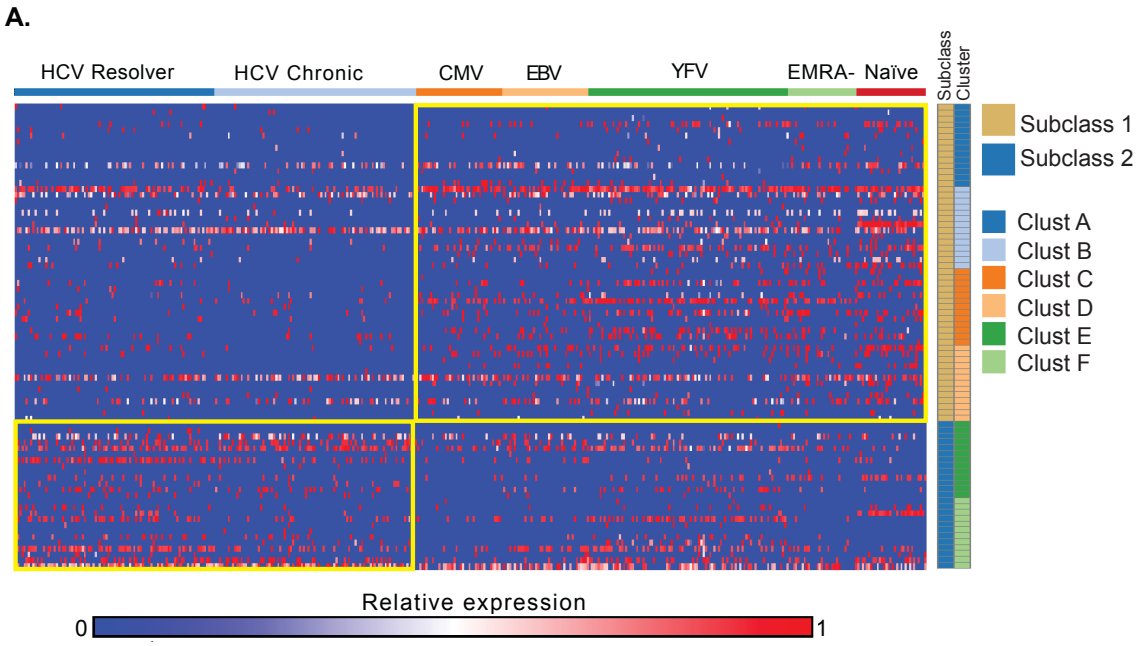
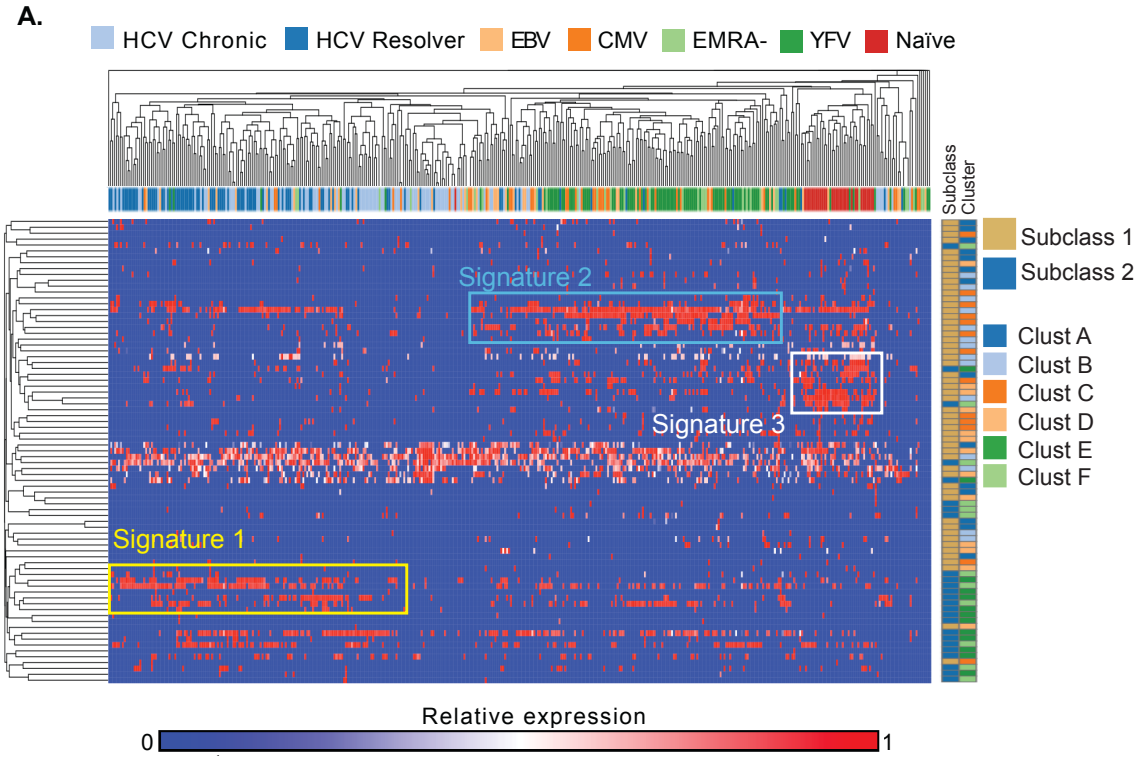


Figure 3.8 Supervised analysis of single cell gene expression data. (A) Heatmap demonstrating relative expression of 79 of the 87 (8 genes were not expressed in any cell) subclass associated genes expressed in 424 single cells (187 HCV, 40 CMV, 40 EBV, 93 YFV-specific, 32 $EMRA^-$, 32 naive $CD8^+$ T cells) analysed on Biomark platform. Each column represents and individual cell and each row an individual gene, coloured to indicate normalised expression. Low relative expression indicated by blue and high relative expression in red. Supervised analysis based on cell type and gene cluster identifies 2 broad gene signatures (A and B). Signature A is upregulated in HCV-specific $CD8^+$ T cells regardless of chronic or resolver status and signature B is upregulated in CMV, EBV, YFV-specific as well as EMRA- and Naive $CD8^+$ T cells.



B.

Signature 1	Signature 2	Signature 3
Clusters E,F	Clusters B,C	Clusters B,C,D
SLC39A8	CAT	DCK,
SKIL	U2AF1	USP33
SMASN1	ATF7IP	SOD2
CREM	GZMA	CCDC91
NR4A2	PTGER2	ARHGEF6
FOSL2	PRKACB	EIF1AY
THBS1	M6PR	ARH6EF18
	DZIP3	LEF1
		CCR7
		FLI1

Figure 3.9. Unsupervised analysis of single cell gene expression data. (A) Heatmap demonstrating 3 sets of gene signatures identified after hierarchical clustering of gene expression values and subclass associated gene clusters. Each column represents and individual cell and each row an individual gene, coloured to indicate normalised expression as in Figure 3.8. Signature 1 appears enriched in HCV-specific CD8+ T cells, Signature 2 in YFV, CMV, EBV-specific CD8+ T cells , as well as EMRA- CD8+ T cells. Signature 3 appears mostly enriched in Naive CD8+ T cells. (B) Table with list of individual genes in signatures 1 to 3 with predominate subclass associated clusters.

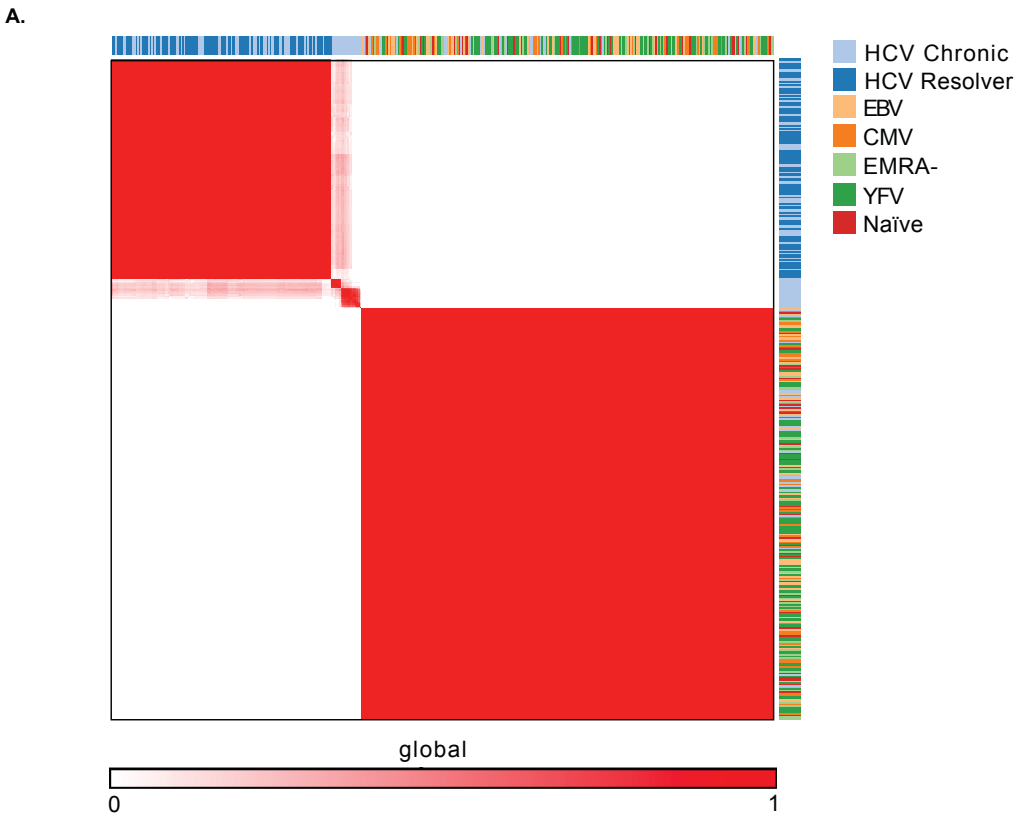


Figure 3.10. Unsupervised analysis of single cell gene expression data identifies 2 broad molecular phenotypes. (A) Consensus hierarchical clustering of expression profiles from 424 viral-specific, EMRA$^-$ and naive CD8$^+$ T cells. Samples segregate into 2 phenotypes, HCV-specific (top left) and all other CD8$^+$ T cells (bottom right). Clustering is based on the top 10% of genes by variance across the dataset. Sample similarity (1-Pearson correlation coefficient) is annotated with color from low (white) to high (red). Novel phenoytpes within the HCV-specific pool were not successfully identified.

3.5 Discussion

Findings:

1. **Unsupervised analysis of genome-wide expression profiles of HCV-specific $CD8^+$ T cells identifies two distinct subclasses of samples.**
2. **HCV subclasses 1 and 2 correlate with clinical outcome such as developing chronic infection or resolving virus.**
3. **Functional network analysis is able to identify gene clusters associated with biological pathways that segregate HCV subclass 1 and 2 samples through hierarchical clustering.**
4. **Single cell analysis of HCV-specific $CD8^+$ T cells using subclass-specific gene clusters was unable to identify novel subsets of HCV-specific $CD8^+$ T cells, but could differentiate from other viral-specific responses such as CMV, EBV and YFV.**

3.5.1 Subclass discovery and correlation with clinical outcome

Understanding the host immune response is key to development of prophylactic and therapeutic strategies to infectious pathogens. Our lack of understanding of the fundamental pathways modulating robust and protective immunologic responses is highlighted by our inability to generate broadly protective vaccines against HCV, HIV, mycobacterium tuberculosis, malaria and influenza virus. Moreover, we lack sufficient surrogate markers of these processes to guide vaccine development. Flow cytometry has been an integral tool in the discovery of fundamental cell processes involved with innate and

adaptive immunity (142). However, over the last 15 years technological advances have enabled immunologists to use genome-wide expression profiles to serve as much broader surrogates for biological phenotypes of cell populations within the immune system (143, 144).

Unsupervised analysis of genome-wide expression profiles of HCV-specific $CD8^+$ T cells allows discovery of 2 novel, biologically distinct subclasses that correlate with clinical outcome. Samples from HCV subclass 1 displayed a higher proportion of samples from donors that resolve infection and subclass 2 samples contained a greater number of donors that develop chronic infection. This algorithm for novel T cell discovery is advantageous to traditional analytic approaches in several ways. Firstly, most investigation for phenotypes of cells in the adaptive immune response are driven by the rational selection of known biologically relevant characteristics, for instance those that develop chronic HCV infection and those that completely resolve virus. The disadvantage of this knowledge-driven approach is that the result captured is only as good as the body of knowledge used. This study uses a data-driven approach whereby correlates between patterns of gene expression and phenotype of interest were sought. This unbiased, unsupervised technique assumes no a priori assumption of what genes are likely to be induced and opens the door to novel biologic discovery. Secondly, although gene-by-gene analysis of genomic transcriptional profiles has played an important discovery tool in the past, such as CD127 (106) and PD1 (107) , it understates the interaction of many genes that may be important in the pathway. Instead, these results were obtained through identifying

coordinated patterns of gene expression highlighting important interactions between hundreds of genes. Thirdly, whereas traditional analytical approaches for gene expression data focus on a handful of ranked genes that are differentially expressed, integration of GSEA helps identify groups of biologically meaningful genes using catalogues of annotated gene sets. Such techniques allow for novel molecular phenotypes to be associated to a biological context such as HCV subclass 1 associated with included metabolism, proliferation, DNA repair, Immune response and HCV subclass 2 with Inflammation and AP1 transcription factors.

However, there are several disadvantages with the above approaches. Firstly, GSEA are based in gene sets from GO and KEGG where annotations are assigned by knowledge from the assigner, rather than direct experimentation. Therefore, the fraction of genes that are incorrectly annotated is unknown. Secondly, due to the wide variation of genetic diversity in human samples, a large cohort of patients must be used to identify significant subclasses with a fair representation of phenotypes. This sample set only contained 28 donors, which were predominately patients that develop chronic infection and taken from different time points. Therefore depending on the time point that samples are taken results can be biased by factors such as viral load, which will be high in chronic disease. Thirdly, genes likely to belong to the same biological process may not be co-regulated at the transcriptional level and may be overshadowed by other members of the gene set. Such limitations are resulting in an increasing

trend to gather experimentally derived datasets and may greatly influence the discovery novel gene biomarkers.

3.5.2 Functional Network Analysis

The ability to identify molecular phenotypes by gene markers depends on 2 critical factors. The first is the minimum number and network of genes required to distinguish the phenotypes and the second is the availability of technologies that can accommodate enough parameters. The first question can be answered by using a novel functional network analysis to choose candidate genes (comprising 87 genes) representative of the major biological processes associated with HCV subclass 1 and 2. I tested the effectiveness of these genes to be able to segregate samples into 2 subclasses creating a heatmap of the original gene expression data. The results demonstrated that this technique was able to capture 2 broad molecular phenotypes. However, Further testing on a larger number of samples and additional datasets would be desirable to confirm the robustness of these gene clusters to segregate samples. In answer to the second question, it was clear that flow cytometry did not possess the number of parameters required to demonstrate the HCV molecular phenotypes. However, I demonstrate how recent nano-fluidic technologies (fluidigm) can allow for up to 96 gene-markers to be analysed on a single cell (119, 120).

3.5.3 Single cell analysis of HCV-specific CD8$^+$ T cells

The analysis of single naive and effector CD8$^+$ T cells demonstrated some encouraging results during initial experiments to optimise the Biomark single cell

analysis platform. There was a good correlation in the mean expression values between samples containing single cells and those containing 100 cells. In addition genes such as Granzyme A and CCR7 demonstrated both linear and bimodal expression patterns, which is consistent with what one would expect from cells expressing transcript during intervals rather than continuously.

Supervised and unsupervised analyses of HCV-specific $CD8^+$ T cells within the space of the identified 87 gene clusters did not identify novel subsets of HCV-specific $CD8^+$ T cells. However, HCV-specific $CD8^+$ T cells were differentiated from other viral-specific cells and naive cells when all analysed together by hierarchical clustering algorithms. Samples generally separated into HCV-specific, latent virus-specific (CMV, EBV, YFV, $EMRA^-$) and naive $CD8^+$ T cells. This could have been due to a number of reasons. Firstly, experimental conditions may not have been optimised enough as only 11% of data points were detected. Therefore, there may not have been sufficient data to be able to capture novel clusters of samples. I suspect that use of fresh cells would help in increasing the yield of gene expression data as a considerable degree of cell death occurs when freezing and thawing lymphocytes. Unfortunately, being able to identify HCV-specific $CD8^+$ T cell responses from donors through tetramer screening prior to processing for single cell gene expression analysis would be a logistically difficult and costly task. Secondly, the choice of genes used for single cell analysis may not have been sufficient in number or type to necessarily identify novel subsets with the pool of HCV-specific $CD8^+$ T cells of these donors. This highlights the third reason that it would be ideal to perform a single cell

analysis on the exact same sample as that used for gene expression analysis by affymetrix microarrays. Unfortunately, this was not possible due to lack of donor samples. Fourthly, cell numbers may have been insufficient. The analysis of 187 HCV-specific CD8$^+$ T cells from 2 HCV donors, may well be too small number to represent the whole HCV-specific CD8$^+$ T cell pool in these donors. Unfortunately, there are no known methods to be able to predict the required minimum number of cells or genes required to perform single cell analyses that will yield these types of results. However, study of the current literature on the 87 subclass-specific genes revealed ENTPD1 (CD39) to be of particular interest owing to its involvement in modulating T cell responses in cancer and chronic infection models. ENTPD1 (CD39) an ectoenzyme, which hydrolyses extracellular ATP to adenosine via the CD73 pathway has recently been discovered as being highly expressed on CD4 T regulatory cells and is upregulated during infection and malignancy (140, 141). It is being increasingly implicated as a modulator of T cell inhibition but has not been investigated in the context of chronic HCV infection on CD8$^+$ T cells. Therefore, in Chapter 4 I demonstrate how I next tested the phenotype of CD39 by flow cytometry in chronic HCV and HIV infection to determine if it is a potential marker of T cell exhaustion and if its expression changed with clinical outcome.

3.5.4 Identification of novel markers of T cell dysfunction?

The identification of signature-based based biomarkers of clinical phenotype has been demonstrated in several studies (145, 146). Functional

network analysis identifies a group of markers associated with an HCV CD8$^+$ T cell response correlated with resolution of HCV infection in subclass 1 samples. I demonstrate in subsequent experiments (see chapter 4) that one such marker known as CD39 (ENTPD1) is highly expressed in HCV-specific CD8$^+$ T cells and correlates with clinical parameters. CD39 comes from the family of ectoenzymes know as ENTPDases, and is proposed to have anti-inflammatory action by producing cAMP from the hydrolysis of ATP, and subsequent CD73 mediated conversion of cAMP to adenosine which inhibits T cells (147-150).

3.5.5 Conclusion

Unsupervised analysis of genome-wide expression profiles of HCV-specific CD8$^+$ T cells can be used to identify novel subclasses of T cell response in HCV infection that can be correlated with clinical outcome. Employing GSEA and functional network analyses, clusters of genes that are subclass-specific can be used to identify novel markers of CD8$^+$ T cells dysfunction, such as CD39 (ENTPD1). Single cell analyses using the Biomark (Fluidigm) platform could potentially identify novel subsets of HCV-specific CD8$^+$ T cells using subclass-specific markers but is limited unless key experimental conditions are optimised sufficiently such as the appropriate number and type of genes selected and number of cells tested.

CHAPTER 4

CD39 expression on HCV-specific CD8$^+$ T cells

4.1 Introduction

After acute viral infection, naive antigen-specific $CD8^+$ T cells undergo activation and rapidly expand in number over the course of a few weeks. These cells differentiate further to gain effector functions such as the ability to produce the antiviral cytokines IFNγ, TNFα as well as granzymes and perforins (55). During this time activation markers are transiently upregulated but once the peak of expansion has been reached, approximately 95% of these effector cells die leaving a contracted population of memory $CD8^+$ T cells that confers long term immunity by retaining the ability to re-expand and gain effector function on subsequent exposure with the same pathogen (151).

In contrast, the persistence of viral antigen during chronic viral infection such as Hepatitis C virus can produce a significantly altered differentiation state of virus-specific $CD8^+$ T cells that display impaired proliferative and functional capabilities that can lead to eventual deletion (152). These traits are collectively termed T cell exhaustion and well described in both animal and human models of chronic viral infection (72, 111, 152). Approximately 30% of people infected with HCV are able to spontaneously clear the virus, whilst the remaining 70% go on to develop chronic infection and T cell exhaustion. The key steps that lead to progressive chronic infection rather than viral control and effective immunity are poorly understood. Identifying reversible mechanisms of T cell exhaustion is therefore a major goal in the treatment of chronic viral infections such as HCV.

The function of exhausted $CD8^+$ T cells is regulated by cell intrinsic and extrinsic mechanisms. Cell intrinsic mechanisms include the prolonged or high

expression of multiple inhibitory receptors such as PD-1, 2B4, Lag-3, TIM-3 and CD160 which is well described in multiple animal and human models of chronic infection including: lymphocytic choriomeningitis virus 1 (LCMV), simian immunodeficiency virus (SIV), human immunodeficiency virus 1 (HIV), Hepatitis B virus (HBV) and Hepatitis C virus (HCV) (70, 107, 111, 153-156). This imbalance between cellular stimulation and inhibition by the above receptors is a key target for potential immunotherapies for chronic infection.

Expression of PD-1 appears to be a particularly important feature of exhausted $CD8^+$ T cells, as the majority of exhausted cells in mouse models of chronic infection express this receptor, and blockade of the PD-1:PD-L1 axis can restore the function of exhausted $CD8^+$ T cells in humans and mouse models (107). However, in humans, many inhibitory receptors can also be expressed by a large fraction of fully functional memory $CD8^+$ T cells (157). PD-1, for instance, can be expressed by as much as 60% of memory $CD8^+$ T cells in healthy individuals, making it challenging to use this marker to enumerate exhausted $CD8^+$ T cells in humans, particularly when the antigen-specificity of potentially exhausted $CD8^+$ T cells is not known (157). Several studies have suggested that co-expression of multiple inhibitory receptors on HCV-specific $CD8^+$ T cells identifies further states of T cell exhaustion than PD-1 can do alone (77, 81). Therefore, the discovery of novel markers of $CD8^+$ T cell exhaustion in chronic HCV infection is a key area for further investigation.

In chapter 3 I identified CD39 as a gene that is differentially expressed in 2 novel subclasses of HCV-specific $CD8^+$ T cells. In this chapter I present data

demonstrating CD39 as a novel marker of T cell exhaustion in human chronic viral infection. CD39 (ENTPD1) is an ectonucleotidase originally identified as an activation marker on human lymphocytes, but has subsequently been shown to be a hallmark feature of regulatory T cells (101, 158, 159). It exists as an integral membrane protein that phosphohydrolyses extracellular ATP and to a lesser extent ADP into adenosine monophosphate (160). Human CD39 consists of 510-amino acid proteins characterised by two transmembrane domains, a small cytoplasmic domain comprising the NH2- and COOH-terminal segments, and a large extracellular hydrophobic domain with five apyrase conserved regions, which are essential for the catabolic activity of the enzyme (see Figure 4.0) (160).

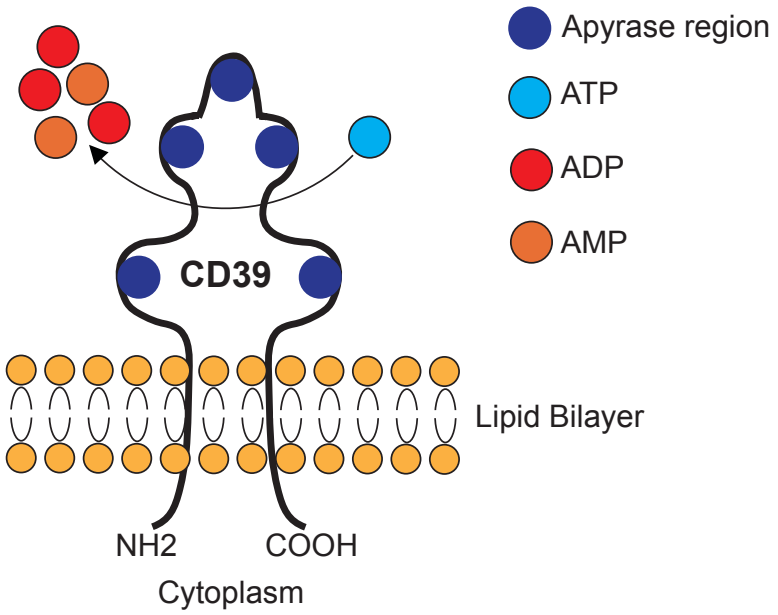


Figure 4.0: Structure of CD39 on T cell.

CD39 becomes metabolically active upon its localisation on the cell surface. Once it's enzymatic activity is complete, adenosine monophosphate is then processed into adenosine by CD73, an ecto-5'-nucleotidase. Adenosine is a potent immunoregulator that binds to A2A receptors expressed by lymphocytes causing accumulation of intracellular cAMP, preventing T cell activation and NK cytotoxicity (147, 161). Loss of CD39 in Tregs markedly impairs their ability to suppress T cell activation, suggesting that the juxtacrine activity of CD39 may be an important regulatory mechanism at the site of inflamed tissues (101). However, $CD8^+$ T cells have generally been reported to be $CD39^-$ (102, 162) and the expression of this marker on exhausted T cells has not been examined.

In this study, we demonstrate that, in contrast to $CD8^+$ T cells from healthy donors, antigen-specific $CD8^+$ T cells responding to chronic viral infection in humans express high levels of biochemically active CD39. $CD39^+$ $CD8^+$ T cells highly co-express PD-1 and together correlate with clinical outcome in human chronic viral infections. Thus, CD39 represents a novel pathological marker of exhausted $CD8^+$ T cells in chronic HCV and HIV infection.

4.2 Aims/hypotheses

Specific Aim 1 (SA1): Use flow cytometry to phenotype CD39 on bulk and viral-specific $CD8^+$ T cells in human chronic and latent viral infections.

Hypothesis 1: CD39 expression on $CD8^+$ T cells is a marker of T cell exhaustion.

Specific Aim 2 (SA2): Quantify CD39 activity on $CD8^+$ T cells in human HCV infection.

Hypothesis 2: CD39 is metabolically active on the surface of $CD8^+$ T cells in HCV infection.

Specific Aim 3 (SA3): Correlate CD39 expression on $CD8^+$ T cells with clinical parameters.

Hypothesis 3: CD39 expression changes at different stages of disease progression in HCV.

4.3 Methods

4.3.1 Human Subjects

Healthy human donors were recruited at the Kraft Family Blood Donor Center, Dana-Farber Cancer Institute (DFCI) with written informed consent following approval by Partners Institutional Review Board.

All human subjects with HCV infection were recruited at the Massachusetts General Hospital Gastrointestinal Unit and the Department of Surgery with written consent in accordance with the IRB approved study: “Cell mediated immunity in Hepatitis C virus infection”; Protocol # 1999- P-004983/54; MGH Legacy #: 90-7246 (appendix Figure 9.1). HCV chronics (n = 27) were defined by positive anti-HCV antibody and detectable viral load. HCV resolvers (n = 14) were defined by positive anti-HCV antibody but an undetectable viral load for at least 6 months. All HCV patients were treatment naive and obtained between 5.9 and 237.3 weeks post infection. HCV RNA levels were determined using the VERSANT HCV RNA 3.0 (bDNA 3.0) assay (Bayer Diagnostics).

All HIV infected cohorts were recruited after written informed consent from the Ragon Institute, Massachusetts General Hospital, Boston USA and the Peter Medawar Building for Pathogen Research, Oxford, UK where ethics approval was given by the Oxford Research Committee. HIV controllers included elite controllers (n = 5) defined as having HIV RNA below the level of detection (<75 copies per ml); viremic controllers (n = 7) with HIV RNA levels < 2,000 copies per ml. HIV chronic progressors (n = 28) were defined as having > 2,000 copies per

ml. All subjects were off therapy. Viral load in chronic infection was measured using the Roche Amplicor version 1.5 assay.

PBMCs were obtained by density centrifugation at 2000 rpm at 25°C for 30 minutes with RPMI (Life technologies) and Ficoll-Plaque (GE Healthcare). Red blood cells (RBC) then lysed (RBC lysis buffer, ebioscience) for 10 minutes at room temperature before being cryopreserved in 4% formaldehyde and 10% Fetal Bovine Serumor further processed.

4.3.2 HLA Class I Tetramers

Major histocompatibility complex (MHC) class I HIV Gag-specific tetramers were generated by the Goulder laboratory as previously described (121) (One-pot, mix-and-read peptide-MHC tetramers); and Alter laboratory as described (121). HCV-specific MHC class I multimers (tetramers and pentamers) were provided by the Lauer laboratory (Proimmune). CMV- and EBV-specific MHC class I dextramers conjugated with FITC and APC were purchased from Immudex. YFV-specific tetramers were conjugated with APC and donated by the Ahmed laboratory (Emory University, Atlanta, Georgia, USA) (Appendix Table 9.4).

4.3.3 Antibodies and flow cytometry

The following fluorochrome-conjugated antibodies were used: anti-huCD8α, anti-huCD4, anti-huCD3, anti-huCD39, anti-huPD1, anti-huCD25, anti-huCCR7, anti-huCD45RA, anti-huCD160, anti-huCD161. Peripheral blood

mononuclear cells (PBMCs) were thawed in R10 medium (RPMI medium, 10% fetal calf serum [FCS], 1% penicillin/streptomycin, 1% glutamine); rested for 1 hour at 37°C in 5% CO2. 5 million PBMC cells were re-suspended in 50µl of buffer (PBS, 0.1% FBS) and stained with the LIVE/DEAD Fixable Aqua Dead Cell Stain Kit (Invitrogen) for exclusion of dead cells prior to staining with MHC Class I multimer (tagged with APC or PE) for 10 minutes at room temperature to identify antigen-specific cells. Surface staining was performed for 25 minutes in the dark at 4°C and cells were fixed using 4% paraformaldehyde for 20 minutes prior to analysis. Intracellular staining was performed following surface stains and fixed and permeabilized using the FoxP3/Transcription Factor Staining Buffer Set (eBioscience). Cells were sorted by BD FACS ARIA II and all other analyses were performed on BD LSR II and BD LSR Fortessa flow cytometers equipped with FACSDiva v6.1. Gates were set using Full Minus One (FMO) controls. Data were analysed using Flowjo software (Treestar).

4.3.4 Cell culture and stimulation of primary T lymphocytes

Primary human $CD8^+$ T cells were isolated by CD8 negative selection MACS kits (Miltenyi) and were cultured in RPMI-1640 medium supplemented with 10% FBS, 10 mM HEPES, 50U/ml of penicillin, 50µg/ml of streptomycin and 50µM β-mercaptoethanol. Naive cells were stimulated for 5 days with soluble plate-bound anti-CD3 (4µg/ml; 2C11; BD Pharmingen) and anti-CD28 (4µg/ml; 37.51; BD Pharmingen) for the generation of $CD8^+CD39^+$ T cells that were used for optimisation of ATP hydrolysis assay experiment.

4.3.5 Cell culture of CEM cells

CEM cells were cultured in RPMI-1640 medium supplemented with 10% FBS, 10mM HEPES, 50U/ml of penicillin, 50µg/ml of streptomycin and 50µM β-mercaptoethanol at 37°C in 5% CO2 prior to use in ATP hydrolysis experiments.

4.3.6 Luciferase assay for ATP analysis

Measurement of ATP hydrolysis was performed on healthy primary $CD3^+$ T cells by addition of exogenous ATP (sigma) at a concentration of 100µM to cell cultures enriched and depleted for CD39 expression. Cells were incubated in cell culture media (RPMI, 10%FBS, 100iu/ml penicillin, 100µg/ml streptomycin) for 1 hour after addition of a CD39 inhibitor Sodium polyoxotungstate 1 (POM 1). Supernatant was removed and mixed at a 1:1 ratio with cell-titer glo (sigma) and incubated for 10 minutes. Luminescence of cell titer glo bound to ATP was measured at 520nm on fluostar Omega luminometer.

4.3.7 HPLC analysis of ATP levels

The concentrations of ATP hydrolysed by effector $CD39^+CD8^+$ from samples of chronic HCV infected subjects (n = 6) were assessed by high performance liquid chromatography (HPLC) as previously described (123). Briefly, 10,000 $CD39^+$ $CD8^+$ T cells were sorted and placed on ice to minimize ATP production by cells. 20mM of ATP was added and incubated for 1 hour at 37°C in 5% CO2 to allow for cellular activity to increase and CD39-mediated ATP hydrolysis to occur. Samples were then placed in an ice bath for 10 minutes to

halt enzymatic activity, collected, and centrifuged for 10 minutes at 2,000rpm and 0°C. Cells were discarded and supernatant centrifuged again to remove remaining cells (5,000rpm, 5 minutes, 0°C). The resulting RPMI samples (160ml) were treated with 10ml of an 8M perchloric acid solution (Sigma-Aldrich, St. Louis, MO) and centrifuged at 13,000rpm for 10 minutes at 0°C to precipitate proteins. In order to neutralize the pH of the resulting solutions and to remove lipids, supernatants (80ml) were treated with 4 M K_2HPO_4 (8ml) and tri-N-octylamine (50ml). These samples were mixed with 50ml of 1,1,2-trichloro-trifluoroethane and centrifuged (13,000rpm, 10 minutes, 0°C) and this last lipid extraction step was repeated once. The resulting supernatants were subjected to the following procedure to generate fluorescent etheno-adenine products: 150ml supernatant (or nucleotide standard solution) was incubated at 72°C for 30 minutes with 250mM Na_2HPO_4 (20ml) and 1 M chloroacetaldehyde (30ml; Sigma-Aldrich) in a final reaction volume of 200ml resulting in the formation of 1,N6-etheno derivatives as previously described (123). Samples were placed on ice, alkalinized with 0.5M NH_4HCO_3 (50ml), filtered with 1 ml syringe and 0.45 mM filter and analyzed using a Waters HPLC system (Waters, Milford, MA, USA) and Supelcosil 3μM LC-18T reverse phase column (Sigma) consisting of a gradient system described previously, a Waters autosampler, and a Waters 474 fluorescence detector (124). Empower2 software was used for the analysis of data and all samples were compared with water and ATP standard controls as well as a sample with no cells to determine background degradation of ATP.

4.4 Results

4.4.1 CD39 is expressed by CD8$^+$ T cells responding to chronic infection

Adenosine is a potent regulator of T cell function, but it is not known whether purinergic pathways are involved in the regulation of the immune response to chronic viral infection (163). To address this question, we began by surveying the expression of CD39, an important regulator of adenosine production, in CD8$^+$ T cells from healthy adult subjects without chronic viral infection. Consistent with previous reports (102) we found that only a small fraction (mean 7%) of CD8$^+$ T cells in healthy individuals expressed CD39 (Figure 4.1B). We next focused on CD39 expression by antigen-specific CD8$^+$ T cells specific for latent viruses in healthy subjects, and found that only a low frequency of CMV- or EBV-specific CD8$^+$ T cells expressed CD39 (Figure 4.1A, 4.1B)(mean 3% and 7% respectively).

To analyse CD39 expression by T cells specific for chronic, rather than latent, viruses, we analysed peripheral blood samples of individuals with either HCV or HIV infection. We measured CD39 expression in 34 subjects with acute HCV infections (13 with acute resolving infection and 21 with chronically evolving infection), and in 40 subjects with HIV infection (28 chronic progressors and 12 elite controllers of infection) (Appendix Table 9.1-9.2). Analysis of CD8$^+$ T cells specific for HIV or HCV showed that both virus-specific populations showed a significantly higher frequency of cells (58% of HCV-specific CD8$^+$ T cells; 31% of HIV-specific CD8$^+$ T cells) expressing CD39 than CD8$^+$ T cells specific for EBV or CMV, or in total CD8$^+$ T cell populations from healthy individuals (Figure 4.1A,

4.1B). Furthermore, virus-specific $CD8^+$ T cells from HCV-infected subjects showed significantly greater CD39 expression (as assessed by analysis of mean fluorescence intensity (MFI)) than did those from HIV-infected subjects.

The frequency of CD39-expressing cells in the virus-specific population was significantly higher than in the total $CD8^+$ T cell population in subjects with both chronic infections (Figure 4.1C, 4.1D). However the fraction of $CD8^+$ T cells expressing CD39 in the $CD8^+$ T cell compartment of individuals with chronic infection was slightly increased compared to healthy controls (Figure 4.1E), consistent with the presence of other, unmeasured virus-specific $CD8^+$ T cells in the tetramer$^-$ fraction of $CD8^+$ T cells. Thus CD39 is expressed infrequently by $CD8^+$ T cells in healthy donors, but marks a large fraction of pathogen-specific $CD8^+$ T cells in patients with chronic infection.

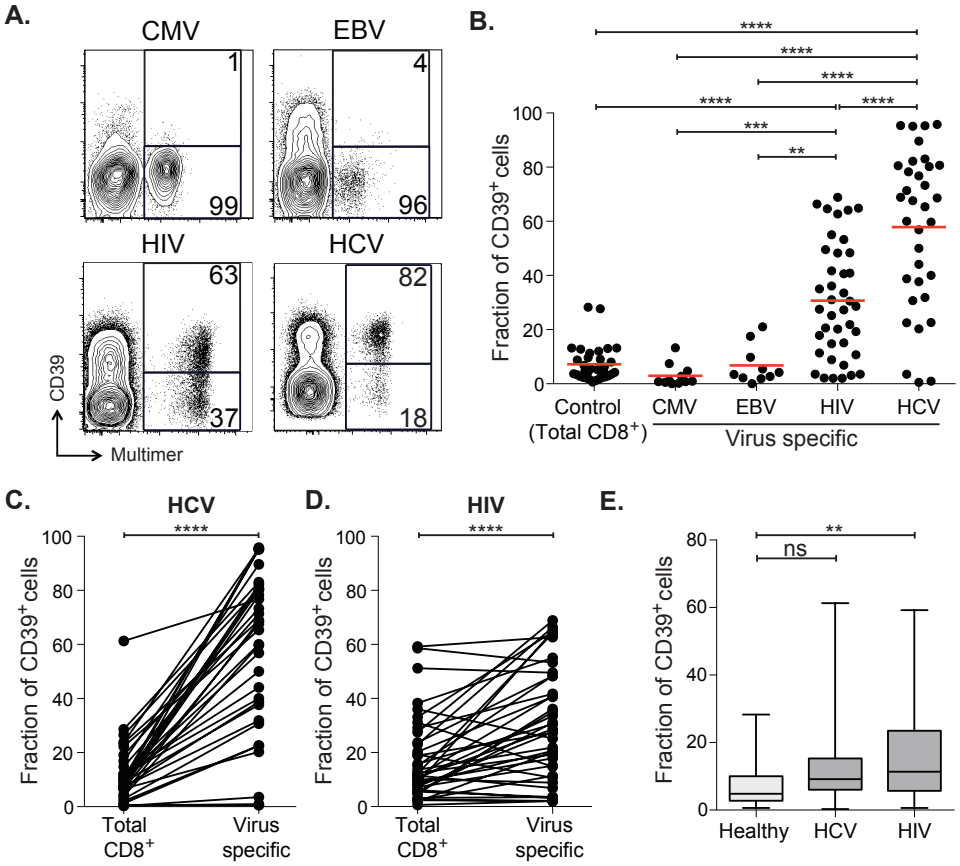


Figure 4.1. CD39 is highly expressed by viral-specific CD8+ T cells in chronic viral infection. (A) Flow cytometry analysis of CD39 protein expression on chronic (HIV, HCV) and latent (CMV, EBV) virus-specific CD8+ T cells**. (B)** Percentage of CD39 expression in human viral-specific CD8+ T cells (HIV-, HCV-, CMV- and EBV-specific) were compared with total CD8+ T cells form healthy donors using flow. In (A) and (B), 11 CMV and 10 EBV samples were tested. **(C, D)** Comparison of CD39 protein expression in total CD8+ T cells and virus-specific CD8+ T cells from patients with HCV (C) and HIV (D) infections. **(E)** Fraction of CD39+ on total CD8+ T cells in healthy, HIV and HCV infected donors. Error bars represent SEM. Statistical significance was assessed by one-way ANOVA (A), paired (C, D), or unpaired (E) Student's t-test. Ns=not significant, **P<0.01, ***P<0.001, ****P<0.0001

4.4.2 CD39 is predominately expressed by effector and central memory $CD8^+$ T cells

In order to further phenotype the pattern of expression of $CD39^+$ T cells in the bulk $CD8^+$ T cell population we used the antibody markers CD45RA and CCR7 to delineate four defined CD8 T cell memory differentiation states including effector memory ($EMRA^+$ and $EMRA^-$), central memory (CM) and naive subsets (49) (Figure 4.2A). We then assessed the levels of CD39 expression in each of these four $CD8^+$ T cell populations in healthy as well as HIV and HCV infected samples (Figure 4.2B). We found that the highest levels of CD39 expression were found in the $EMRA^-$ (means of $EMRA^-$, 9.7%, 8.1%, 9.4% in Healthy, HIV and HCV respectively) and CM (means of CM- 9.7%, 11.3%, 14.8% in Healthy, HIV and HCV respectively) populations across all three groups of samples with CD39 expression being highest in the CM subsets in the chronic infection samples (Figure 4.2B). By far the lowest level of CD39 expression was found in the naive $CD8^+$ T cell population across all samples (means of naive, 1.6%, 3.1%, 1.2% in Healthy, HIV and HCV respectively) followed by the $EMRA^+$ group (means of $EMRA^+$, 3.3%, 6.1%, 1.4% in Healthy, HIV and HCV respectively) (Figure 4.2B). These results would be generally consistent with the finding that CD39 expression is reduced on naive compared with antigen-experienced $CD8^+$ T cells. Therefore CD39 is more highly expressed in the effector memory and central memory groups in chronic viral infection.

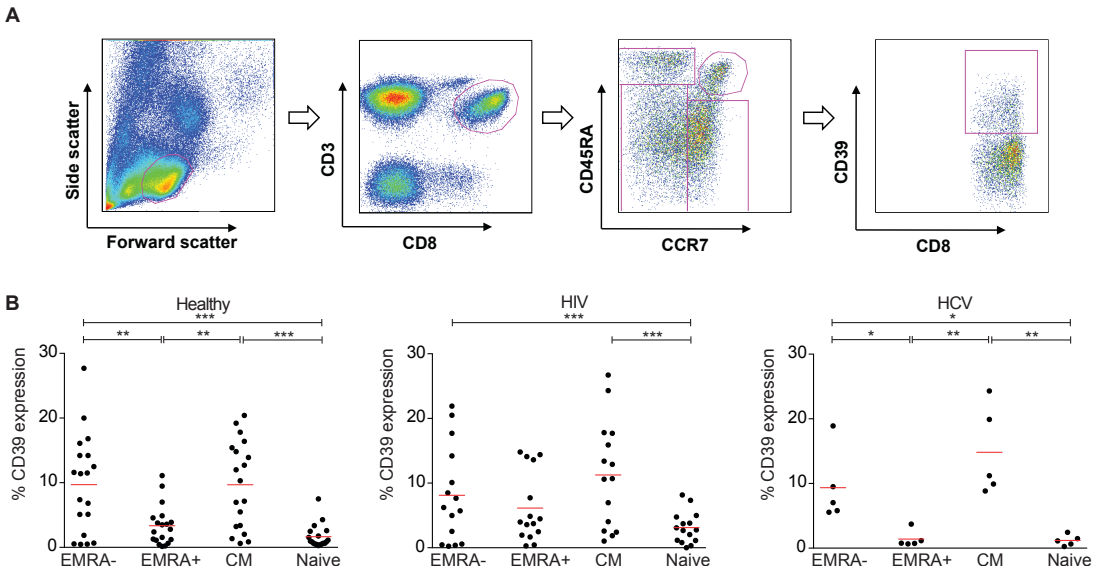


Figure 4.2. Flow cytometry analysis of human CD8+ T cells demonstrates that CD39 is most highly expressed in effector memory and central memory compartments. PBMC samples from healthy, HIV and HCV infected donors were analysed to determine the distribution of $CD8^+CD39^+$ T cells within effector memory, central memory and naive compartments. (A) Flow cytometry gating strategy of PBMCs from an HCV infected donor demonstrating CCR7 and CD45RA markers effectively identifying effector memory, central memory and naive $CD8^+$ T cells for CD39 expression analysis. (B) Scatter plot demonstrating that $EMRA^-$ and CM populations of human $CD8^+$ T cells have statistically significant higher levels of CD39 expression than naive and $EMRA^+$ $CD8^+$ T Cells. . Statistical significance was assessed by one-way ANOVA (B). *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001.

4.4.3 Detection of ATP hydrolysis on $CD3^+$ T cells

CD39 belongs to the family of ENTPDases, which hydrolyse extracellular adenosine triphosphate (ATP) into adenosine diphosphate (ADP) and cyclic adenosine monophosphate, which is subsequently dephosphorylated to adenosine by CD73. The functional role of CD39 in the context of HCV infection is not clear so we aimed to first confirm CD39 expressing cells' functional role as a hydrolyser of ATP in healthy human T cells. In order to achieve this I began by enriching T cells with CD39 expression in order to increase the total number of CD39 expressing cells so as to reach the lower limit of detection of an ATP hydrolysis assay. After stimulating PBMC cells from healthy human donors with soluble anti-CD3/28 (Figure 4.3A) I used flow cytometry to confirm that the percentage of CD39 expressing cells was significantly increased in activated $CD3^+$ T cells so that I may obtain a higher yield of cells for ATP hydrolysis quantification (Figure 4.3B). Next by positive selection with magnetic columns to isolate $CD3^+CD39^+$ and negative selection to isolate $CD3^+CD39^-$ T cells I tested the levels of ATP hydrolysis from these T cell populations using 200,000 cells (Figure 4.3C). A positive control CD39 expressing T cell line called CEM cells was also tested at different concentrations of the well established CD39 inhibitor POM 1 (140). Using a luciferase-based assay I found that ATP hydrolysis was detectable from $CD3^+CD39^+$ T cells (ranging from 98% to 55%) and that this was significantly greater than $CD3^+CD39^-$ T cells (29.2% to 2.2%)(Figure 4.3C). In addition the percentage of ATP hydrolysis dropped as CD39 inhibitor concentration is increased in both the $CD3^+CD39^+$ T cells and $CD39^+$ CEM cell

line. Therefore, ATP hydrolysis is quantifiable at relatively high cell numbers by use of the cell titer glo luciferase based assay. However, this does not allow the accurate quantification of CD39 activity in small groups cells such as CD39$^+$CD8$^+$ T cells in chronic HCV infection.

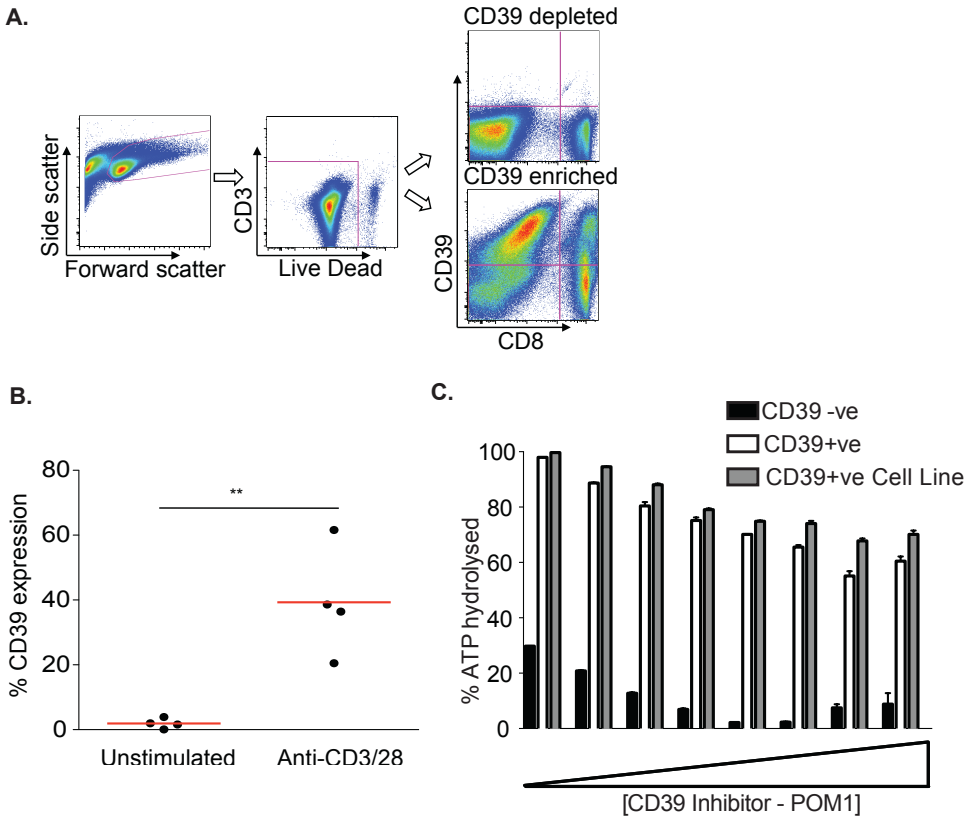


Figure 4.3. Functional activity of CD39 on human CD8$^+$ T cells was assessed by a luciferase assay on TCR stimulated CD3$^+$ T cells from healthy human donors. (A) Flow cytometry analysis demonstrating good enrichment of CD39$^+$ T cells from positive magnetic selection of CD3$^+$CD39$^+$ T cells. **(B)** Scatter plot demonstrating statistically significant upregulation of CD39 in healthy human CD8$^+$ T cells after 5 days TCR stimulation by soluble anti-CD3/28 antibodies. **(C)** Cell titer glo assay demonstrating greater ATP hydrolysis (and its inhibition) by CD3$^+$CD39$^+$ T cells and positive control CEM T cell line expressing CD39 than CD3$^+$CD39$^-$T cells. Statistical significance was assessed by paired Student's t-test (B). **P<0.01.

4.4.4 $CD4^+$ T reg positive control population

In order to explore a method of more accurately quantifying CD39 activity on small numbers of T cells, I next tested primary T cells to look for a more comparable positive control than a pre-existing T cell line. Human $CD4^+$ T regulatory cells (Tregs) are well recognised to be enriched with CD39 and the definition of this population of cells has ranged from $CD4^+CD25^+$, $CD4^+CD25^{High}$ or, $CD4^+CD25^{High}FOXP3^+$. It has been reported that 90% of $Foxp3^+$ T regs are $CD39^+$ and that $CD4^+$CD39 T cells largely overlap with $CD4^+CD25^+FOXP3^+$ T regs (100, 164). I therefore used flow cytometry analysis and tested the enrichment of CD39 in Tregs using the flow cytometry markers CD3, CD4 and CD25 (Figure 4.4A). I found that there was a distinct population of (165) $CD4^+CD25^+CD39^+$ T cells that represents the CD39 enriched Treg group of cells. This population of $CD4^+CD25^+$ T cells is significantly enriched with CD39 when compared with $CD4^+CD25^-$ T cells (mean 64% $CD25^+$, mean 36% $CD25^-$) (Figure 4.4B). Therefore, primary $CD4^+$ Tregs are significantly enriched with CD39 and are a viable alternative positive control for the detection of CD39 activity in T cell subgroups.

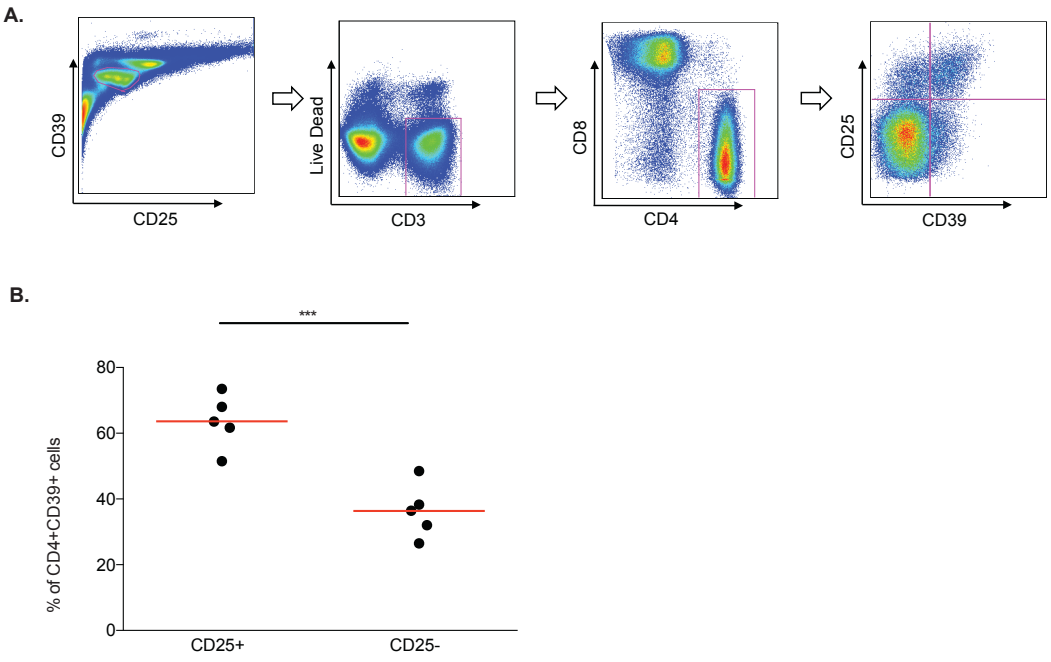


Figure 4.4. Enrichment of CD39 in CD4+ T regulatory cells (CD4+ Tregs). Quantification of CD39 activity on CD8+ T cells requires positive control measurements and adequate cell numbers for assay detection. CD4+ Tregs are known to express CD39 and the marker CD25 enriches for Treg cells, which express adequate numbers of CD39+ T cells for analysis by HPLC. **(A)** Flow cytometry analysis demonstrating subset of CD4+CD25+ T regulatory cells enriched for CD39 with adequate cell numbers for detection by HPLC. **(B)** Scatter plot demonstrating statistically significant increase in CD39 expression with CD4+CD25+ T cells compared with CD25-CD4+ T cells. Statistical significance was assessed by paired Student's t-test (B). ***P<0.001.

4.4.5 Optimisation of ATP hydrolysis quantification by rpHPLC

Reverse phase high performance liquid chromatography (rpHPLC) has been used to detect low quantities of ATP, ADP, cAMP and adenosine in the extracellular environment of many human cell types (165, 166). In the context of lymphocytes, most of these studies have concentrated on $CD4^+$ Treg cells, which are in relative abundance in CD39 expression compared with other T cell populations (167). No studies have analysed adenyl purines in the extracellular environment of human $CD8^+CD39^+$ T cells in the context of HCV infection. Nor have ATP levels been detected from lymphocytes as low as 10,000 in number. In order to accurately quantify ATP hydrolysis in small populations $CD8^+CD39^+$ T cells I first sorted TCR stimulated $CD8^+CD25^+$ T cells enriched with CD39 so that I may have enough cells for a HPLC assay (Figure 4.5A). I began by testing the cellular threshold for detection of ATP hydrolysis in TCR stimulated $CD8^+CD39^+$ by measuring the degradation of exogenous ethenylated ATP at cell numbers ranging from 0 to 100,000. I found that there was a linear relationship between the percentage of ATP hydrolysis and cell number between 0 and 10,000 $CD8^+CD39^+$ T cells (R^2 = 0.996, $P<0.05$)(Figure 4.5B). However, there was a non-linear relationship at 50,0000 cells suggesting that maximal ATP hydrolysis was achieved and therefore HPLC measurement of ATP are unreliable in accurately quantifying ATP hydrolysis at these higher numbers of activated $CD8^+$ T cells. As previously shown by luciferase assays I next tested whether POM 1 induced inhibition of ATP hydrolysis and was detectable by rpHPLC. My results demonstrate that ATP hydrolysis drops after addition of POM 1 beginning at a

concentration of 50μM and upto 100μM after exogenous addition of 5,000nM of ATP (39% ATP hydrolysis at 0μM and 6% at 100μM of POM1) in 10,000 CD8$^+$CD39$^+$ T cells (Figure 4.5C). CD8$^+$CD39$^-$ T cells demonstrated low levels of ATP hydrolysis overall as expected (4.7% 0μM and 1.3% at 100μM of POM 1). Therefore, rpHPLC is able to reliably and accurately quantify CD39 activity on primary human CD8$^+$ T cells at low cell numbers between 5,000 and 10,000.

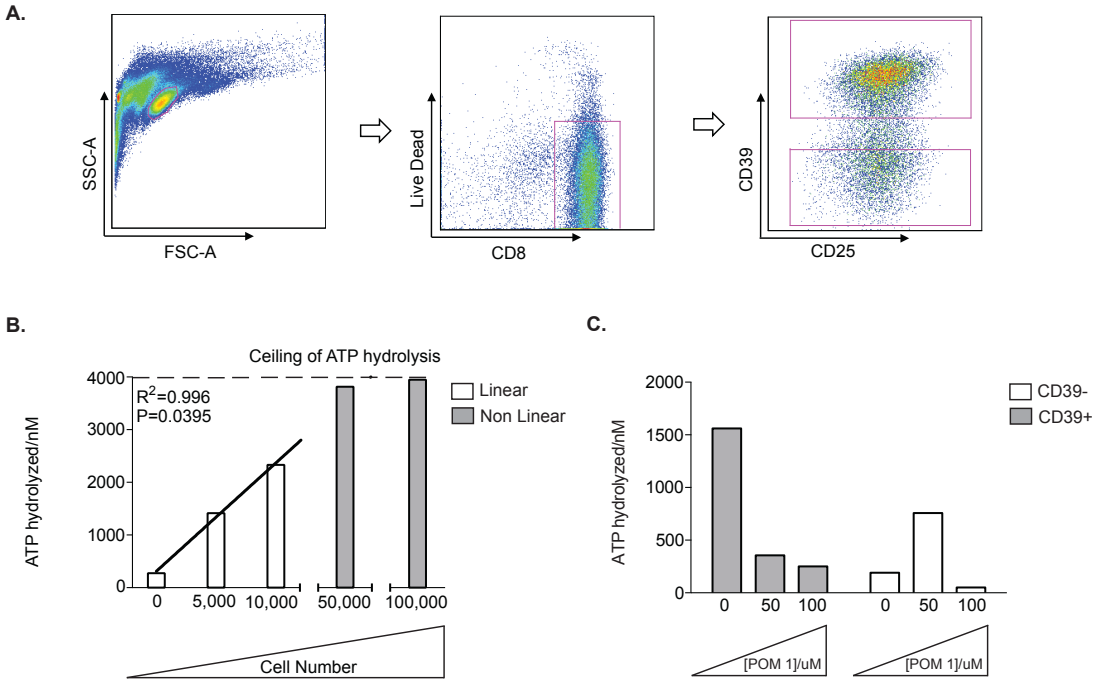


Figure 4.5. Detection of CD39 activity on CD8$^+$CD39$^{+/-}$ and CD4$^+$CD25$^+$CD39$^+$ T cells from healthy and HCV infected donors was performed by reverse phase High Performance Liquid Chromatography (rpHPLC). (A) Flow cytometry analysis demonstrating TCR stimulated CD8$^+$CD39$^+$ T cells sorted for ATP hydrolysis quantification. **(B)** Histogram demonstrating detection of ATP hydrolysis down to 5,000 CD8$^+$CD39$^+$ TCR stimulated T cells (A). **(C)** Bar chart demonstrating inhibition of CD39 mediated ATP hydrolysis by the CD39 inhibitor Sodium polyoxotungstate 1 (POM 1). Statistical significance was assessed by linear regression

4.4.6 CD39 expressed by $CD8^+$ T cells in HCV infection hydrolyses ATP

The upregulation of CD39 on virus-specific $CD8^+$ T cells in chronic HIV and HCV infection suggests that CD39 may have a functional role in the regulation of the T cell response to viral antigens. A number of studies have shown that T cells can control ATP release from their intracellular compartment in response to various extracellular stimuli supporting the role of purinergic signaling as having an active role when T cells respond to antigen (168-170). CD39 (ENTPD1) converts extracellular ATP to AMP, and CD73 degrades AMP allowing for homeostatic control of adenosine concentrations. Several studies have observed that CD39 deficient mice demonstrate a reduced ability to mediate inflammatory responses attributable to reduced ATP hydrolysis, adenosine formation and the adenosine A2A receptor activation (171). Although much interest has been concentrated on examining the expression of CD39 on the surface of $CD4^+$ Tregs, there is evidence that this pathway is responsible for the inhibition of cytotoxic $CD8^+$ T cells as well (101, 172). Several techniques now allow for ATP hydrolysis to be directly quantified from small numbers of cells (123, 166). Therefore, we asked whether CD39 expressed on the surface of $CD8^+$ T cells from patients infected with chronic HCV was functional by using ATP hydrolysis as a surrogate marker of CD39 activity.

To address this, we used flow cytometry to sort populations of T cells from the $CD4^+$ and $CD8^+$ T cell compartments that did and did not express CD39. Using 6 patients infected with chronic HCV we sorted three subsets of 10,000 lymphocytes including a positive control closely resembling Tregs ($CD4^+CD25^+CD39^+$), $CD8^+CD39^+$ and negative control ($CD8^+CD39^-$) T cells

(Figure 4.6A, 4.6B). As previously described, I used rpHPLC to quantify CD39 activity by measuring ATP hydrolysis from the supernatants of these three subsets of cells after addition of pre-determined amounts of ethynlated ATP and incubating in culture media. In addition, the mean fluorescence intensities of CD39 (MFI) of the sorted T cell populations were analysed and compared with quantities of ATP that were hydrolysed.

MFI data was normalised to the positive control Treg population and demonstrated that the highest MFI being $CD4^+CD25^+CD39^+$ T cells, followed by the $CD8^+CD39^+$ T cells (mean 51.1 %), which was significantly higher in CD39 expression than the negative control (mean 0.9 %) $CD8^+CD39^-$ T cell group (Figure 4.6C). This trend in CD39 expression matches the same trend in detectable ATP hydrolysis in respective cell populations. Again, data was normalised to the positive control Treg cell group. Our results demonstrate the highest percentage of ATP hydrolysis was performed by the positive control, followed by the $CD8^+CD39^+$ population (mean 72.3%), which was also significantly greater than the $CD8^+CD39^-$ population (Mean 41.7%, Figure 4.6D). This highlights that CD39 expression is significant on $CD8^+$ T cells in patients with chronic HCV infection as well as functionally active and capable of hydrolysing ATP. Therefore CD39 may contribute to modulation of the $CD8^+$ T cell response in chronic HCV infection.

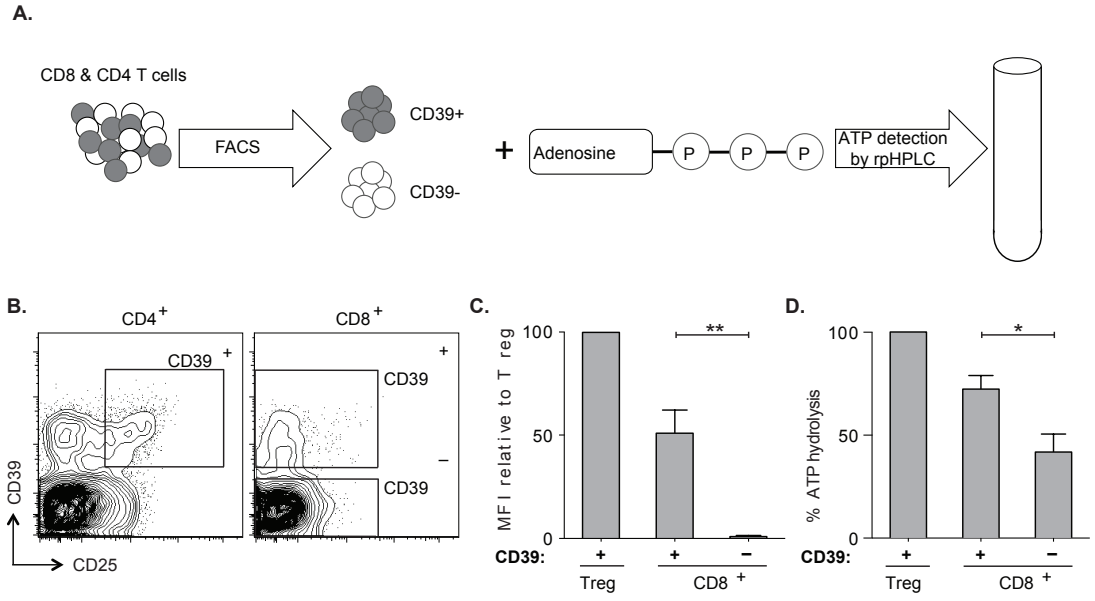


Figure 4.6. CD39 activity is quantifiable on CD8+ T cells in chronic HCV infection. **(A)** Diagram representing workflow for analysing CD39 activity on $CD8^+CD39^+$ T cells in HCV infection. Cell populations were sorted by flow cytometry and exogenous ethenylated ATP added. Supernatent was then analysed for remaining ATP levels by rpHPLC. **(B)** Flow cytometry sorting gates of $CD8^+CD39^{+/-}$ T cells and $CD4^+CD25^+CD39^+$ Tregs used for rpHPLC analysis of CD39 activity. **(C)** Summary of CD39 expression level from cells in (B) relative to Treg. **(D)** Percentage ATP hydrolysis detected in $CD8^+$ T cell populations relative to Treg. Data represent 6 patients with chronic evolving HCV infection. Error bars represent s.e.m. Statistical significance was assessed by paired Student's t-test (C, D). *P<0.05, **P<0.01.

4.4.7 CD39 is elevated in activated T cells

Previous studies have documented that CD39 is upregulated during activation states in both $CD4^+$ and $CD8^+$ T cells (102, 159). Having confirmed already that CD39 is upregulated upon activation of the TCR complex in $CD4^+$ and $CD8^+$ T cells, I next tested a range of activation markers on healthy (n=9),

HCV (n=5) and HIV (n=21) $CD8^+$ T cells to determine whether CD39 was co-expressed with markers of activation between samples with and without chronic infection. Using flow cytometry I analysed the expression of the activation markers CD69, CD38 and HLA DR. I found that the activation marker CD69 was significantly raised in $CD39^+CD8^+$ T cells from samples with chronic HCV and HIV infection (mean HCV 2%, $p<0.01$; mean HIV 12%, $p<0.001$) whereas CD69 was not significantly raised in $CD39^+CD8^+$ T cells from healthy donors (Figure 4.7A). However, both CD38 and HLA DR expression were significantly increased on $CD39^+CD8^+$ T cells from healthy donors (mean CD38 37%, $p<0.001$; mean HLA DR 52% $p<0.001$) and in chronic HIV (mean CD38 71%, $p<0.001$; mean HLA DR 60% $p<0.001$). But only HLA DR was raised on $CD39^+CD8^+$ T cells in HCV infection (mean HLA DR 37% $p<0.0001$)(Figure 4.7B, 4.7C). Therefore CD39 is upregulated in a wide range of activation states on $CD8^+$ T cells from both healthy and chronic infection such as HCV. This suggests that the role of CD39 in modulating T cell responses in chronic infection is coupled to its cellular activation status, rather than merely being a marker of activation alone.

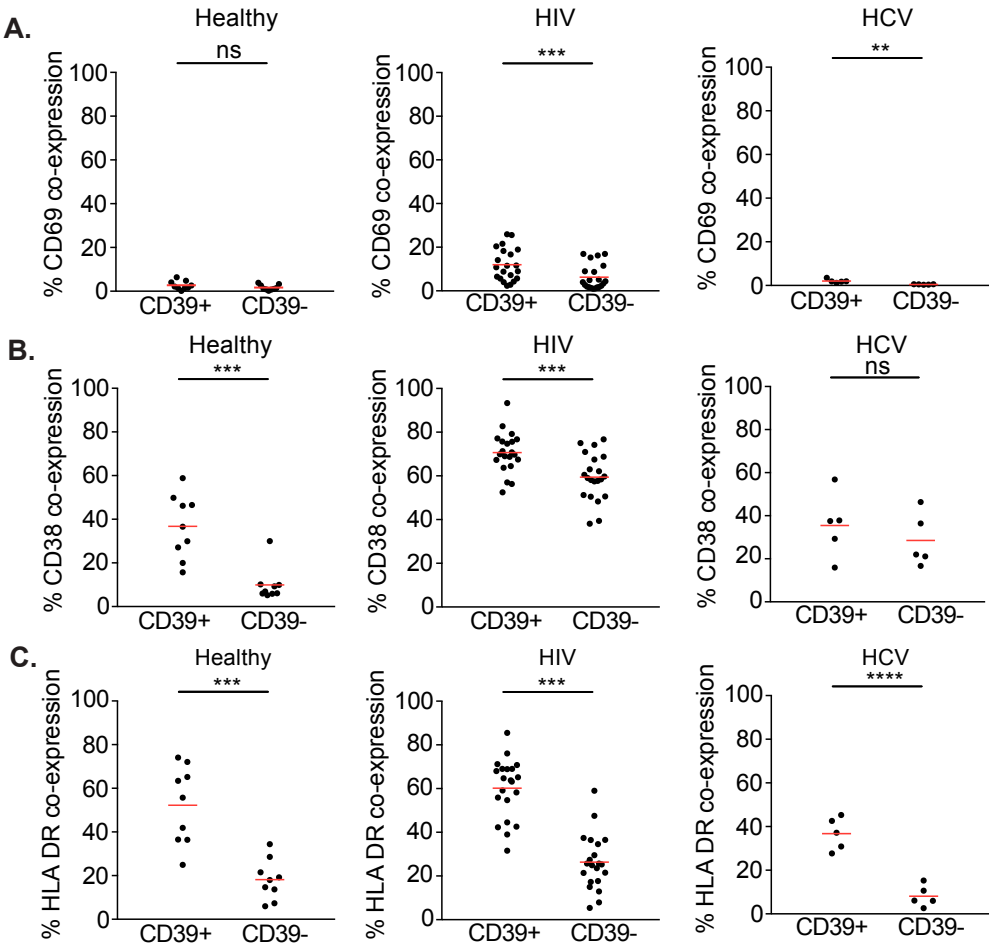


Figure 4.7. Flow cytometry analysis of human CD8+ T cells identifies increased expression of activation markers with CD39 across healthy human donors and those with chronic viral infections. CD8+ T cells from donors with and without chronic viral infections were stained with CD69, CD38 and HLA DR. **(A)** CD69 expression was increased in CD8+C39+ T cells from samples with chronic HCV and HIV infection but not healthy donors. **(B,C)** Scatter plots demonstrating increased levels of CD38, HLA DR expression with CD8+CD39+ T cells in samples with and without chronic viral infection. Statistical significance was assessed by paired Student's t-test (A-C). ns=not significant, **P<0.01, ***P<0.001, ****P<0.0001.

4.4.8 PD-1 is co-expressed with $CD8^+CD39^+$ T cells in chronic HCV and HIV infection

HCV-specific $CD8^+$ T cells in chronic viral infection display an increase in activation markers due to their profile as exhausted T cells. I next tested if $CD39^+$ T cells expressed markers of T cell exhaustion and compared healthy donors (n =19) with those with chronic HCV (n=5) and HIV (n = 18) infections. In order to profile exhaustion markers I tested the inhibitory markers CD160, PD-1 and CD161, which are highly upregulated on HCV-specific $CD8^+$ T cells. My results demonstrate that there was no statistically significant difference in CD160 expression between $CD39^+$ and $CD39^-$ $CD8^+$ T cells in all sample groups (Figure 4.8A). CD161 demonstrated increased expression in $CD8^+CD39^-$ T cells in healthy donors (mean CD161 20%, $p<0.001$), whereas there was no significant difference in CD161 expression in HCV or HIV infection (Figure 4.8B). However, PD-1 expression was significantly raised in $CD8^+CD39^+$ T cells in chronic HCV and HIV infection (mean HIV 48 %, $p<0.001$; mean HCV 37%, $p<0.01$) but there was no difference in PD-1 expression between $CD39^+$ and $CD39^-$ T cells in healthy donors (Figure 4.8C). Since PD-1 is a well recognised marker of exhaustion, These data suggests that CD39 expression is associated with increased PD-1 expression and a state of $CD8^+$ T cell exhaustion in the context of chronic viral infection.

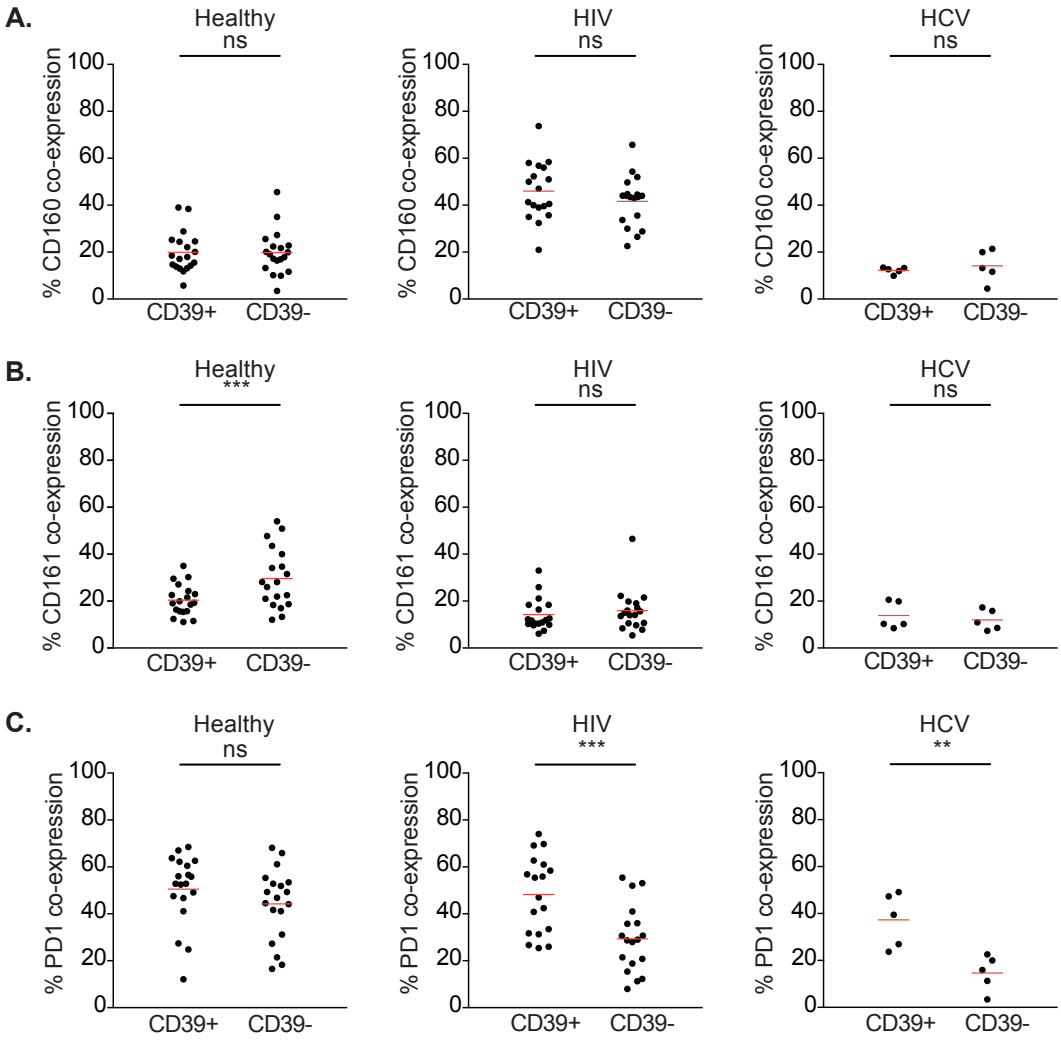


Figure 4.8. Flow cytometry analysis of human CD8+ T cells identifies increased co-expression of CD39 with PD-1 in donors with chronic viral infection. $CD8^+$ T cells from donors with and without chronic viral infections were stained with CD160, CD161 and PD-1 to look for co-expression of CD39 with markers of T cell exhaustion. **(A-B)** Scatter plots demonstrating no significant difference in percentage expression of CD160, CD161 on $CD8^+CD3^+$ and $CD8^+CD39^-$ T cells. **(C)** Significantly increased levels of PD-1 expression on $CD8^+CD39^+$ T cells in HIV and HCV infection. ns=not significant, **P<0.01, ***P<0.001.

4.4.9 CD39 is co-expressed with PD-1 on virus-specific $CD8^+$ T cells

T cell exhaustion is described as a state of T cell dysfunction that occurs during many chronic viral infections (55, 65). The cell surface receptor PD-1 has been identified as a major inhibitory receptor pathway that plays a central role in immunoregulation and often referred to as a marker of T cell exhaustion (107). PD-1 is highly expressed in chronic HCV and HIV where antigen load is high (173, 174). No studies have analysed the phenotypic relationship of CD39 with PD-1 in the context of chronic viral infection on viral-specific $CD8^+$ T cells. Therefore, we tested 40 samples infected with HIV (28 chronic progressors, 7 viraemic controllers and 5 elite controllers) and 39 samples with HCV (21 chronically infected and 13 resolvers) to investigate the relationship between CD39 and PD-1 expression on viral-specific $CD8^+$ T cells in patients with chronic viral infection. We quantified levels of PD-1 and CD39 expression using their mean fluorescence intensity (MFI) values and percentage expression relative to unstained samples by flow cytometry.

Flow cytometry analysis demonstrated a high degree of CD39 and PD-1 expression in viral-specific $CD8^+$ T cells in HIV and HCV. Some HCV-specific $CD8^+$ T cells from patients with chronic disease at 82 weeks after infection demonstrated a novel population of highly expressing PD-1^{High} and $CD39^{High}$ $CD8^+$ T cells, which was less distinct in HIV infection (Figure 4.9A, 4.9B). In addition, analysis of the mean fluorescence intensities of conjugated fluorochromes to anti- PD-1 and anti-CD39 antibodies revealed a statistically significant positive correlation between CD39 and PD-1 expression (R^2=0.56, P

<0.0001 for HCV and R^2=0.3, P<0.05 in HIV) on HCV and HIV-specific $CD8^+$ T cells (Figure 4.9C, 4.9D). These data suggest that CD39 is upregulated on highly exhausted PD-1 expressing $CD8^+$ T cells in chronic HCV and HIV infection.

We next asked if CD39 expression on viral-specific $CD8^+$ T cells in chronic viral infection could differentiate between patients with varying disease severity. We found that when separating patients that are clinically defined as having chronic HIV or HCV infection and those that resolve virus, expression of CD39 is higher in viral-specific $CD8^+$ T cells in chronic HIV (mean 37.2 % chronic progressors and 15.5 % in viraemic and elite controllers combined) but not significantly higher in chronic HCV infection (Appendix Figure 9.4A, 9.4B). PD-1 expression demonstrated the opposite finding by being significantly decreased in HCV^-specific $CD8^+$ T cells form patients that resolve virus (mean 90.2% in chronics and 40.4 % resolvers) whereas not significantly changed between HIV-specific $CD8^+$ T cells (mean 74.2 % chronic progressors and 80.0 % in viraemic and elite controllers combined) from patients that develop chronic infection or resolve virus (Appendix Figure 9.4A, 9.4B). However, we discovered that when expression of both CD39 and PD-1 was taken into account, the percentage of $CD39^+$ $PD\text{-}1^+$ viral-specific $CD8^+$ T cells demonstrated a statistically significant increase in expression from chronic infection on both HCV^- (mean 56.8 % chronic infection and 22.4 % in resolvers) and HIV-specific (mean 30.9 % chronic progressors and 10.5 % in viraemic and elite controllers combined) $CD8^+$ T cells (Figure 4.9E, 4.9F). Thus identifying the population of $CD39^+$ $PD\text{-}1^+$ viral-specific $CD8^+$ T cells as a potential clinical correlate in chronic viral disease.

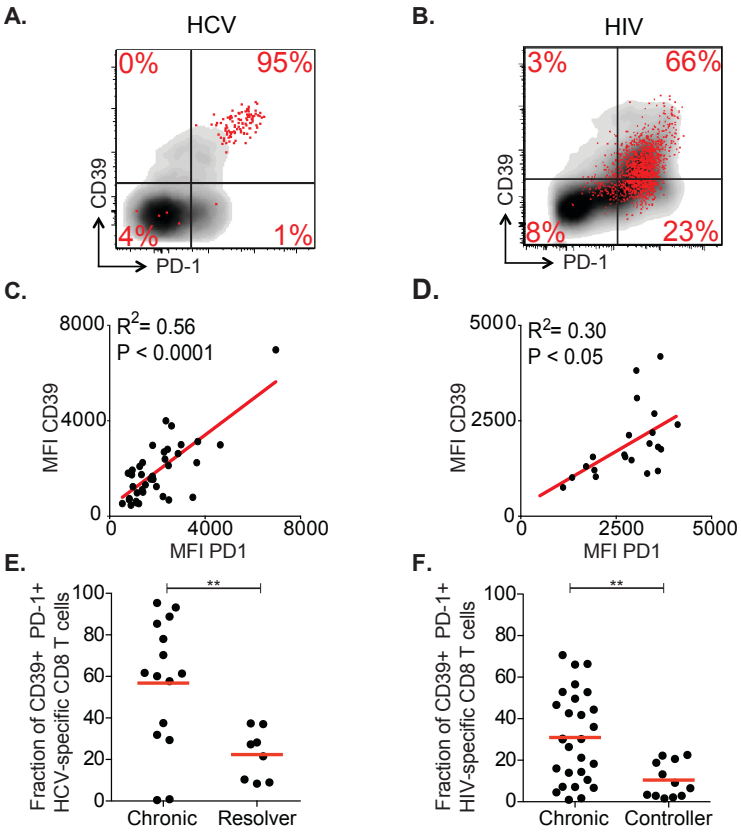


Figure 4.9. CD39 is co-expressed with PD-1 on virus-specific CD8+ T cells. (A,B) Flow cytometric analysis of CD39 versus PD-1 expression in chronic HCV (A) and HIV infections (B). Representative plots demonstrate total (gray) and virus-specific (black) CD8+ T cells. **(C, D)** Correlation between CD39 and PD-1 expression of HCV-(C) and HIV-specific CD8+ T cells (D). 40 samples with HIV (21 chronic progressors, 7 viraemic controllers and 5 elite controllers) and 39 patients with HCV (21 chronically infected and 13 resolvers) infection were tested. **(E, F)** Percentage of CD39+ PD-1+ virus-specific CD8+ T cells compared between HCV chronics and resolvers (E) or HIV progressors and controllers (F). Statistical significance was assessed by Linear regression (C, D) or unpaired Student's t-test (C). **P<0.01.

4.4.10 CD39 expression on viral-specific CD8$^+$ T cells correlates with viral load in chronic HCV and HIV infection

Having defined the expression patterns of CD39 and PD-1 on viral-specific CD8$^+$ T cells in HCV and HIV infection that can help to distinguish between patients clinically defined as having resolved virus or not. We next sought to determine how closely associated CD39 expression is with viral load, the main clinical parameter segregating patients from chronic and resolved infection states. It is well recognised that PD-1 expression is increased on the surface of viral-specific CD8$^+$ T cells in the face of persistent virus in both human and mouse chronic infection models. However, there are no studies that correlate the expression pattern of CD39 on viral-specific CD8$^+$ T cells with viral load in human chronic viral infections such as HIV and HCV. To determine this we measured the expression of CD39 on viral-specific CD8$^+$ T cells from 28 patients with chronic HCV infection and 21 patients with chronic HIV infection and correlated these findings with their corresponding viral load titers. In addition, we also measured PD-1 expression on the same cell populations to act as a positive control and recapitulate previous studies.

We found that both the HCV and HIV sample groups demonstrated a statistically significant positive correlation between viral load and MFI of CD39 expression on viral-specific CD8$^+$ T cells (Figure 4.10A, 4.10B). The correlation between viral load and CD39 expression on HCV-specific CD8$^+$ T cells was stronger and more statistically significant than on HIV-specific CD8$^+$ T cells (R^2 = 0.40, $p<0.0005$ on HCV and R^2 = 0.35, $p<0.005$ for the HIV-specific cohort). PD-

1 expression showed a similar trend with both HCV and HIV-specific $CD8^+$ T cells demonstrating positive correlations with viral load titers ($R^2 = 0.27$, $p<0.005$ on HCV and $R^2 = 0.18$, $p<0.06$ for the HIV-specific cohort), however only HCV was statistically significant (Figure 4.10C, 4.10D). These data suggest that CD39 expression is more closely correlated to viral antigen load than PD-1 on HCV- and HIV-specific $CD8^+$ T cells.

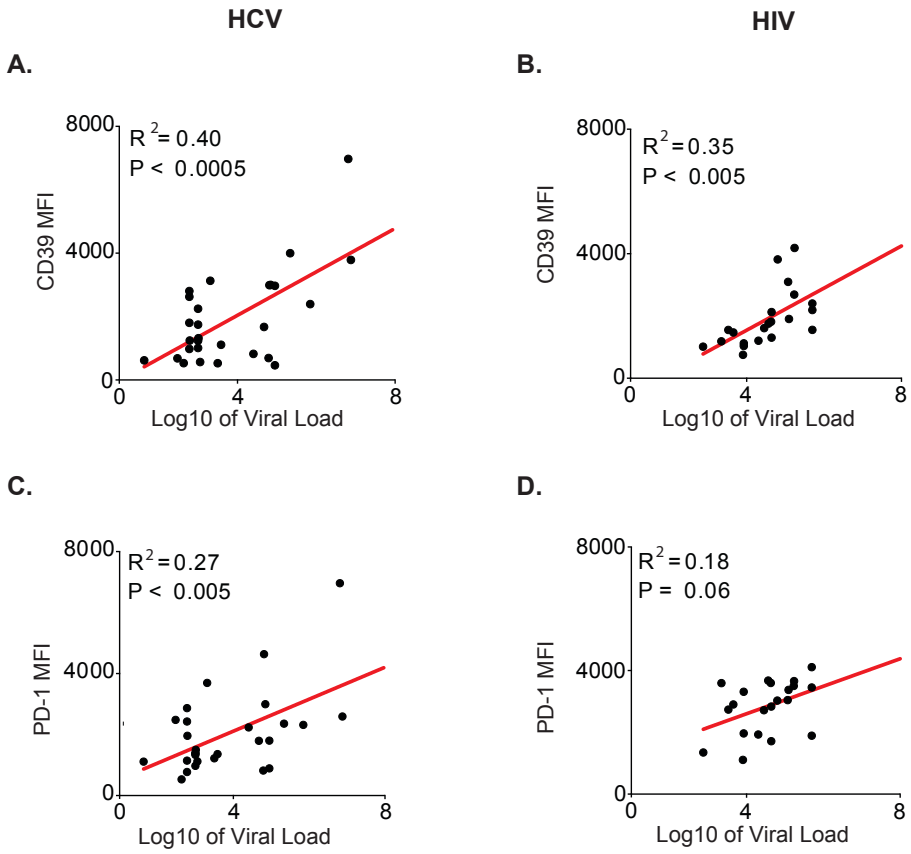


Figure 4.10. CD39 expression on viral-specific CD8+ T cells correlates with viral load in chronic HCV and HIV infection. Flow cytometric analysis of 28 chronic HCV and 21 progressor HIV infection samples after staining with CD39 and PD-1 monoclonal antibodies. **(A, B)** Significant positive correlation between MFI CD39 expression of virus-specific $CD8^+$ T cells and Log_{10} of viral load counts in HCV (A) and HIV (B). **(C, D)** Significant correlation between PD-1 expression of virus-specific $CD8^+$ T cells and Log_{10} of viral load counts in HCV (C) from but no significant correlation between MFI PD-1 and Log_{10} of viral load in HIV (D). Statistical significance was assessed by Linear regression (A – D).

4.5 Discussion

Findings:

1. **CD39 is highly expressed by viral-specific $CD8^+$ T cells in chronic viral infections such as HCV and HIV.**
2. **CD39 is functionally active on the surface of $CD8^+$ T cells from donors with chronic HCV infection.**
3. **CD39 is highly co-expressed with the exhaustion marker PD-1 on HCV and HIV-specific $CD8^+$ T cells.**
4. **CD39 expression correlates with viral load and when co-expressed with PD-1 is associated with poor clinical outcome.**

The expression of CD39 has been studied on numerous cell populations including B cells, natural killer (NK) cells, dendritic cells, Langerhans cells, monocytes, macrophages, mesangial cells, neutrophils, and regulatory T cells (Tregs) in particular (150). The importance of CD39 on T cells has been increasingly recognised over the last few years. $CD4^+$ Tregs have received special attention by researchers, as there is growing evidence that the $CD4^+$ Treg -mediated CD39-CD73-adenosine pathway is able to regulate effective T cell immunity to tumors and pathogens and autoimmune disease (101, 159). Tregs display a distinct population of $CD39^+$ subpopulation of cells and have been shown to be the major rate-limiting step in T cell mediated inhibition via the purinergic pathway in both murine and human models. (101). Emerging evidence has identified that a small subpopulation of $CD8^+$ T cells express CD39 and are

also involved in immunosuppression via adenosine (102, 175). These subpopulations have been sometimes referred to as $FOXP3^+CD8^+$ regulatory T cells and have been studied in the context of pathogenic infections such as Mycobacterium tuberculosis in humans and simian immunodeficiency virus in primates (176). However, the phenotype of CD39 on viral-specific $CD8^+$ T cells in chronic HCV or HIV infection in humans has never been observed. I show that high- expression levels of the ectonucleotidase CD39 is characteristic of $CD8^+$ T cells specific for chronic viral infections in humans but is otherwise rare in the $CD8^+$ T cell compartment of healthy donors.

4.5.1 CD39 expression in chronic viral infection in humans

Previous data has indicated that CD39 expression is mostly restricted to $CD4^+$ regulatory T cells and small subpopulations of $CD8^+$ T cells (102). I demonstrate that in the bulk population of $CD8^+$ T cells in healthy donors only a small minority of $CD8^+$ T cells expresses CD39. However, CD39 is highly expressed by virus-specific $CD8^+$ T cells in two human chronic infections (HIV and HCV). This suggests that high expression of CD39 by $CD8^+$ T cells is a pathological occurrence in chronic viral infection where $CD8^+$ T cells receive chronic antigenic stimulation. Interestingly the proportion of HCV-specific $CD8^+$ T cells enriched with CD39 is significantly greater than HIV-specific $CD8^+$ T cells, which may be due to various reasons. Firstly, by virtue of them being different viruses there will be an induction of different T cell responses to viral antigen. For example the pattern of inhibitory receptors expressed by HIV-specific $CD8^+$ T

cells will be different to HCV-specific CD8$^+$ T cells, which could result in the up or down regulation of proteins such as CD39 via pathways which are as of yet unknown. Secondly, increased CD39 expression on HCV-specific CD8$^+$ T cells may be related to differences in the timing of blood sampling during the course of infection, which may affect the extent of antigen-load and inflammation in the two infections. Thirdly, the number of viral-specific CD8$^+$ T cells that are present in the CD8$^+$ T cell compartment may be greater in HIV infection, but only a small fraction being detected by tetramer response on analysis. This can artificially skew data to appear as if CD39 expression is highest in HCV-specific CD8$^+$ T cell pool. Finally, the extent of T cell exhaustion of viral-specific CD8$^+$ T cells detected could alter results. HIV samples may demonstrate lower degrees of T cell exhaustion, such that if CD39 were a marker of exhaustion then it's expression would be less than in HCV.

4.5.2 Quantification of CD39 enzymatic activity in HCV

The presence of CD39 on the cellular surface of CD8$^+$ T cells suggests that it has an active role in cellular processes. It has been described as an activation marker on lymphoid cells and it's degradation products are used to produce adenosine, which is a critical regulator of innate and adaptive immune responses, inhibiting T lymphocyte proliferation and the secretion of inflammatory cytokines including IL-2, TNF α, and IFN- γ (177). Therefore, the functional ability of CD39 to hydrolyse ATP in the context of chronic viral infection is central to its role in modulating immunogenic responses to viral pathogens. My results

demonstrate that CD39 is functionally active on the surface of $CD39^+CD8^+$ T cells in samples with chronic HCV infection and therefore may contribute to $CD8^+$ T cell inhibition in HCV infection.

It is clear from these data that CD39 mediated hydrolysis of ATP is significantly greater in $CD8^+CD39^+$ T cells than in $CD8^+CD39^-$ T cells. Although this is not a surprise, there are two important observations that can be derived from theses results. The first is that although the MFI of CD39 in the $CD39^-$ population was extremely low, there was still a detectable level of ATP hydrolysis. This could be due to other ectonucleotidases that perform ATP hydrolysis, but to a lesser extent than CD39 (ENTPD1). There are several other ENTPDases that can hydrolyse ATP, although ENTPD1 is well recognised as being by far the major contributor for this process (178). Also, the rpHPLC assay used to detect degradation of ATP by detecting fluorescent tagged ethene to ATP molecules may have detected exaggerated ATP hydrolysis due to background degradation of ATP secondary to light exposure. Secondly, $CD4^+$ Tregs demonstrate a higher degree of CD39 expression and ATP hydrolysis than in $CD8^+CD39^+$ T cell subsets. This finding may be of significance as it suggests the relative effects of T cell inhibition from $CD4^+$ Tregs may well be greater than $CD8^+$ T cells via the purinergic pathway, although the cellular contexts may be completely different.

4.5.3 CD39 as a marker of T cell exhaustion in HCV and HIV

The state of $CD8^+$ T cell exhaustion is characterised by widespread changes in gene expression relative to functional memory $CD8^+$ T cells (153). The inhibitory receptor PD-1 is an important pathway regulating $CD8^+$ T cell exhaustion and previous studies have shown that blockade of the PD-1:PD-L1 pathway alone can regulate function of viral-specific $CD8^+$ T cells in LCMV, SIV, HIV and HCV infections (89, 107, 111, 174, 179, 180). However, in humans, identifying specific markers of T cell exhaustion that are not shared by more functional $CD8^+$ T cell populations has been challenging (157). These data indicate that CD39 expression is a novel marker of $CD8^+$ T cell exhaustion in chronic viral infections such as HCV and HIV.

Various pieces of evidence suggest that CD39-expressing $CD8^+$ T cells are specific to T cell exhaustion. First, human virus-specific $CD8^+$ T cells co-express CD39 with PD-1, which is an inhibitory receptor expressed by the majority of exhausted T cells (107, 153). Second, CD39 expression correlates with viral load in subjects with HIV and HCV infection suggesting that the conditions of high levels of inflammation and antigen load that lead to exhaustion also increase CD39 expression in the virus-specific pool of $CD8^+$ T cells, as has been observed for PD-1 (111, 180). Thirdly, co-expression of CD39 and PD-1 on viral-specific cells significantly correlates with disease progression in both HCV and HIV, suggesting that these are amongst the most highly exhausted group of cells due to chronic antigen stimulation.

The expression of a range of molecules that inhibit T cell function has been used to identify exhausted $CD8^+$ T cells in several studies of human chronic infection and cancer. However, there are important distinctions between the pattern of CD39 expression and that of inhibitory receptors. Many inhibitory receptors, such as PD-1 and CD244 are also expressed by a substantial fraction of $CD8^+$ T cells in healthy donors that are not exhausted (111, 157). In contrast, CD39 expression is found only in a very small minority of $CD8^+$ T cells from healthy donors. This suggests that CD39 expression, particularly in combination with PD-1, may be useful as a specific phenotype of exhausted $CD8^+$ T cells, at least in HCV and HIV infection. Moreover, CD39 may therefore provide a useful marker to isolate exhausted $CD8^+$ T cells in settings such as tumor-specific responses where many reagents are available to identify antigen-specific T cells. Importantly, while CD39 is rare in the $CD8^+$ compartment in healthy donors, it is expressed by $CD4^+$ Tregs (as is PD-1) making it difficult to distinguish between exhausted $CD4^+$ T cells and Tregs by this marker alone. It is not clear why the small fraction of $CD8^+$ T cells expressing CD39 in healthy donors exists. Perhaps it could be due to a subclinical pathological response or, part of homeostatic control of T cell inhibition.

4.5.4 Conclusion

It is conceivable that CD39 expression may promote T cell exhaustion through mediation of T cell inhibition. For instance, the expression of CD39 might enable $CD8^+$ T cells to provide negative regulation via adenosine in the

same manner as Tregs (101). The fact that CD39 requires both a substrate (ATP) and a downstream enzyme (CD73) to generate adenosine could provide a mechanism to ensure that this negative signaling occurred only in certain contexts such as in inflamed, damaged tissues where the extracellular concentrations of ATP are high and CD73-expressing cells are present. In the following chapters I demonstrate further investigation of CD39 in the context of T cell exhaustion by use of genome-wide expression profiles as well as mouse studies.

CHAPTER 5

Gene expression profiling of $CD39^+$ versus $CD39^-$ $CD8^+$ T cells in HCV infection

5.1 Introduction

The failure to develop broadly effective vaccines to viral pathogens such as HCV highlights our lack of understanding of the fundamental pathways that control the development of a robust and protective $CD8^+$ and $CD4^+$ T cell response during infection. The limitation of our understanding is tightly linked with the analytical tools we have at our disposal to be able to de convolute the highly complex regulatory system of the lymphocyte response. Flow cytometry quickly became a central tool in the discovery of cellular phenotypes involved in the adaptive immune response and continues to play an essential role today (181). Recent advances have allowed variations of flow cytometry based on mass spectrometry that can theoretically allow for up to 100 intracellular and surface markers to be analysed (182). However, the memory lymphocyte compartment is made of a multitude of phenotypically and functionally different cell types, a large extent of which is as of yet undiscovered (43). Therefore, techniques that are able to employ a larger number of parameters must be used to identify novel phenotypes within cellular states that we currently deem as homogenous.

Over the last 15 years the use of large scale, high throughput gene expression platforms have become more widely available and now populations of immune cells can be profiled by not tens of markers but rather thousands using genome-wide expression profiles (109, 183). The detailed and large scale information provided by a microarray analysis of up to 20,000 genes allows subtle but important distinctions to be discovered. Enriched pools of genes may

serve to represent the minimal subset of genes that sufficiently define a phenotype or whose function it is central to.

The use of genome-wide expression profiles has been a powerful tool in characterising the phenotype of T cells in the context of chronic viral infection in humans and animal models (109, 153). Common pathways that are highlighted in both are the upregulation of markers of T cell exhaustion such as the inhibitory receptor PD-1, which has been identified as playing a central role in the development of T cell exhaustion. Since the discovery of PD-1 a number of other inhibitory receptors have been discovered including CTLA4, CD244, CD160 and TIM-3 that are significantly upregulated in HCV infection (77, 81-83). The differential expression of these markers help to identify the state of T cell exhaustion and emerging data suggests that co-expression of multiple inhibitory receptors on HCV-specific CD8$^+$ T cells in the blood leads to more severely exhausted phenotype than PD-1 expression alone (152). Initial clinical trials on ant-PD-1 antibodies as immunotherapeutic interventions against HCV produced varied responses (86). Therefore, understanding the patterns of inhibitory receptors that influence CD8$^+$ T cell immunity as well as identifying novel markers of T cell dysfunction is key to creating a definitive profile of exhaustion so that its prevention and/or reversal can be achieved.

Over recent years, CD39 (ENTPD1) has been increasingly characterised on the lymphocyte population. It is an ectonucleotidase that is highly expressed on regulatory T cells and hydrolyses extracellular ATP and ADP into cAMP, which is then processed into adenosine by the ecto-5'-nucleotidase, CD73 (101,

158, 159). Adenosine is a potent immunoregulator that binds to A2A receptors expressed by lymphocytes causing accumulation of intracellular cAMP, preventing T cell activation and NK cytotoxicity (147, 161). Loss of CD39 in Tregs markedly impairs their ability to suppress T cell activation, suggesting that the juxtacrine activity of CD39 may be an important inhibitory mechanism (101). However, although CD39 expression is well recognised on $CD4^+$ subsets, very little attention has been focused towards its presence on $CD8^+$ T cells, where it is much less abundant. Furthermore, no studies have used genome-wide transcription profiling to phenotype CD39 on $CD8^+$ T cells in the context of chronic viral infection.

Therefore, I demonstrate how genome-wide expression profiling can provide a phenotypic analysis of CD39 expressing $CD8^+$ T cells in chronic HCV infection. Here I discover that effector CD39 expressing $CD8^+$ T cells are enriched with signatures of T cell exhaustion, in particular the inhibitory receptor PD-1, which is recognised to play a central role in exhaustion of HCV-specific $CD8^+$ T cells in chronic HCV infection.

5.2 Aims/hypotheses

Specific Aim 1 (SA1): Use gene expression profiling to phenotype $CD8^+CD39^{+/-}$ T cells from patients with HCV infection.

Hypothesis 1: $CD39^+CD8^+$ T cells have a distinctly different molecular phenotype to $CD39^-CD8^+$ T cells.

Specific Aim 2 (SA2): Use gene expression analysis software (Gene e) to Identify genes enriched in $CD8^+CD39^+$ T cells.

Hypothesis 2: $CD8^+CD39^+$ T cells are enriched in genes associated with T cell exhaustion.

Specific Aim 3 (SA3): Use GSEA analysis to identify clusters of genes that are coordinately upregulated in $CD8^+CD39^+$ T cells.

Hypothesis 3: GSEA analysis can identify gene signatures enriched in $CD8^+CD39^+$ T cells in HCV infection.

5.3 Methods

5.3.1 Human Subjects

All human subjects with HCV infection were recruited at the Massachusetts General Hospital Gastrointestinal Unit and the Department of Surgery with written consent in accordance with the IRB approved study: “Cell mediated immunity in Hepatitis C virus infection”; Protocol # 1999- P-004983/54; MGH Legacy #: 90-7246. A total of 9 HCV infected donors were used. Those with chronic HCV (n = 4) were defined by positive anti-HCV antibody and detectable viral load. HCV resolvers (n = 5) were defined as having an undetectable viral load for at least 6 months. All HCV patients were treatment naive and obtained between 27.9 and 237.3 days post infection.

5.3.2 Antibodies and flow cytometry

The following fluorochrome-conjugated antibodies were used: anti-huCD8α, anti-huCD4, anti-huCD3, anti-huCD39, anti-huCCR7, anti-huCD45RA. Peripheral blood mononuclear cells (PBMCs) were thawed in R10 medium (RPMI medium, 10% fetal calf serum [FCS], 1% penicillin/streptomycin, 1% glutamine); rested for 1 h at 37°C in 5% CO2. 5 million PBMC cells were re-suspended in 50µl of buffer (PBS, 0.1% FBS) and stained with the LIVE⁄DEAD Fixable Aqua Dead Cell Stain Kit (Invitrogen) for exclusion of dead cells. Cells were sorted by BD FACS ARIA II and flow cytometers equipped with FACSDiva v6.1. Gates were set using Full Minus One (FMO) controls. Data were analysed using Flowjo software (Treestar).

5.3.3 Microarray data acquisition and analysis (please refer to methods chapter 2)

Effector $CD39^+$ and $CD39^-$ $CD8^+$ T cells from HCV samples were sorted and pelleted and re-suspended in TRIzol (Invitrogen). RNA extraction was performed using the RNAdvance Tissue Isolation kit (Agencourt). Concentrations of total RNA were determined with a Nanodrop spectrophotometer or Ribogreen RNA quantification kits (Molecular Probes/Invitrogen). RNA purity was determined by Bioanalyzer 2100 traces (Agilent Technologies). Total RNA was amplified with the WT-Ovation Pico RNA Amplification system (NuGEN) according to the manufacturer's instructions. After fragmentation and biotinylation, cDNA was hybridised to Affymetrix HG-U133A2.0 microarrays (Appendix Table 9.10).

5.3.4 Bioinformatics

Prior to analysis, microarray data were pre-processed and normalised using robust multichip averaging (RMA), as previously described (125). Differentially gene expression and consensus clustering was performed using the web based server GENE-E software (www.broadinstitute.org/cancer/software/GENE-E/), and gene set enrichment analysis was performed as described previously (Methods Chapter 2) using either gene sets from web based GSEA server MSigDB (116) or published resources (126, 127)

5.4 Results

5.4.1 Successful cDNA amplification from $CD8^+CD39^+$ and $CD8^+CD39^-$ T cells

Effector $CD8^+$ T cells were isolated from 9 HCV infected patients and were sorted into $CD39^+$ and $CD39^-$ fractions using flow cytometry (Figure 5.1A). Naive $CD8^+$ T cells were excluded as they are known to have low levels of CD39 expression and would dilute the yield of $CD39^+$ cells collected for gene expression profiling. 2 samples were unsuccessfully sorted either due to poor staining for CD39 or insufficient cell numbers. However a total of 16 samples were collected and processed for RNA extraction and cDNA amplification for microarray data acquisition (Appendix Table 9.10). Cell inputs were variable ranging from as low as 433 to 250,000 cells. Good amounts of cDNA were yielded for microarray analysis ranging from 167.52ng/μl to 360.32ng/μl, which demonstrated a statistically significant positive correlation with cell input number (Figure 5.1B).

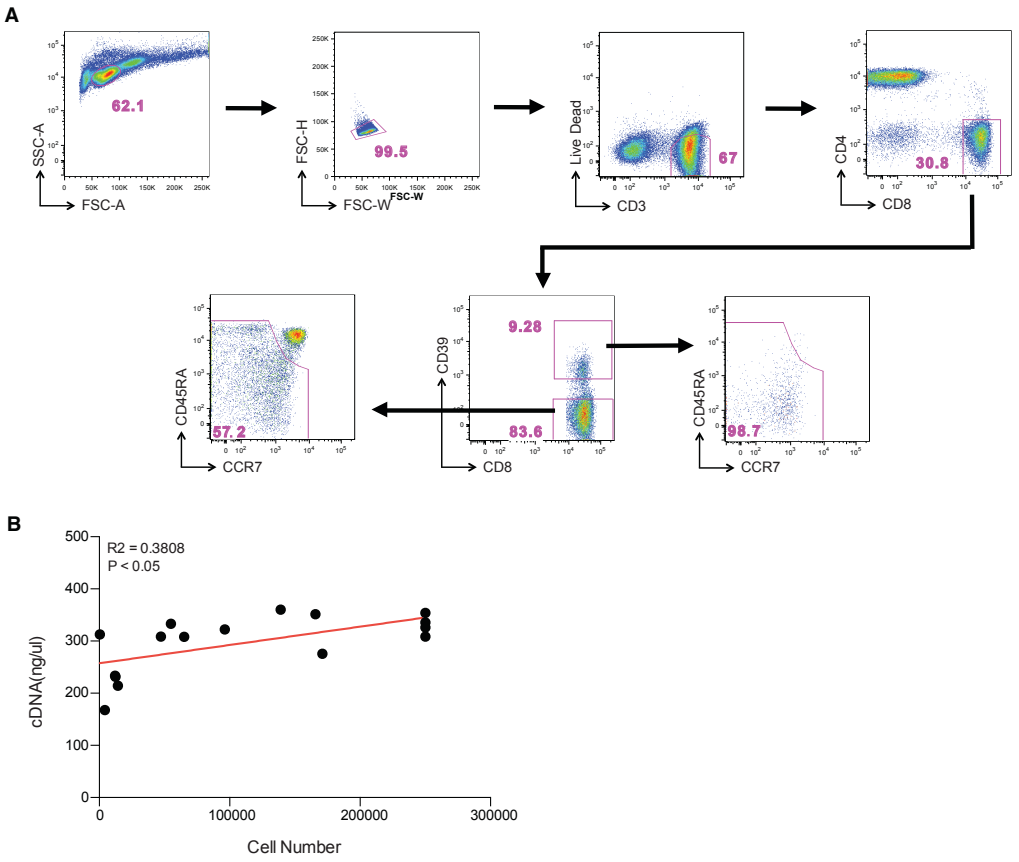


Figure 5.1. Successful cDNA amplification from CD8^{+}CD39^{+} and CD8^{+}CD39^{-} T cells. Human CD8^{+}CD39$^{+/-}$ T cells were isolated by flow cytometry from HCV infected patients. RNA extraction and cDNA amplification was performed in preparation for gene expression analysis. **(A)** Representative gating strategy to sort CD8^{+}CD39$^{+/-}$ from PBMC samples of donors with HCV infection (3 with acute resolving and 5 with chronic infection). **(B)** Graph with total cDNA yields form RNA extracts demonstrating successful cDNA amplification.

5.4.2 Microarray quality metrics analysis identifies one outlier sample

All samples were processed for quality control (QC) assessment of data quality. A sample was deemed to have failed QC assessment if the degree of deviation from all other samples was comparatively high in the data set and

therefore labeled an outlier. Figure 5.2A demonstrates a false colour heatmap of the distances between arrays. The colour scale is chosen to cover the range of distances encountered in the dataset. Patterns in this plot can indicate clustering of the arrays either because of intended biological or unintended experimental factors (batch effects). The distance d_{ab} between two arrays *a* and *b* is computed as the mean absolute difference (L_1-distance) between the data of the arrays (using the data from all probes without filtering). In formula, d_{ab} = mean | M_{ai} - M_{bi} |, where M_{ai} is the value of the *i*-th probe on the *a*-th array. Outlier detection was performed by looking for arrays for which the sum of the distances to all other arrays, $S_a = \Sigma_b d_{ab}$ was exceptionally large. One such array was detected, and it is marked by an asterisk, *. Figure 5.2B shows a bar chart of the sum of distances to other arrays S_a, the outlier detection criterion from Figure 5.2A. The bars are shown in the original order of the arrays. Based on the distribution of the values across all arrays, a threshold of 7.37 was calculated, which is indicated by the vertical line. One array exceeded the threshold and was considered an outlier.

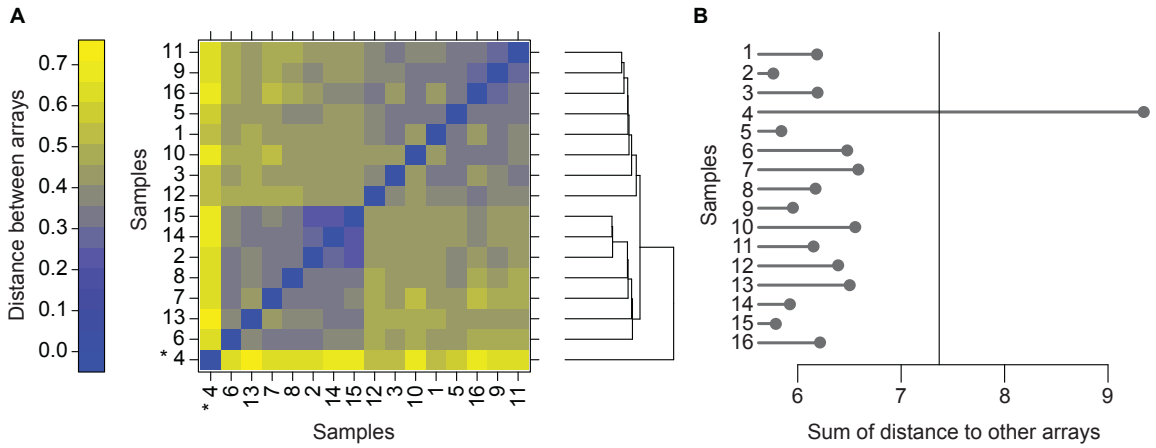


Figure-5.2. Microarray quality metrics analysis identifies one outlier sample. (A) Heatmap demonstrating distance in data between samples calculated by the mean absolute difference between data from all probes without any filtering. 1 outlier is marked by an asterisk, *, and identified by looking for arrays for which the sum of the distances to all other arrays, was exceptionally large. **(B).** Bar chart demonstrating the sum of distances of each array sample to all others. One array exceeded the threshold for distribution (sample 4, threshold 7.37) and was therefore excluded from all other analyses.

5.4.3 Consensus clustering of gene expression profiles identifies 2 broad molecular phenotypes

After removal of samples that failed QC analysis, gene expression data was variance filtered to include only the top 10% of the most variant genes across the dataset. I next sought to identify the differences in gene signatures between CD39^{+} and CD39^{-} CD8^{+} T cells in HCV infection. I first performed an unsupervised analysis of gene expression profiles using hierarchical clustering on the GENE-E software package to determine whether differences in gene expression can accurately identify CD39 expression. Hierarchical clustering (HC) recursively merges samples with other samples, according to their pair-wise "distance" (with the closest item pairs being merged first). As a result, it produces a tree structure, referred to as dendogram, whose nodes correspond to: i) the original items (these are the leaves of the tree); and ii) the merging of other nodes (these are the internal nodes of the tree). If k clusters are required (k 1), the merging proceeds until k nodes are left. This method identifies 2 distinct groups of samples were segregated and corresponded very closely to CD39^{+} and CD39^{-} populations, suggesting that that CD39 expression demarcates two types of CD8^{+} T cells with markedly different patterns of gene expression (Figure 5.3A).

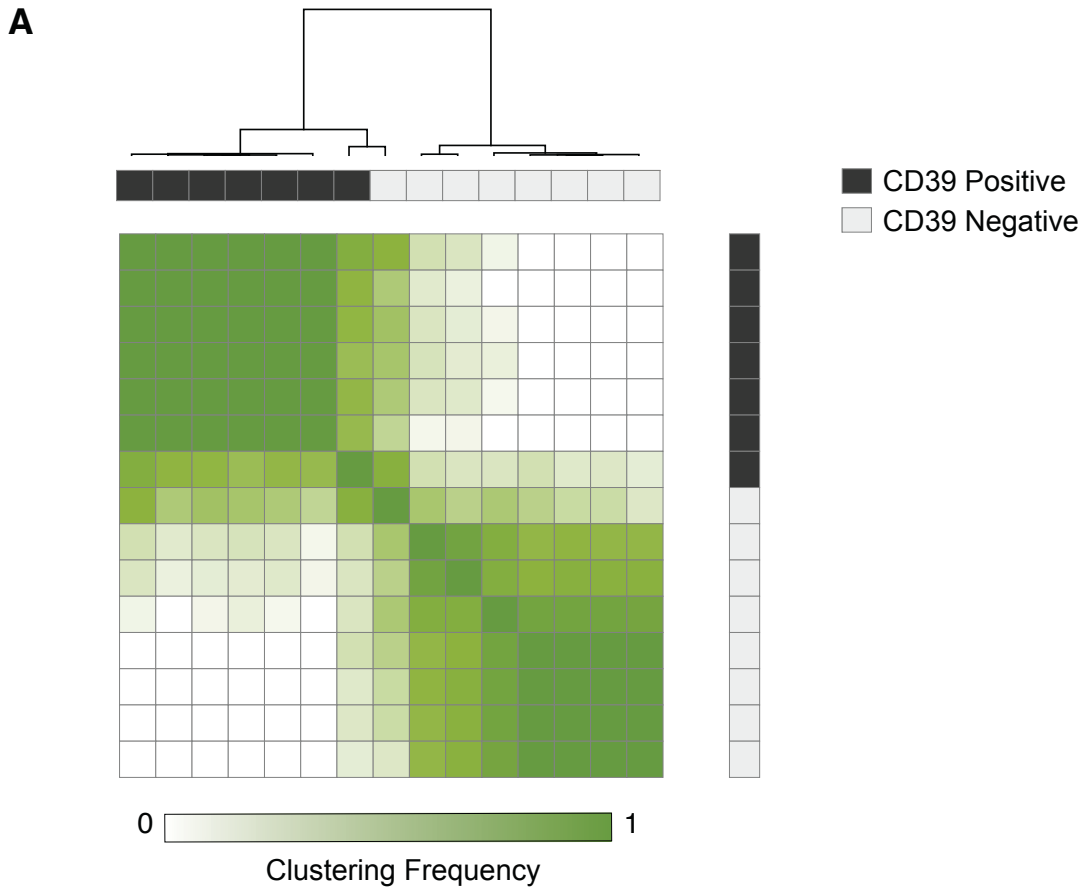


Figure 5.3. Consensus clustering of gene expression profiles identifies 2 broad molecular phenotypes. (A) Consensus clustering of expression profiles from CD39^{+} (black) and CD39^{-} (grey) CD8^{+} T cells from 8 HCV infected patients. Sample similarity is annotated with color form low (white) to high (green) and identifies that gene transcription profiles of CD39^{+} and CD39^{-} CD8^{+} T cells segregate into 2 broadly different molecular phenotypes.

5.4.4 Gene transcriptional analyses reveals markers of exhaustion upregulated in CD8$^+$CD39$^+$ T cells

Having determined through unsupervised analysis that the gene signatures of CD39$^+$ cells are distinctly different from CD39$^-$ CD8$^+$ T cells, I next performed a supervised analysis and identified 745 genes differentially expressed between CD39$^+$ and CD39$^-$ CD8$^+$ T cells with a False Discovery Rate (FDR) cut off of 0.15 (FDR<0.15). Using GENE-E software I obtained a ranked list of the top and bottom 500 genes that were differentially expressed between transcription profiles of CD39$^+$ and CD39$^-$ CD8$^+$ T cells and displayed this as a heatmap (Figure 5.4A). Each coloured cell in the heat map represents the gene expression value for a probe in a sample. The largest gene expression values are displayed in red, the smallest values in blue, and intermediate values in shades of red or blue. Within the list of genes upregulated in CD39$^+$ CD8$^+$ T cells were the inhibitory receptors PD-1 and CTLA-4, suggesting that CD39 shares a phenotype with cells that express multiple markers of T cell inhibition. KLRG-1, which is recognised to play multiple roles in memory and effector CD8$^+$ T cell biology, was more highly expressed in the CD39$^-$ CD8$^+$ T cell population. A full list of genes differentially expressed between CD39$^+$ and CD39$^-$ is in Chapter 9 (Appendix Table 9.11). To identify biological processes that were differentially active in CD39$^+$ vs. CD39$^-$ cells, gene set enrichment analysis (GSEA) was performed using the Gene Ontology collection of gene sets (113). There was no significant enrichment of GO terms in the CD39$^-$ CD8$^+$ subset. In contrast, multiple gene sets significantly enriched (FDR < 0.1) in CD39$^+$ population, almost

all of which were related to mitosis and cell cycle related genes, which were visualised with Enrichment Map software (184) (Figure 5.4B) (Appendix Figures 9.5, 9.6). This suggests that CD39$^+$ CD8$^+$ T cells show coordinate upregulation of genes related to proliferation.

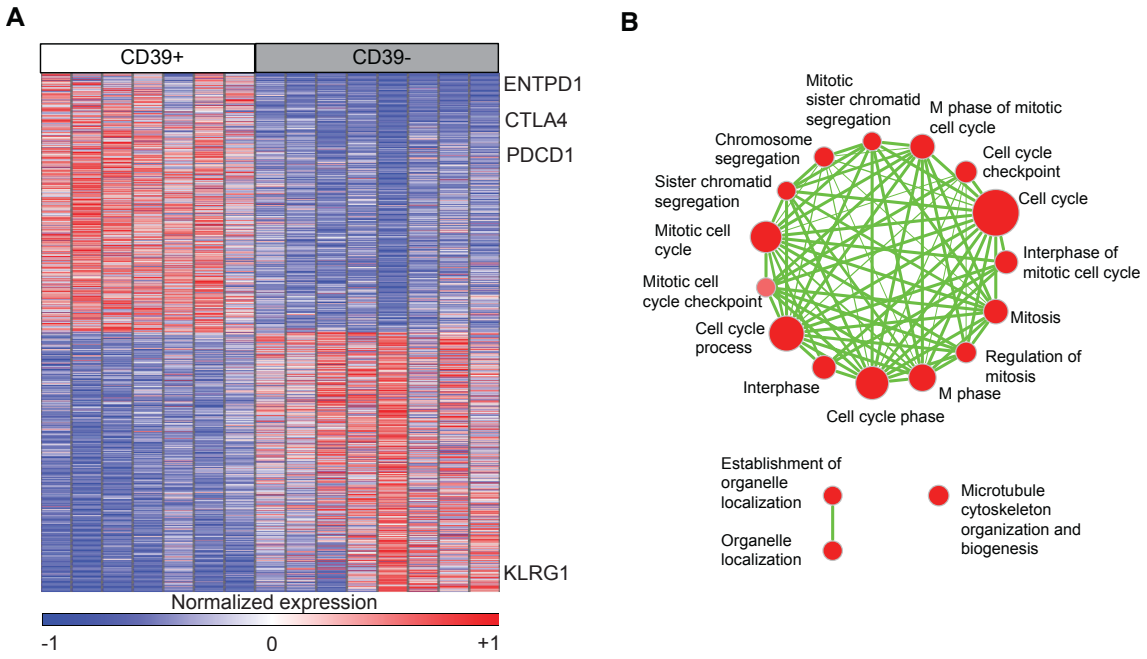


Figure-5.4. Gene transcriptional analyses reveals markers of exhaustion and proliferation upregulated in CD8$^+$CD39$^+$ T cells. (A) Heat map demonstrating genes differentially expressed in CD39$^+$ and CD39$^-$ CD8$^+$ T cells ranked by fold change. Each column representing an individual sample and each row an individual gene, coloured to indicate normalised expression. Low relative expression indicated by blue and high relative expression in red. The exhaustion markers CTLA4 and PDCD-1 (PD-1) are over expressed in CD39$^+$ whereas the marker KLRG1 is overexpressed in CD39$^-$ CD8$^+$ T cells. **(B)** Gene set enrichment map displaying Gene Ontology gene sets enriched (FDR < 0.1) in CD39$^+$ CD8$^+$ T cells from (A). Nodes (in red) are sized in proportion to gene set size; connecting line thickness represents extent of gene member overlap between gene sets.

5.4.5 Nearest neighbour analysis and GSEA identifies ENTPD1 as being highly correlated with PD-1 and enriched with a $PD1^{High}$ signature.

Having clustered genes in a supervised fashion by fold change to identify genes that are upregulated in $CD39^+$ $CD8^+$ T cells, I used an alternative clustering technique to look for additional functionally coordinated genes. By applying a nearest neighbor analysis using the GENE-E platform I sought to identify clusters of genes that are similar to CD39 with regards to their patterns of gene expression across all samples. This method of focusing on nearest neighbours rather than absolute fold change allows the capture of genes with high connectivity and related biological processes (185). Nearest neighbor analysis demonstrates again that the markers PD-1 and CTLA4 are amongst genes that are most similar to CD39 by their pattern of gene expression (Figure 5.5A, Appendix Figure 9.12). I then used GSEA analysis to test whether the PD-1 signature was similar to the $CD39^+$ gene signature. Comparison of genes that are differentially expressed in $CD39^+$ vs $CD39^-$ and $PD\text{-}1^{High}$ vs $PD\text{-}1^{Int}$ in chronic clone 13 infection in mice demonstrated a significant enrichment of genes (Figure 5.5B). This highlights again that the CD39 signature is enriched with markers of T cell inhibition.

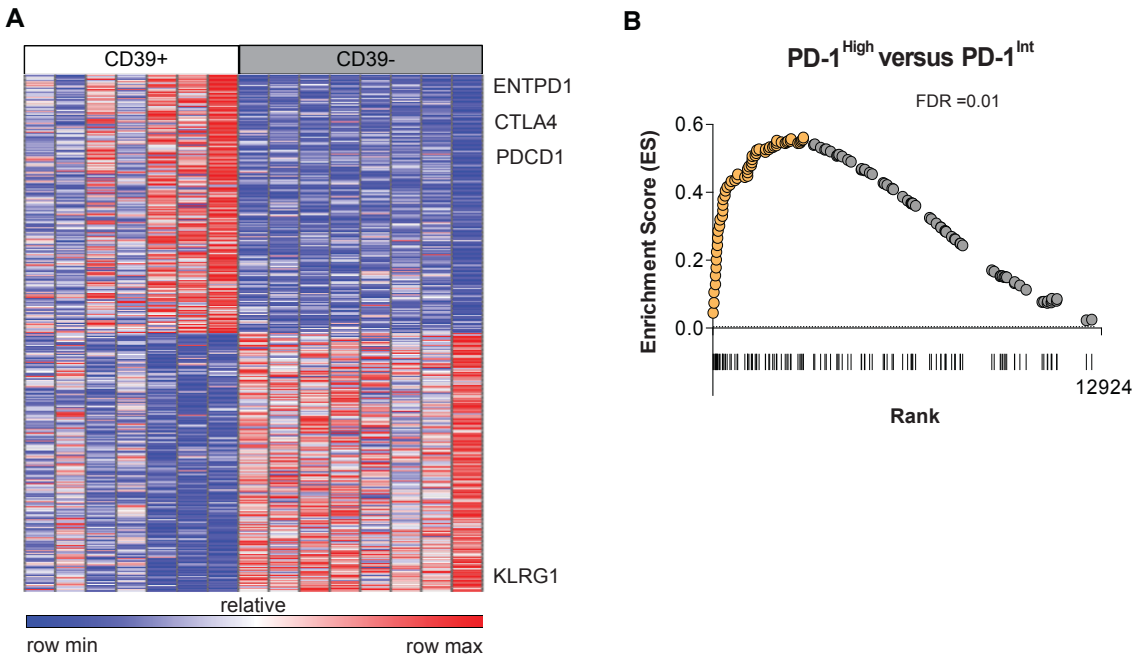


Figure-5.5. Nearest neighbour analysis identifies ENTPD1 as being highly correlated with PD-1 as well as enriched with PD-1^{High} vs PD-1^{Int} signature in chronic clone 13 infection in mice. (A) Nearest neighbour analysis demonstrating genes that are most highly correlated with ENTPD1 in terms of their patterns of gene expression ranked by Pearson correlation. Each column representing an individual sample and each row an individual gene, coloured to indicate normalised expression. Low relative expression indicated by blue and high relative expression in red. Again exhaustion markers, CTLA4 and PDCD-1 are present. **(B)** GSEA analysis demonstrating enrichment of PD-1^{High} vs PD-1^{Int} signature (from mouse clone 13 LCMV chronic infection model) with genes differentially expressed between CD39^{+} vs CD39^{-} CD8^{+} T cells. Leading edge genes (those most enriched) are highlighted in orange.

5.4.6 GSEA identifies exhaustion signature from mouse model of chronic infection enriched in CD8$^+$CD39$^+$ T cells

Because CD39 is most highly expressed by CD8$^+$ T cells in chronic but not acute/latent infection, we tested whether the profile of CD39$^+$ CD8$^+$ T cells was enriched for genes expressed by exhausted CD8$^+$ cells. Previous studies of gene expression in CD8$^+$ T cells in the mouse model of chronic viral infection with the Clone 13 strain of LCMV have identified global signatures of T cell exhaustion that are conserved in exhausted CD8$^+$ T cells in humans (109, 126, 186). I therefore curated a signature of genes upregulated by exhausted CD8$^+$ T cells responding to chronic infection relative to functional memory CD8$^+$ T cells generated by acute infection (LCMV Armstrong strain). I found that the exhausted CD8$^+$ T cell signature from LCMV model was significantly enriched in CD39$^+$ vs. CD39$^-$ CD8$^+$ T cells in subjects with HCV infection (Figure 5.6A). I focused on the “leading edge” genes contributing most to the enrichment, which include those most upregulated both in the mouse exhausted signature and in the human CD39$^+$ profile. As expected, the leading edge genes included PD-1 (PDCD1), a feature of both CD39$^+$ CD8$^+$ T cells and of exhausted CD8$^+$ T cells (Figure 5.6B). In addition, we found upregulation of many genes associated with proliferation including BUB1, TOP2A and MKI67 were shared by mouse exhausted CD8$^+$ T cells and CD39$^+$ CD8$^+$ T cells. Thus CD39$^+$ CD8$^+$ T cells in HCV infection share transcriptional features with exhausted CD8$^+$ T cells in a mouse model of chronic infection that are predominantly related to biological pathways representing proliferation.

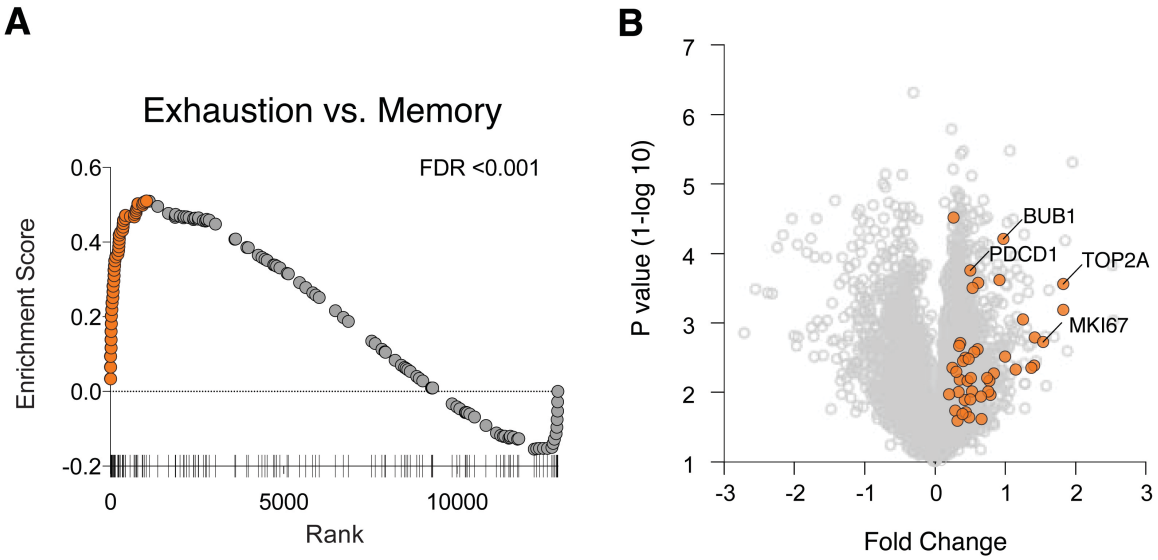


Figure-5.6. GSEA analysis identifies exhaustion signature from mouse model of chronic infection enriched in CD8^{+}CD39^{+} T cells. (A) Gene set enrichment analysis of a signature of 200 genes upregulated in exhausted CD8^{+} T cells from the mouse model of chronic viral infection versus acute infection (day 30 post infection) in the ranked list of genes differentially expressed in CD39^{+} vs. CD39^{-} CD8^{+} T cells. Leading edge genes are indicated by orange symbols. **(B)** Volcano plot of all genes (grey) or exhausted leading edge genes (orange).

5.5 Discussion

Findings:

1. **Unsupervised analysis of gene transcription profiles from $CD8^+$ T cells from donors with HCV infection demonstrate 2 distinct molecular phenotypes segregated by CD39 expression.**
2. **$CD39^+$ $CD8^+$ T cells in HCV infection are enriched with genes associated with T cell exhaustion in chronic infection from both human and animal models.**
3. **GSEA analysis demonstrates that markers of T cell proliferation are enriched in both $CD39^+$ $CD8^+$ T cells and exhausted $CD8^+$ T cells.**

The phenotype of T cell exhaustion is complex and demonstrates a distinct change in gene expression compared to memory $CD8^+$ T cells (153). The identification of the exact state of $CD8^+$ T cell exhaustion in the context chronic viral infection is important to aid in the development of successful and novel therapeutics. Unfortunately, use of techniques such as flow cytometry are limited by the number of parameters available. However, methods such as genome-wide expression profiling are able to provide an alternative means to broadly characterise the phenotype of T cell exhaustion, identify novel markers and biological pathways that are coordinately upregulated. I have demonstrated how genome-wide expression profiling identifies CD39-expressing $CD8^+$ T cells as associated with markers of T cell inhibition and pathways of T cell proliferation. These data suggest that CD39 expression on $CD8^+$ T cells is a marker of T cell exhaustion in the context of chronic HCV infection.

5.5.1 Unsupervised and supervised analyses

My previous data has demonstrated that CD39 is highly expressed in the viral-specific $CD8^+$ T cells in chronic viral infections such as HCV and HIV. However, genome-wide expression profiling was performed on effector $CD8^+CD39^+$ T cells from HCV infection due to the potential low yield of cell numbers from antigen-specific cells that may not have been sufficient for gene transcriptional profiling. However, consensus hierarchical clustering was able to distinguish 2 broadly different phenotypes of gene transcription profiles that segregated into $CD39^+$ and $CD39^-$ $CD8^+$ T cells. Whilst this is informative and highlights that there are differences between the 2 cell types worth investigating, it does not provide information on the specific individual or clusters of genes that contribute to these differences.

Analysis of gene expression profiles by comparing fold change allows the opportunity to identify genes that are specifically upregulated in $CD39^+$ and $CD39^-$ populations. Interestingly, my results demonstrate that genes that are associated with inhibition and exhaustion of T cells such as CTLA4 and PD-1 are upregulated in CD39 enriched effector CD8's. In addition, a nearest neighbour analysis, which clusters genes that are most closely correlated with regard to their patterns of gene expression, also identified PD-1 and CTLA4 as highly correlated with CD39. These data provides 2 independent analytical tools that associate markers of T cell inhibition with CD39 expression.

5.5.2 GSEA

Even though CD39 is enriched with 2 genes known to play a role in T cell exhaustion (PD-1 and CTLA4), further analysis is warranted to support the assumption that CD39 also plays a similar role in chronic viral infection. For example, PD-1 is also increased on activated T cells, therefore CD39 may represent an activated $CD8^+$ T cell only. In addition, both PD-1 and CTLA 4 may rise due to reasons related to factors such as sample selection and not associated with CD39 expression at all. Therefore, an additional analytical approach was used, known as GSEA. This technique identified groups of coordinately upregulated genes in the signature produced by $CD39^+$ T cells. Interestingly there was enrichment the $PD\text{-}1^{High}$ vs $PD\text{-}1^{In}$ signature. These data strongly suggest that CD39 is not just a marker of activation or T cell inhibition but more likely a marker of $CD8^+$ T cell exhaustion in chronic HCV infection.

The definition of T cell exhaustion has changed over the years with the availability of more complex methods of phenotyping these cells resulting in the discovery of multiple markers and functional states (152, 153). Additional analysis of gene expression profiles of $CD39^+$ versus $CD39^-$ $CD8^+$ T cells in HCV infection showed that the $CD39^+$ fraction was strongly enriched for genes related to proliferation. These signatures were also found to be enriched in the mouse model of exhaustion. This can be explained by data from the mouse model of chronic infection demonstrating that the survival of exhausted $CD8^+$ T cells is dependent on continuous antigenic exposure (187). This would explain the increased expression of genes associated with proliferation in $CD39^+$ $CD8^+$ T

cells in HCV infection and highlights a novel signature that has been identified in this subset of cells.

5.5.3 Conclusion

Gene expression profiling provides further supportive evidence that CD39 is a marker of T cell exhaustion in chronic HCV infection and the enrichment of its gene signature with exhaustion signatures suggests that CD39 may demarcate a particular subset of exhausted $CD8^+$ T cells. So far we may only speculate that since it is highly expressed on exhausted antigen-specific cells from chronic infection, its role is tightly linked to the dysfunction of $CD8^+$ T cells by providing further inhibitory signals via the purinergic pathway. Therefore, further interrogation of CD39 expression in human and animal models is warranted to identify which cellular pathways may be of interest to help restore T cell function in chronic viral infection.

CHAPTER 6

Phenotype of CD39 in acute and chronic LCMV infection in mice

6.1 Introduction

Elucidating the Immune responses to viral pathogens has long been a major goal for researchers as the disease burdens of chronic viral infections such as HCV, HBV and HIV have a significant impact on morbidity and mortality worldwide. Evolutionary mechanisms have emerged that allow for effective T cell responses to many viruses, which has served to broaden our understanding of effective adaptive T cell immunity where re-exposure of the same antigen results in a more rapid and robust response that quickly eliminates infection (188, 189). The differences between this and ineffective, dysfunctional T cell responses in chronic infections, where viral antigens persist, are key to the development of novel therapeutic vaccines and treatments. It has been possible to directly compare acute and chronic T cell responses to the same virus in certain diseases such as HCV, where 70-75% of patients will develop chronic infection and the remainder will completely eradicate virus and be cured. Such comparisons have helped identify cardinal differences such as the process of T cell exhaustion in chronic infection where dysfunctional T cells can be recognised by expressing a range of functional and phenotypic profiles such as increased expression of inhibitory receptors as well as reduced proliferation and cytokine secretion. However, the wide genetic heterogeneity in human donors has made it difficult to isolate individual pathways and phenotypes that are of key significance. As a result our understanding of effective T cell immunity is still lacking and we have yet to develop totally effective treatments to many of these chronic viral infections.

Chapter 6. Phenotype of CD39 in acute and chronic LCMV infection in mice

The study of acute Armstrong and chronic Clone 13 strains of LCMV infection in genetically identical mice has allowed mechanistic discovery in antigen presentation, T cell effector functions as well as memory formation (140). It was in these mouse models of infection where $CD8^+$ T cell exhaustion was first characterized and where virus-specific $CD8^+$ T cells were stratified into different degrees of exhaustion with progressive loss of effector cytokines, reduced cytotoxicity, loss of proliferation and eventual deletion in the most exhausted terminal effector $CD8^+$ T cells (64, 66, 190, 191).

Comparison of genome-wide transcription profiles of effector and memory $CD8^+$ T cells from acute LCMV infection with profiles of exhausted viral-specific $CD8^+$ T cells from chronic LCMV infection has demonstrated overexpression of key markers of exhaustion (153). The inhibitory receptor PD-1 was shown to be highly overexpressed in exhausted $CD8^+$ T cells in chronic LCMV infection, which is in keeping with previous data that recognizes PD-1 as central to regulating T cell exhaustion (192). However, additional inhibitory receptors that are markers of T cell exhaustion were also upregulated in chronic infection including 2B4, LAG-3 and CTLA-4 (153). This highlights that PD-1 alone does not determine the fate of $CD8^+$ T cell exhaustion, instead the number and combination of markers in different environmental contexts will all contribute to the state of exhaustion. It is this pattern in chronic viral infection that if determined could allow for novel, effective therapies to be generated. For example, in vivo blockade CTLA-4 or LAG-3 alone demonstrated no change in virus-specific responses during chronic LCMV infection but blockade of LAG-3 with PD-1 demonstrates a synergistic

improvement in virus-specific $CD8^+$ T cells (70). In fact it is now well recognised that progressive exhaustion of $CD8^+$ T cells in chronic viral infection results in an increase in the number of coregulated inhibitory molecules (70, 152). This mouse model of acute and chronic LCMV infection that has identified functional $CD8^+$ T cell exhaustion has been confirmed in human models of chronic infection. Upregulation of multiple inhibitory receptors, in particular PD-1 which when blocked results in an improved effector $CD8^+$ T cell response, is demonstrated in HCV (174, 179), HIV (89, 111) and HBV (193) infection. Therefore, the study of T cell exhaustion in mouse models of chronic viral infection have directly lead to the discovery of key markers of T cell dysfunction in chronic viral infection in humans. However, complete reversal of $CD8^+$ T cell exhaustion by blockade of these markers has so far been unachievable suggesting that there may still exist novel exhaustion phenotypes that are as of yet undiscovered.

CD39 (ENTPD1) is an ectonucleotidase originally identified as an activation marker on human lymphocytes, but has subsequently been shown to be a hallmark feature of regulatory T cells (101, 158, 159). CD39 hydrolyses extracellular ATP and ADP into adenosine monophosphate, which is then processed into adenosine by CD73, an ecto-5'-nucleotidase (178). Adenosine is a potent immunoregulator that binds to A2A receptors expressed by lymphocytes causing accumulation of intracellular cAMP, preventing T cell activation and NK cytotoxicity (147, 161). Loss of CD39 in Tregs markedly impairs their ability to suppress T cell activation, suggesting that the juxtacrine activity of CD39 serves to negatively regulate T cell function (101). This mechanism of T cell inhibition in

Tregs may not be isolated in the $CD4^+$ compartment. Previous studies have generally reported $CD8^+$ T cells as being $CD39^-$ (102, 158, 162, 182) but growing evidence suggests that $CD8^+$ T cells can also express CD39 and may act as a regulatory cell mediating inhibition of effector cells in the context of viral infection (102). The expression of this marker on exhausted T cells in chronic viral infection models such as LCMV has not been examined.

In this study I demonstrate that LCMV-specific $CD8^+$ T cells in mice infected with chronic clone 13 strain of LCMV exhibit a novel $CD39^{High}$ population of highly exhausted $CD8^+$ T cells that does not exist on $CD8^+$ T cells from mice infected with acute Armstrong strain of LCMV. In addition, this $CD39^{High}$ $CD8^+$ T cell population of cells demarcates severe exhaustion with high coregulation of the inhibitory receptors PD-1, 2B4 and LAG-3. Thus CD39 provides a specific, pathological marker of exhausted $CD8^+$ T cells in chronic viral infection in a mouse model of chronic viral infection.

6.2 Aims/hypotheses

Specific Aim 1 (SA1): Phenotype CD39 expression on $CD8^+$ T cells in acute Armstrong and chronic Clone 13 LCMV mouse model of infection.

Hypothesis 1: CD39 expression is increased on LCMV-specific $CD8^+$ T cells in chronic Clone 13 strain of LCMV infection compared to acute Armstrong.

Specific Aim 2 (SA2): Investigate the functional role of CD39 on $CD8^+$ T cells in the context of chronic LCMV infection in mice.

Hypothesis 2: In vivo inhibition of CD39 activity by POM 1 will increase functional capacity of $CD8^+$ T cells during chronic viral infection.

Specific Aim 3 (SA3): Characterise the phenotype and functional properties of $CD8^+$ T cells in CD39 null mice during chronic LCMV infection.

Hypothesis 3: CD39 null mice will demonstrate a reduced ability to control chronic LCMV infection compared to WT controls.

6.3 Methods

6.3.1 HLA Tetramers

Mouse MHC class I tetramers of H-2D^b complexed with LCMV $GP_{276-286}$ and GP_{33-41} as previously described (66, 122). Biotinylated complexes were tetramerised using allophycocyanin-conjugated streptavidin (Molecular Probes). The complete list of multimers can be found in supplemental materials (Appendix Table 9.4).

6.3.2 Antibodies and flow cytometry

The following fluorochrome-conjugated antibodies were used: anti-mCD8, anti-mCD4, anti-mCD3, anti-mCD39 anti-mCD244, anti-mPD-1, anti-mLag3, anti-mCD44(IM7), anti-mCD160 , anti-mIFN γ , anti-mTNF α (all from Biolegend). 5 million splenocytes were re-suspended in 50ml of staining buffer (PBS, 0.1% FBS) and stained with the LIVE⁄DEAD Fixable Aqua Dead Cell Stain Kit (Invitrogen) for exclusion of dead cells prior to staining with MHC Class I multimer for 10 minutes at room temperature to identify antigen-specific cells. Surface staining was performed for 25 minutes in the dark at 4°C and cells were fixed using 4% paraformaldehyde (Electron Microscopy Sciences) for 20 minutes prior to analysis. Intracellular staining was performed following surface stains and fixed and permeabilsed using the FoxP3/Transcription Factor Staining Buffer Set (eBioscience). Cells were sorted by BD FACS ARIA II and all other analyses were performed on BD LSR II and BD LSR Fortessa flow cytometers equipped

with FACSDiva v6.1. Gates were set using Full Minus One (FMO) controls. Data were analysed using Flowjo software (Treestar).

6.3.3 Mice and infections

All mice were used according to the Harvard Medical School Standing Committee on Animals and National Institutes of Animal Healthcare Guidelines. Wildtype C57BL/6J mice were purchased from The Jackson Laboratory and CD39 null mice kindly donated from the Robson laboratory, Beth Israel Hospital, Boston, USA. 6-8 week old female mice were infected with 2 x 10^5 p.f.u. LCMV-Armstrong intraperitonealy or 4 x 10^6 p.f.u. LCMV-Clone 13 intravenously and analysed at indicated time points by homogenising the spleen into a single cell suspension, ACK lysis of red blood cells, followed by antibody staining. Viruses were propagated by triple-plaque purification on Vero cells and grown in BHK-21 cells (128). Virus stocks at passage 1 or 2 were harvested for experiments. Virus was obtained from the Wherry laboratory, University of Pennsylvania, PA, USA. In vivo inhibition of CD39 with POM 1 in Wild type mice was achieved by administering 10mg/kg/day of the CD39 inhibitor sodium polyoxotungstate (POM 1) (n=5) or, 1 ml of placebo (n=5) water intra-veneously at Day 0 after infection with chronic clone 13 strain of LCMV.

6.3.4 Peptide stimulations

Single cell splenocyte suspensions were stimulated with purified peptide at 1/100 concentration diluted with RPMI and supplemented with 5% FBS (Hyclone, Logan, UT), 20mM HEPES, L-glutamine, antibiotics and 2μg/ml Brefeldin A (Sigma) at 37°C, 6% CO_2 for 6 hours. After stimulation, cells were stained with surface antibodies first followed by intracellular antibodies as above according to the manufacturers' guidelines using fixation and permeabilisation buffers. Once staining was complete samples were fixed in PBS + 4% formaldehyde and analysed. Non-specific cytokine production after incubation with medium alone was subtracted using unstimulated isotype controls.

6.4 Results

6.4.1 CD39 is increased in exhausted $CD8^+$ T cells in the mouse model of chronic viral infection

Because the mouse signature of $CD8^+$ T cell exhaustion was significantly enriched in the transcriptional prolife of $CD39^+$ $CD8^+$ T cells in HCV-infected patients, I next asked if CD39 is upregulated by $CD8^+$ T cells in the mouse model of chronic viral infection. To address this question I compared two well described mouse models of viral infection using two strains of LCMV, Armstrong causing acute infection that is resolved in up to 8 days and Clone 13 that persists in mice for up to 3 months. Using flow cytometry I quantified CD39 expression and compared it to that of the central marker of $CD8^+$ T cell exhaustion, PD-1 (107, 153).

Despite the percentage expression of CD39 on antigen experienced $CD8^+$ T cells ($CD44^+$) from mice infected with both Clone 13 and Armstrong strains of LCMV being high (Figure 6.1A-C), there was a statistically significant difference at days 8 and 15 post infection with clone 13 demonstrating higher levels of CD39 (mean 98.9% day 8 and 95% at day 15). However, there was no significant difference in CD39 expression at Day 36 post infection (p.i.) between the 2 groups even though the mean percentage for CD39 expression was higher in chronic clone 13 infection (95.4% clone 13 and 92.3% Armstrong) (Figure 6.1D). Having discovered that CD39 is increased in LCMV-specific $CD8^+$ T cells from mice with chronic infection when compared with acute, I next compared PD-1 expression on CD39 expressing $CD8^+$ T cells in both the LCMV-specific

population and bulk CD8's using multiple tetramers (Figure 6.2A). I found that CD39 expressing cells are significantly enriched with PD-1, more so in the virus-specific pool compared with the total $CD8^+$ population (mean 47.5% $Tetramer^+$ compared with 27.5% $CD8^+$), supporting the suggestion that CD39 is a marker of T cell exhaustion in chronic viral infection (Figure 6.2B)

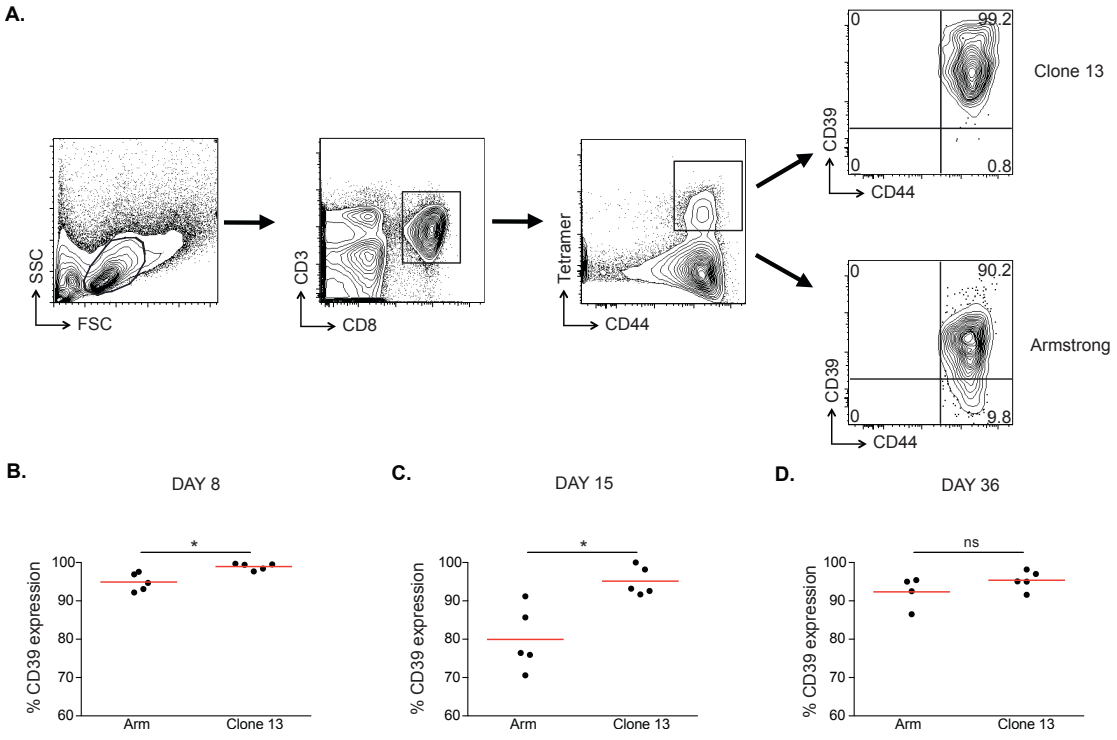


Figure 6.1. Mouse model of viral infection demonstrates higher percentage of CD39 expression on viral-specific $CD8^+$ T cells in chronic clone 13 compared with acute Armstrong. (A) Flow cytometry analysis demonstrating higher percentage of CD39 expression on LCMV-specific $CD8^+$ T cells from mice infected with chronic clone 13 strain of LCMV compared with acute Armstrong. **(B-D)** Scatter plots comparing percentage of CD39 expression on LCMV-specific $CD8^+$ T cells at days 8, 15 and 36 after infection with Armstrong (Arm) or Clone 13 strains of LCMV. Data are

representative of 1 experiment of 5 mice per group. Statistical significance was assessed with unpaired student's t-test. ns=not significant, *P < 0.5.

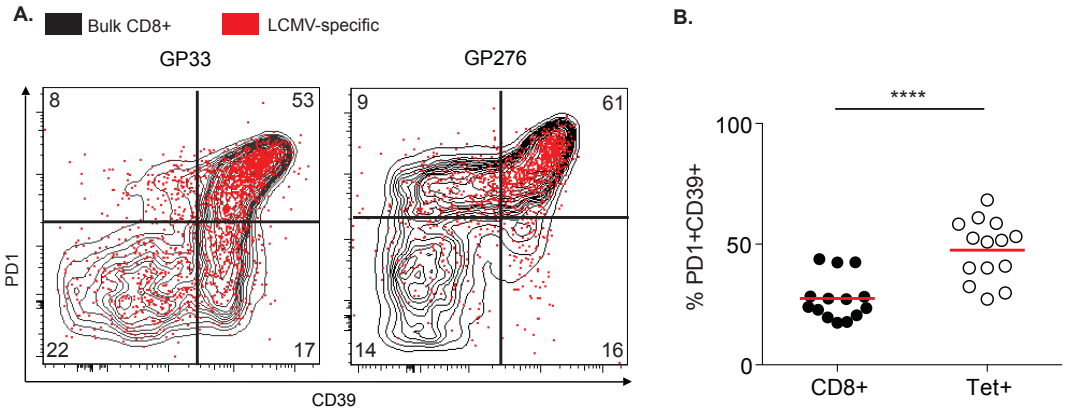


Figure 6.2. CD39 population is enriched with PD-1 in LCMV-specific CD8$^+$ T cells in chronic clone 13 infection. (A) Flow cytometry analysis demonstrating high degree of PD-1 and CD39 co-expression on LCMV-specific CD8$^+$ T cells in chronic clone 13 LCMV infection. **(B)** Scatter plot demonstrating significantly increased expression of PD-1 on CD39 expressing LCMV-specific CD8$^+$ T cells compared with bulk CD8's in chronic clone 13 LCMV infection. Data are representative of 1 experiment of 14 mice across 2 tetramers (GP33, GP276) at day 31 p.i. Statistical significance was assessed with paired student's t-test. ****P < 0.0001.

6.4.2 Splenic CD8$^+$ T cells from mice infected with chronic clone 13 strain of LCMV demonstrate a CD39^{High} population that represents severe T cell exhaustion

Having identified that CD39$^+$ CD8$^+$ T cells are enriched with PD-1 and thus likely represents a marker of T cell exhaustion, I next compared the expression of CD39 in acute and chronic LCMV infection by mean fluorescent intensity (MFI) to

investigate different subsets of $CD8^+$ T cells identifiable by varying CD39 expression alone. I found that splenic $CD8^+$ T cells from mice with clone 13 LCMV infection demonstrated a $CD39^{High}$ subset of cells not present in mice infected with acute Armstrong (Figure 6.3A). When comparing the MFI of CD39 expression on $CD8^+$ T cells form Armstrong and Clone 13 infected mice, there was a significant difference in expression values (Figure 6.3A, 6.3B). Therefore, I examined whether this $CD39^{High}$ population of $CD8^+$ T cells, specifically marked a highly exhausted phenotype in chronic viral infection. To investigate this I used flow cytometric analysis to compare the expression of 3 inhibitory receptors (PD-1, 2B4 and LAG-3) on the surface of $CD39^{High}$ and $CD39^{Int}$ LCMV-specific $CD8^+$ T cells (Figure 6.4A). My results demonstrate that there was a significant enrichment in PD-1 expression in the $CD39^{High}$ population of cells (Figure 6.4B). Comparison of 2 other markers of T cell exhaustion, 2B4 and LAG-3 (Figure 6.4C) demonstrated that $CD39^{High}$ $CD8^+$ T cells were significantly enriched in multiple inhibitory receptors, whereas $CD39^{Int}$ cells were lacking in the same markers of T cell exhaustion (Figure 6.4D, 6.4E).

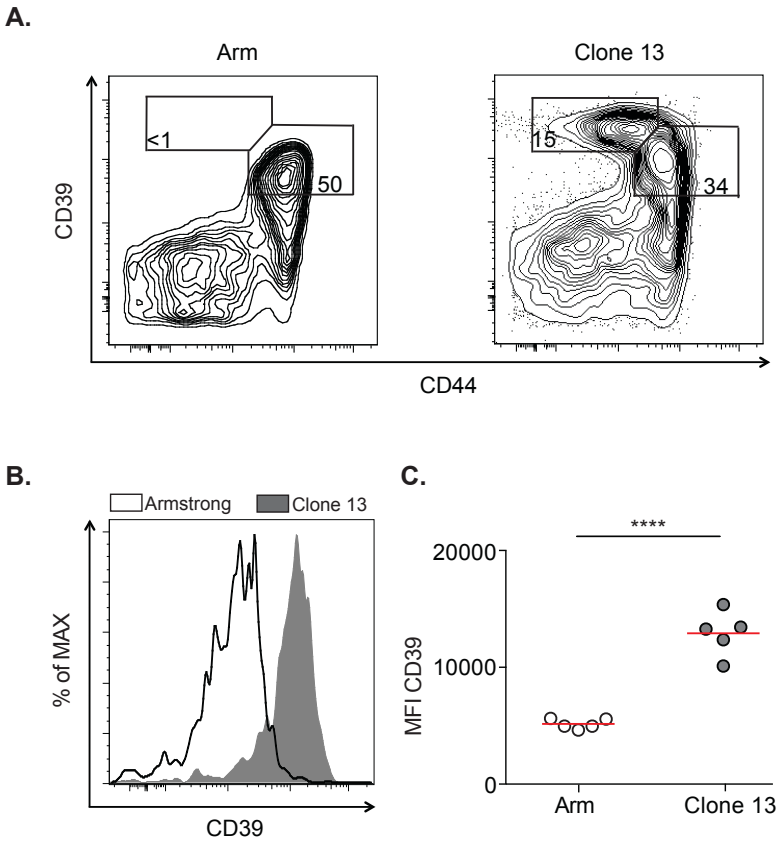


Figure 6.3. CD8$^+$ T cells from mouse spleens with chronic LCMV infection exist as 2 distinct populations of CD39 expressing cells. (A) FACS plots showing expression of CD39 on CD44$^+$CD8$^+$ T cells in spleens of mice following LCMV Armstrong (left) or clone 13 (right) infection and indicating CD39^{High} and CD39^{Int} CD8$^+$ T cells. **(B)** Representative histogram of CD39 expression on CD8$^+$ T cells from chronic clone 13 and acute Armstrong infected mice on day 36 p.i. (C) Scatter plot demonstrating significantly increased MFI of CD39 on CD8$^+$ T cells in chronic clone 13 infection compared with Armstrong. Data are representative of 1 experiment of 5 mice per group across 2 at day 36 p.i. Statistical significance was assessed with unpaired student's t-test. ****P < 0.0001.

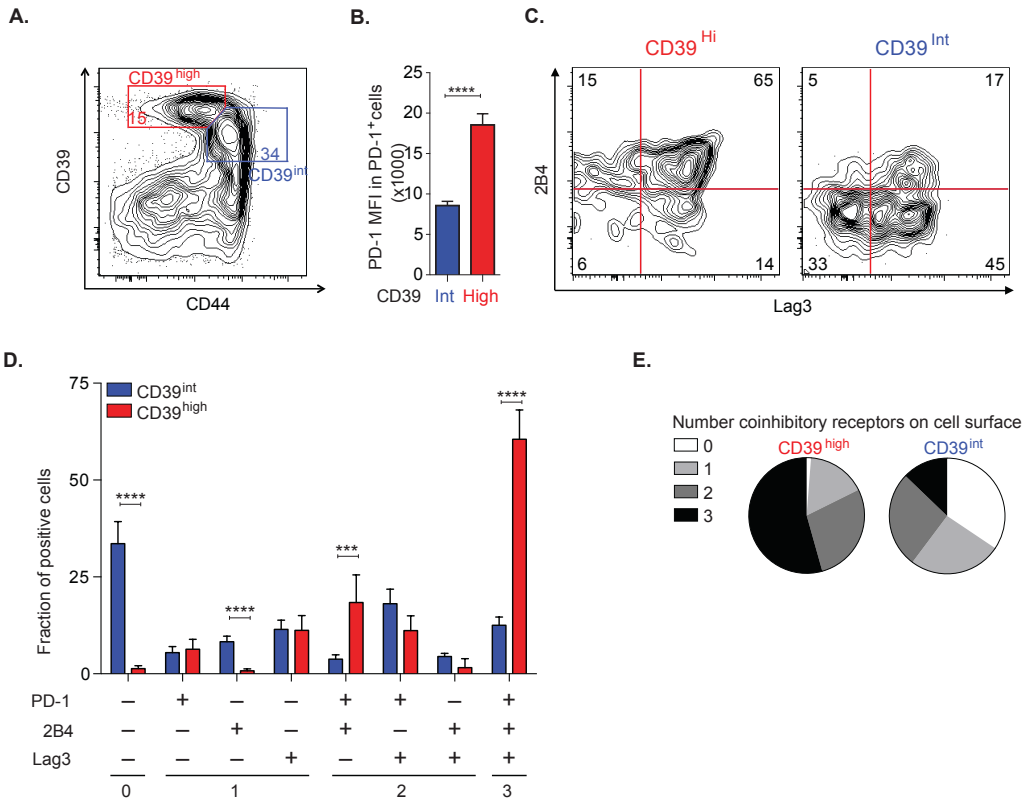


Figure-6.4. CD39^{Hi} cells express multiple inhibitory receptors compared with CD39^{Int} LCMV-specific CD8^{+} T cells in chronic clone13 LCMV infection. (**A**) FACS plot demonstrating expression of CD39 on CD44^{+} CD8^{+} T cells in spleens of mice 36 days following Clone 13 infection and with CD39^{high} (red) and CD39^{int} (blue) CD8^{+} T cells. **(B)** Histogram demonstrating increased MFI of PD-1 in CD39^{High} compared with CD39^{Int} CD8^{+}T cells **(C)** Flow cytometry analysis showing the difference in patterns of inhibitory receptor expression (2B4 and LAG-3) between CD39^{high} and CD39^{int} on LCMV-specific CD8^{+} T cells. **(D**) Fraction of CD39^{high} and CD39^{int} CD44^{+}CD8^{+} T cells expressing different combinations of co-inhibitory receptors PD-1, 2B4, and LAG-3 and (**E**) average number of co-inhibitory receptors expressed on CD39^{int} and CD39^{high} CD8^{+} T cells at d36 p.i. with LCMV Clone 13. Data are representative of 2 experiments of 5 mice per group. Statistical significance was assessed with with Holm-Sidak multiple comparison correction (C). ***P < 0.001, ****P < 0.0001.

6.4.3 In vivo administration of the CD39 inhibitor POM 1 to mice with chronic clone 13 LCMV infection

Having discovered the phenotype of CD39 as being highly expressed in viral-specific $CD8^+$ T cells in chronic LCMV infection as well as $CD39^{High}$ $CD8^+$ T cells being enriched with multiple markers of T cell exhaustion, I next sought to investigate the functional role of CD39 in the context of chronic viral infection. Several well known molecules have been identified as effective inhibitors of CD39 activity (149, 194). I chose to use the CD39 inhibitor POM 1 in vivo to test whether the inhibition of CD39 activity alters the phenotype and/or function of LCMV-specific $CD8^+$ T cells during chronic viral infection in mice. After administration of POM 1 or water controls to mice infected with chronic clone 13 strain of LCMV, I used flow cytometry to analyse splenic $CD8^+$ T cells (Figure 6.5A). I first examined the level of PD-1 expression in both $CD39^{High}$ and $CD39^{Int}$ LCMV-specific $CD8^+$ T cells, but found no significant difference despite using multiple tetramers to isolate LCMV-specific $CD8^+$ T cells (GP33 and GP276) (Figures 6.5 B-E).

I next examined the expression of multiple inhibitory receptors (PD-1, 2B4, CD160 and LAG-3) that are implicated in T cell exhaustion (Figures 6.6 A-I) (77). My results demonstrate no significant differences in the expression of these markers across 2 LCMV-specific tetramers used, except for PD-1 (Figure 6.6F) showing a significant increase in the POM 1 group and LAG-3 increased in controls (Figure 6.6I). With little evidence that in vivo inhibition alters the profile of inhibitory receptor expression on LCMV-specific $CD8^+$ T cells, I next sought to

identify functional differences between the 2 samples groups. 28 days p.i. and after peptide stimulation I used flow cytometry to compare cytokine secretion of $CD8^+$ T cells in POM 1 administered and water control groups (Figure 6.7FA). My results demonstrated no significant change in IFN γ and/or TNF α secretion in all groups (Figure 6.7B-G) except for GP276 stimulated cells that demonstrated a mild significant increase in TNF α (mean 0.78%in POM 1 and 0.48% in control groups) in the POM 1 group (Figure 6.7F).

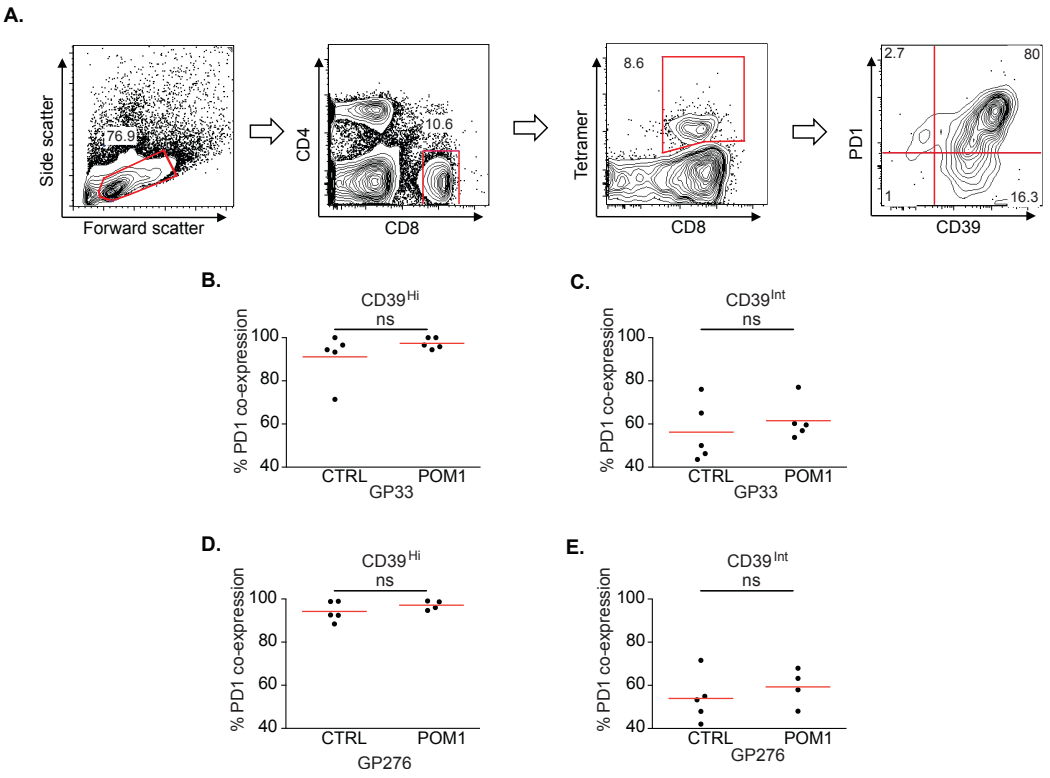


Figure-6.5. In vivo inhibition of CD39 by POM 1 does not alter expression of PD-1 on CD39^{High} or CD39^{Int} LCMV-specific CD8^{+} T cells in mice with chronic clone 13 infection. (A) Flow cytometry analysis of splenic LCMV-specific CD8^{+} T cells from mice infected with chronic clone 13 strain of LCMV. Mice were either administered the CD39 inhibitor POM 1 or water controls and analysed 28 days after infection. **(B-E)** Scatter plots demonstrating percentage of PD-1 expression on CD39^{High} (left) and CD39^{Int} (right) LCMV-specific CD8^{+} T cells from mice in (A). Samples were analysed across 2 tetramers (GP33 (B,C) and (GP276 (D, E)). Data are representative of 1 experiment of 5 mice per group at day 28 p.i. Statistical significance was assessed with unpaired student's t-test.ns = not significant.

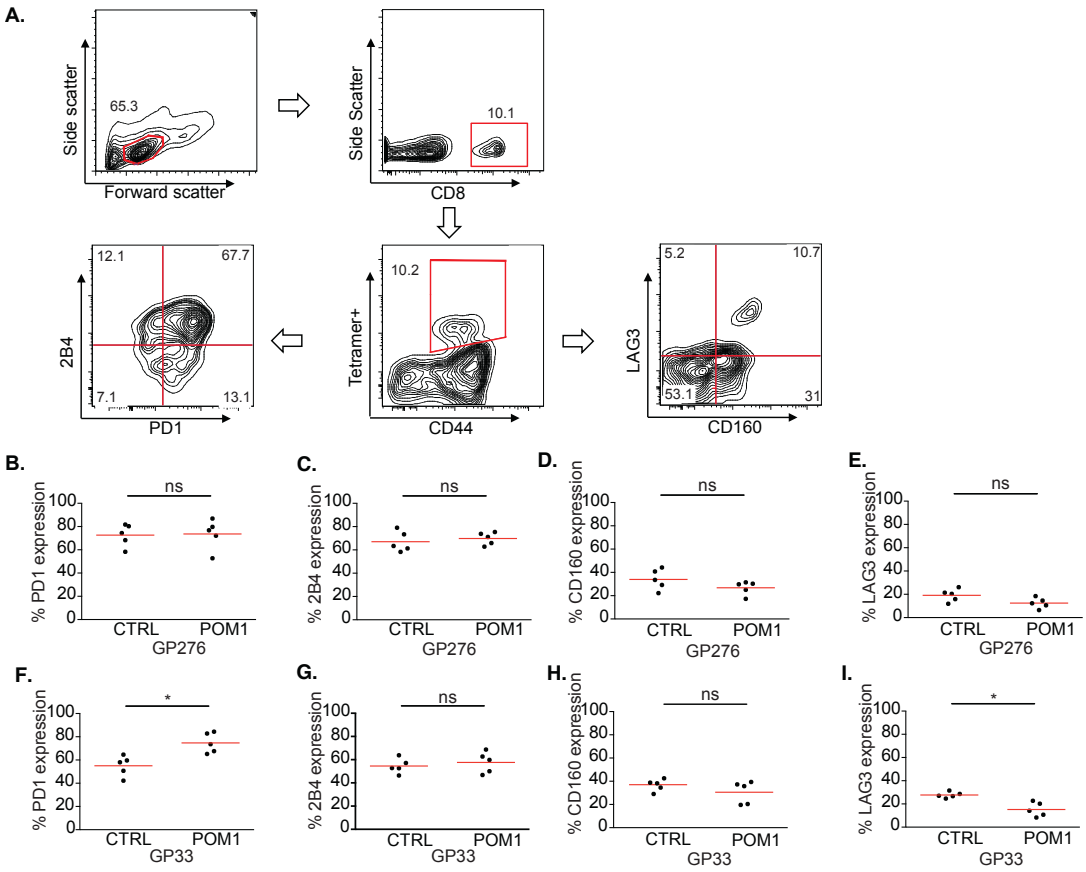


Figure-6.6. In vivo inhibition of CD39 by POM 1 does not alter expression of exhaustion markers on LCMV-specific CD8$^+$ T cells in mice with chronic clone 13 infection. Splenocytes from mice infected with Clone 13 were analysed by flow cytometry to quantify expression levels of exhaustion markers on CD8$^+$ T cells**. (A)** Representative FACS plots of mouse splenocytes analysed after infection with clone 13 strain of LCMV and then injected with POM 1 or, water **(B-I)** Scatter plots demonstrating little to no significant variation of exhaustion marker profile of CD8$^+$ T cells from LCMV infected mice with POM 1 injections compared with water controls. Data are representative of 1 experiment of 5 mice per group at day 28 p.i. Statistical significance was assessed with unpaired student's t-test. ns = not significant, *P < 0.5.

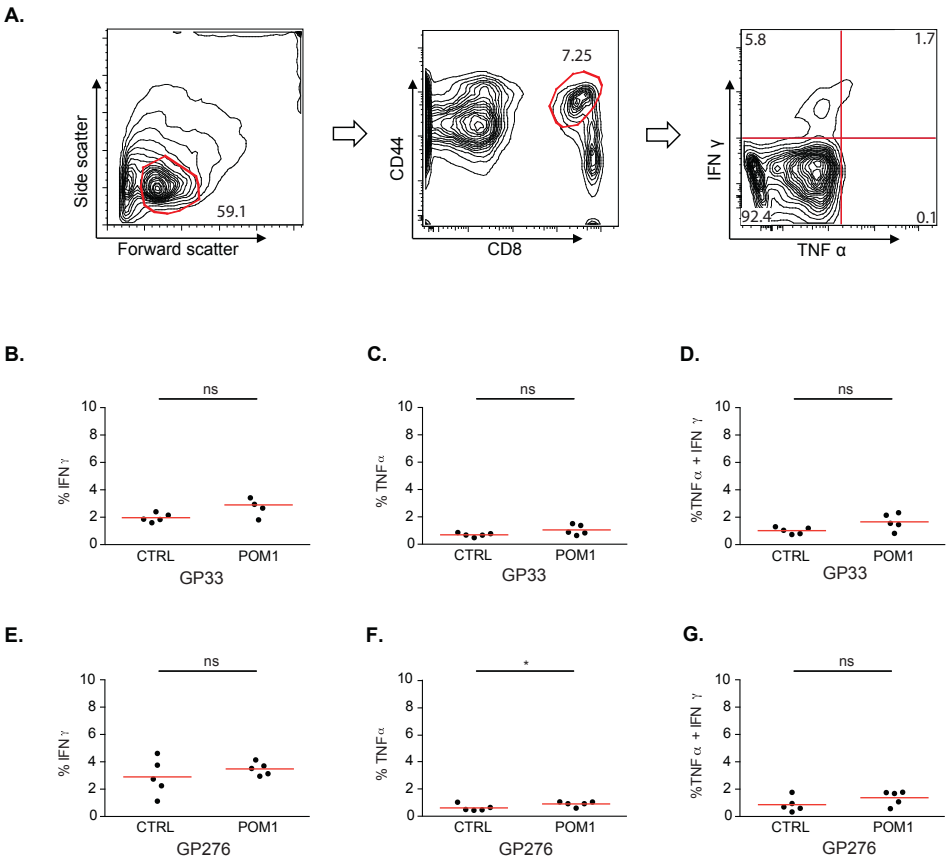


Figure-6.7. In vivo inhibition of CD39 by POM 1 does not alter secretion of IFN γ or TNF α secretion from CD8$^+$ T cells in mice with chronic clone 13 infection. Splenocytes from mice infected with Clone 13 were analyzed by flow cytometry to quantify functional capacity of CD8$^+$ T cells**. (A)** Representative FACS plots of mouse splenocytes analysed after infection with clone 13 strain of LCMV and stimulated for 6 hours with gag peptides GP 33 and GP 276 **(B-G)** Scatter plots demonstrating little to no significant variation of IFN γ or TNF α secretion from CD8$^+$ T cells from LCMV infected mice with POM 1 injections compared with water controls except for TNF α (mean 0.76% in POM 1 group, 0.48% in controls). Data are representative of 1 experiment of 5 mice per group at day 28 p.i. Statistical significance was assessed with unpaired student's t-test. ns=not significant, *P < 0.5.

6.4.4 Mice deficient in CD39 (CD39KO) demonstrate increased mortality in chronic infection compared with WT

To further investigate the functional significance of CD39 expression on $CD8^+$ T cells, I next sought to examine the phenotypic and functional characteristics of $CD8^+$ T cells in CD39 deficient mouse during chronic LCMV infection. After infection with chronic clone 13 strain of LCMV, 3 out of 5 mice deficient of CD39 died beginning form 7 days p.i (Figure 6.8A). Of the remaining 2 CD39KO and 6 WT mice, I used flow cytometric analysis to determine the expression profile of the inhibitory receptors PD-1, 2B4, CD160 and LAG-3 (Figure 6.8B, 6.8C). Due to insufficient numbers statistical analysis was unable to be performed. However, across all markers, the mean level of inhibitory receptor expression was increased in WT mice compared with CD39 deficient mice (WT mean values for PD-1, 2B4, CD160, LAG-3 are 76%, 54%, 28%, 31% respectively and CD39KO, 41%, 26%, 15%, 17%) (Figure 6.8 D-G). Further investigation to test the functional capacity of $CD8^+$ T cells from these groups of mice was performed by using flow cytometry analysis of IFN γ and TNF α cytokine secretion profiles after peptide stimulation of cells (Figure 6.9A). Again, due to insufficient numbers statistical analysis was unable to be performed. However, the mean level of cytokine expression was increased in CD39KO compared with WT mice (mean values for CD39KO after GP33 stimulation for IFN γ and TNF α are 11.7%, 14.7% respectively compared to 8.3%, 2.5% in WT mice) (Figure 6.9B-G). These data point towards CD39 deficiency as perhaps reducing the inhibitory capacity of $CD8^+$ T cells during chronic viral infection.

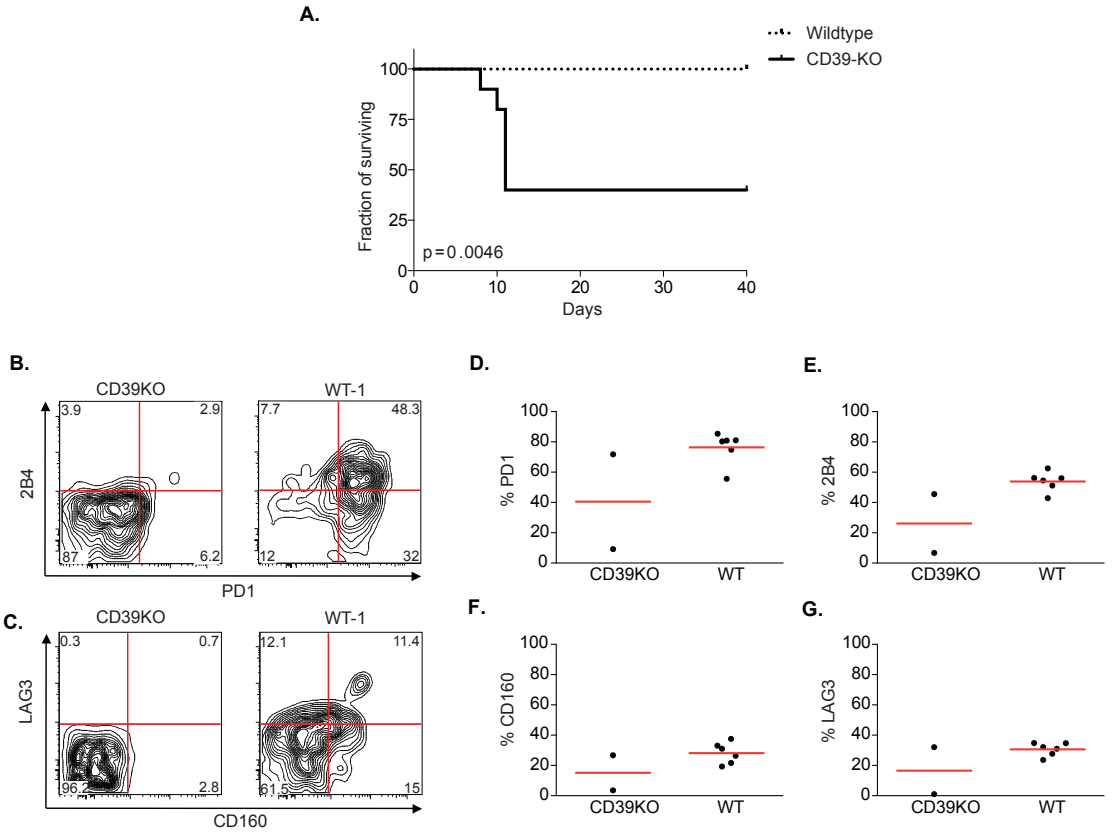


Figure-6.8. CD39 deficient mice exhibit increased mortality to LCMV clone 13 infection and may demonstrate reduced inhibitory receptor expression compared with Wild Type (WT). (A) Kaplan-meier survival graph demonstrating increased mortality of CD39KO mice compared with Wildtype counterparts 7 days after infection with chronic clone 13 strain of LCMV. **(B,C)** Scatter plots demonstrating exhaustion marker expression of CD39KO and Wild type mice 30 days p.i. **(D-F)** Scatter plots demonstrating percentage inhibitory marker expression on LCMV-specific $CD8^+$ T cells from CD39KO and WT mice 30 days after infection with clone 13 LCMV infection. Data are representative of 1 experiment of 5 or 6 mice per group. Statistical significance not calculated due to insufficient numbers for comparison.

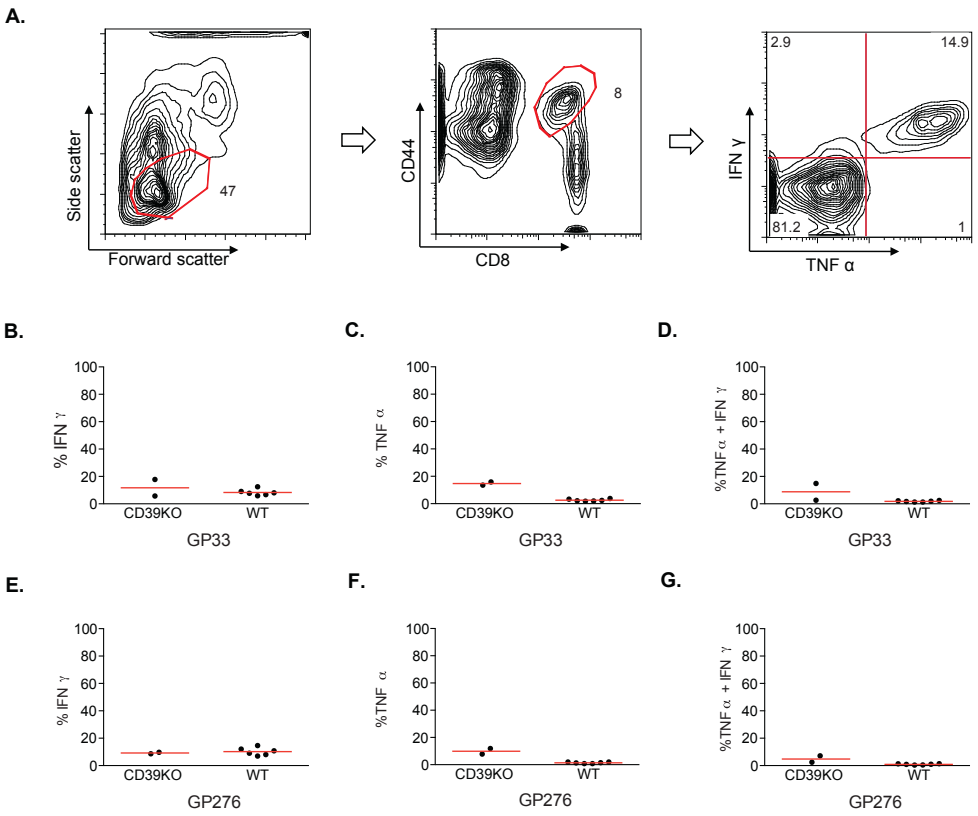


Figure-6.9. Functional analysis of splenic CD8$^+$ T cells from CD39KO and WT mice after infection with chronic clone 13 strain of LCMV. Splenocytes from mice infected with Clone 13 were analysed by flow cytometry to quantify the functional capacity of CD8$^+$ T cells**. (A)** Representative FACS plots of mouse splenocytes analysed 30 days after infection with clone 13 strain of LCMV and stimulated for 6 hours with gag peptides. **(B-G)** Scatter plots of CD8$^+$ T cells stimulated with GAG peptides GP 33 (B-D) and GP 276 (E-G) demonstrating possible increase in IFN γ and TNF α secretion from CD8$^+$ T cells of CD39KO mice compared with WT. Data are representative of 1 experiment of 5 (CD39KO) or 6 (WT) mice per group at day 30 p.i. Statistical significance not calculated due to insufficient numbers for comparison as 60% of mice died from CD39KO group.

6.5 Discussion

Findings:

1. **CD39 is highly expressed on viral-specific $CD8^+$ T cells in chronic clone 13 strain of LCMV infection in mice.**
2. **CD39 is more highly expressed on viral-specific $CD8^+$ T cells in chronic clone 13 strain of LCMV than in acute Armstrong.**
3. **$CD8^+$ T cells from mice with chronic LCMV infection demonstrate 2 distinct subsets of CD39 expression referred to as $CD39^{High}$ and $CD39^{Int}$.**
4. **$CD8^+CD39^{High}$ cells demonstrate the highest levels of PD-1 expression and multiple markers of T cell exhaustion, whereas $CD8^+CD39^{Int}$ express significantly lower levels of inhibitory receptors.**
5. **CD39 null mice demonstrate increased mortality after infection with chronic clone 13 strain of LCMV.**
6. **$CD8^+$ T cells from CD39 null mice demonstrate reduced expression of inhibitory markers and increased cytokine secretion.**
7. **In vivo administration of the CD39 inhibitor POM 1 does not alter $CD8^+$ T cells state.**

The phenotype of $CD8^+$ T cell exhaustion has been very well described in the LCMV mouse model of chronic infection. In particular, the pattern of inhibitory receptor expression appears to play an important role in deciding the fate of

$CD8^+$ T cells (152). In addition, these discoveries have been recapitulated in human models of chronic infection such as HCV (174, 179), HIV (89, 111) and HBV (193). However, despite these key insights, the complete reversal of T cell exhaustion is still unachievable, highlighting our lack of knowledge of key pathways and differentiation states have yet to be discovered. T cell inhibition mediated by purinergic pathways through the hydrolysis of extracellular ATP into adenosine is a growing area of research and within lymphocytes, the main concentration has been toward $CD4^+$ T cells. However, expression of the key regulator of ATP hydrolysis, the ectoenzyme CD39 has not been studied on viral-specific $CD8^+$ T cells in the context of chronic viral infection. Here, I demonstrate that not only is CD39 highly expressed on viral-specific $CD8^+$ T cells in chronic clone 13 strain of LCMV infection compared to acute Armstrong strains, but also the discovery of a $CD39^{High}$ population. This novel subset of $CD8^+$ T cells express the highest levels of the inhibitory receptor PD-1 as well as multiple markers of T cell exhaustion. In addition, CD39 null mice displayed an increased mortality as well as increased $CD8^+$ T cell functional capacity compared with WT mice after infection with chronic LCMV infection. These data suggest that CD39 expression plays an inhibitory role on $CD8^+$ T cells during chronic LCMV infection and that $CD39^{High}$ $CD8^+$ T cells marks a severe exhaustion state.

6.5.1 CD39 is a marker of severe exhaustion in chronic LCMV infection

There are a number of features that indicate that CD39 is a marker of T cell exhaustion. Firstly, the increased expression of CD39 on LCMV-specific $CD8^+$ T cells from mice with chronic clone 13 infection compared with acute Armstrong highlights a significant enrichment of CD39 in exhausted viral-specific $CD8^+$ T cells as most LCMV-specific cells identified by the tetramers GP33 and GP276 are highly exhausted in clone 13 infection. This appears to be a reliable finding since the 2 strains of LCMV allows for an extremely close comparison of the viral-specific response during acute and chronic infection, which can be difficult to find in human models. Secondly, it is well recognised that PD-1 expression is the central hallmark of T cell exhaustion in chronic LCMV infection (192). These findings have been translated to human models of chronic infection such as HCV with in vitro and in vivo experiments using monoclonal antibodies to inhibit PD-1 demonstrating signs of improvement in $CD8^+$ T cell responses (86, 174, 179). My results demonstrate that PD-1 is highly enriched in $CD39^+CD8^+$ T cells during chronic viral infection thus highlighting how CD39 expression mimics PD-1 expression in exhausted T cells. Thirdly, it is well recognised that other surface receptors such as LAG-3 and 2B4 play a significant role in combination with PD-1 to modulate the T cell response to viral antigen (153). However, they can also contribute to the pathological process of T cell exhaustion. I have identified a novel subset of $CD39^{High}$ $CD8^+$ T cells in chronic clone 13 LCMV infection that is not present in acute Armstrong and expresses significantly higher levels of PD-1 and number of inhibitory receptors than $CD39^{Int}$ $CD8^+$ T cells.

These data support CD39 ^{High} as demonstrating a severe exhaustion phenotype on $CD8^+$ T cells in chronic LCMV infection in mice.

6.5.2 In vivo inhibition of CD39

The functional role of CD39 has primarily revolved around its action as a hydrolyser of ATP to cAMP, thereby facilitating the production of adenosine via CD73's action of hydrolysing cAMP. Various inhibitors of CD39 have been described (162, 188, 189) although the most reliable used in vivo is POM 1 (140). Administration of POM 1 has been shown to alter cellular responses in mice, however my results demonstrate very little change in the way of $CD8^+$ T cell expression profile of exhaustion markers such as PD 1, CD160, 2B4 and LAG-3 as well as secretion of the cytokines IFN γ and TNF α. It is possible that the lack of significant findings may have been due to in adequate dose of POM 1 to induce an effect. However, the inhibition of CD39 in vivo would not have been specific to $CD8^+$ T cells expressing CD39. Therefore, due to the widespread expression of CD39 on multiple cell types any results could not be accurately attributed to $CD8^+$ T cells.

6.5.3 CD39KO mice during chronic LCMV infection

The initial studies of LCMV infection on mice deficient of PD-1 demonstrated an increased mortality due to an overactive immune response resulting in immunopathologic damage and death (107). This key study highlighted PD-1's role as an inhibitory receptor that was essential to providing a

balance between T cell inhibition and allowing an adequate functional response to combat pathogens. This eventually led to the discovery of PD-1's pathological role in T cell exhaustion, which is still under heavy investigation today. Using an analogous technique I infected CD39 null mice with chronic clone 13 strain of LCMV and observed a similar response. My data demonstrate that mice begin to die at day 7 resulting in only 40% survival. Analysis of $CD8^+$ T cells from those that did survive demonstrated increase in mean cytokine production as well as reduced expression of inhibitory receptors PD-1, CD160, 2B4 and LAG-3. Although these findings were not statistically significant due to lack of biological replicates secondary to increased mortality, it does highlight the possibility that CD39 may play a significant role in T cell inhibition during chronic LCMV infection. The lack of CD39 may drive an overactive immune response causing death of mice in a similar manner to PD-1 deficient mice. In addition, the reduced expression of inhibitory receptors on LCMV-specific $CD8^+$ T cells in CD39KO mice may mean that the presence of CD39 is required for these markers to be expressed. In any case further experiments at earlier time points would be required to confirm these findings with adequate biological replicates and the cause of death by severe lymphocytic infiltrate would need to be assessed as well. In addition, it is important to realise that these findings would be difficult to attribute to a lack of CD39 expression on $CD8^+$ T cells alone, as CD39 is deficient on all cell populations in these CD39 null mice including $CD4^+$ regulatory T cells.

6.5.4 Conclusion

CD39 expression appears to be a novel marker of T cell exhaustion in chronic viral infection and may serve to inhibit $CD8^+$ T cell function via the purinergic pathway involving the action of CD73 and adenosine. It is important to realise however, that the expression of CD39 and CD73 may differ in mouse and human models and therefore mechanisms of T cell inhibition could differ greatly. Nevertheless, this discovery moves researchers one step closer to developing therapeutic interventions to chronic viral infections, such as vaccinations in combination with other immune-boosting modalities acting as CD39 inhibitors that may provide synergistic effects.

Chapter 7

Discussion

7.1 Introduction

Since the discovery of $CD8^+$ T cell exhaustion defining its exact phenotypic profile in the context of different disease conditions has been a challenging task for immunologists. However, great progress has been made with advances in multi-parameter flow cytometry and the use of genome-wide transcription profiling. The latter has allowed the identification of distinct differences in gene expression patterns of exhausted $CD8^+$ T cells when compared to healthy functional memory $CD8^+$ T cells. (153). In addition the rising complexity of data has required more advanced computational approaches to extract meaningful, biologically relevant information. Such analyses have been able to identify novel phenotypic subclasses of T cells (136) and the discovery of markers that play a critical role in T cell dysfunction. I demonstrate how a systems biology approach utilising genome-wide expression profiles can identify novel subclass-specific markers such as CD39 and be applied to single cell analyses of virus-specific $CD8^+$ T cells. I show that CD39 is highly expressed in virus-specific $CD8^+$ T cells in human (HCV and HIV) and mouse (LCMV) models of chronic viral infection when compared to $CD8^+$ T cells of healthy donors. Furthermore, I demonstrate how CD39 expression correlates with clinical parameters such as viral load and in the mouse model of LCMV infection, $CD39^{High}CD8^+$ T cells express the highest levels of PD-1 and co-express multiple inhibitory receptors. These data indicate that CD39 expression is a marker of severe $CD8^+$ T cell exhaustion in chronic viral infection.

7.2 Identification of the subclass-specific marker CD39

The discovery of CD39 was achieved primarily through the interrogation of genome-wide expression profiles of HCV-specific $CD8^+$ T cells mimicking the path taken of other gene discoveries utilising genomic analyses of virus-specific $CD8^+$ T cells such as BIM in chronic HBV infection and BATF in HIV (109, 110). The key concept in this analysis was the use of an unsupervised analysis that segregated samples into novel subclasses first and then performing an analysis of the genes that are differentially expressed between the subclasses. This did not just involve choosing from a list of rank-ordered genes that were differentially expressed between the 2 subclasses of samples. Instead I chose CD39 as it was a marker identified within a network of 87 genes that were associated with known biological processes and was a key enzyme involved in the purinergic pathway which modulates T cell inhibition. Choosing from a rank-ordered list of differentially expressed genes would have made the discovery of CD39 much more challenging. Therefore, breaking down large data sets into novel subclasses of samples and novel gene clusters with biological context is a valuable technique in immunological discovery.

7.3 Increased CD39 expression in chronic viral infection

My analysis of CD39 by flow cytometry revealed that it was highly expressed in HCV and HIV-specific $CD8^+$ T cells as well as in LCMV-specific $CD8^+$ T cells from mice with chronic clone 13 LCMV infection. However, there were very low levels of CD39 expression on $CD8^+$ T cells from healthy human

donors and on virus-specific $CD8^+$ T cells of latent infections such as CMV and EBV. This is a novel discovery as CD39 has never been characterised on these cell populations and suggests that the expression of CD39 is pathological and related to the development of T cell exhaustion. The fact that a larger fraction of HCV-specific $CD8^+$ T cells express CD39 than do HIV-specific $CD8^+$ T cells may be related to differences in the timing of blood sampling during the course of infection, or may be due to differences in the extent of antigen-load and inflammation in the two infections. My further analysis of CD39 on $CD8^+$ T cells from mice with chronic clone 13 infection identified 2 distinct subpopulations of CD39 expressing cells, one expressing high amounts of CD39 ($CD39^{High}$) and the other expressing intermediate levels ($CD39^{Int}$). $CD39^{High}$ $CD8^+$ T cells were completely absent in mice with acute clone 13 LCMV infection further supporting CD39's pathological role in chronic viral infection. One area that could be improved upon in future experiments is the addition of CD127 as a marker of exhausted viral-specific T cells. This is particularly important in chronic HCV infection where tetramer responses could be from epitopes viral escape mutations have already occurred.

7.4 CD39 is a marker of severe $CD8^+$ T cell exhaustion

There are a number of features that support CD39 expression as a marker of severe $CD8^+$ T cell exhaustion in chronic viral infection. Firstly, CD39 is highly co-expressed with the inhibitory receptor PD-1, which is a central marker of exhausted T cells in chronic viral infection and rises with antigen load (107, 111,

153). Second, CD39 expression correlates with viral load in subjects with HIV and HCV infection suggesting that the pro-inflammatory environment associated with high titers of circulating virus is induces a rise in CD39 expression in the virus-specific pool of $CD8^+$ T cells (111, 180). Thirdly, the enrichment of gene signatures of exhausted mouse $CD8^+$ T cells in $CD39^+$ $CD8^+$ T cells in subjects with HCV infection further reinforces the relationship between CD39 expression and T cell exhaustion. Finally, analysis of global expression profiles of $CD39^+$ *versus* $CD39^-$ $CD8^+$ T cells in HCV-infected subjects showed that the $CD39^+$ cells were enriched for genes related to proliferation and cell cycle pathways. Initially this appeared to be the opposite of what we would expect given the reduced proliferative rates of exhausted $CD8^+$ T cells in vitro (152, 153, 195). However, LCMV mouse models of chronic infection demonstrate that exhausted $CD8^+$ T cells are dependent on continuous exposure to viral antigen to ensure their survival and as a result undergo extensive cell division at a rate higher than seen in physiological homeostatic proliferation of the memory $CD8^+$ T cell pool (187). Exhausted $CD8^+$ T cells therefore have a paradoxical increase in their proliferation *in vivo* despite reduced proliferative potential *in vitro* (195), explaining the increased expression of proliferation-associated genes in $CD39^+$ $CD8^+$ T cells in HCV infection and in mouse exhausted $CD8^+$ T cells (73, 196).

7.5 Terminally exhausted $CD8^+$ T cells are demarcated by CD39

A recent study published by Paley et al. has identified two distinct differentiation states of virus-specific $CD8^+$ T cells in chronic LCMV infected mice

and HCV infected humans (196). The T-box transcription factors T-bet and Eomes characterise these states. T-bet^{High} cells display low intrinsic turnover but are capable of proliferation in response to persisting antigen and give rise to Eomes^{High} terminal effector cells, which have a reduced capacity to undergo proliferation in vivo. Unpublished data (data not shown) from the Haining and Sharpe laboratories (Boston, MA, USA) show that in the LCMV mouse model of chronic infection the CD39^{High} subset of CD8^{+} T cells demarcates terminally exhausted Eomes^{High}/Tbet^{Low} cells as well as expressing higher levels of multiple inhibitory receptors including PD-1, 2B4 and LAG-3. These findings suggest that CD39 may be a marker not only of severe exhaustion, but more specifically of the most terminally exhausted cells (Figure 7.1). The ability to use more specific markers to identify terminally exhausted CD8^{+} T cells in chronic viral infection could aid in understanding how this progeny arises and possibly prevented.

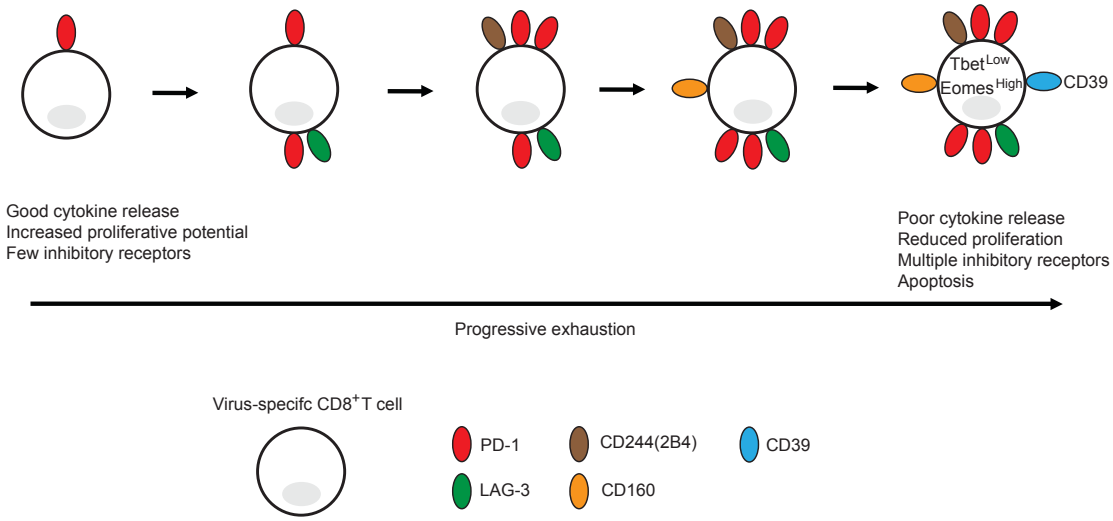


Figure 7.1. CD39 is a novel marker of severe exhaustion in the most terminally exhausted CD8^{+} T cells, which are phenotypically Eomes^{High}Tbet^{Low}.

7.6 Conclusion

The expression of CD39 has been increasingly recognised in multiple cellular contexts. Its enzymatic action of ATP metabolism and adenosine production via the activity of CD73 makes the purinergic system a likely method of providing an autocrine or juxtacrine (Figure 7.2) inhibitory pathway in inflamed tissues with high viral antigen loads in chronic infection. If this is so then inhibition of CD39 may work in conjunction with other strategies in restoring the function of exhausted $CD8^+$ T cells in chronic viral infections.

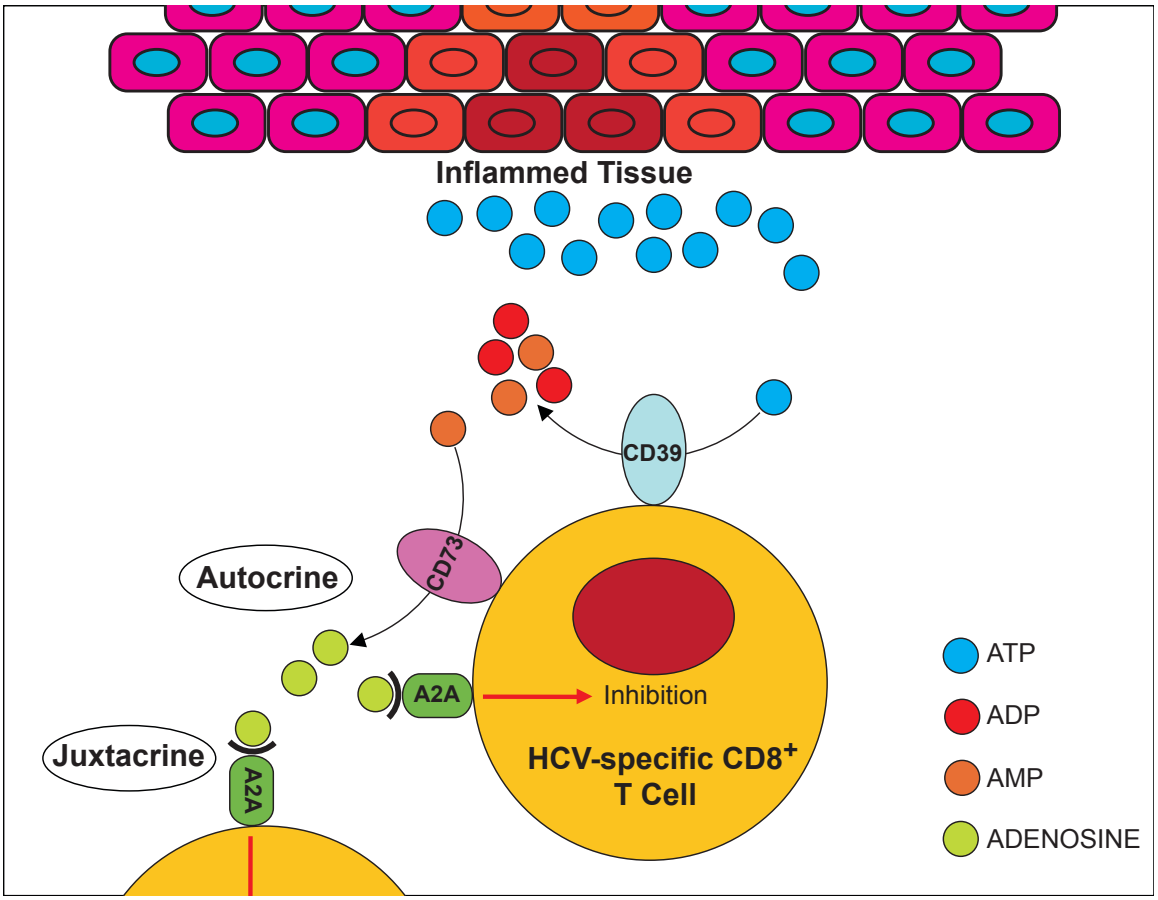


Figure 7.2. Proposed mechanism of T cell inhibition via purinergic pathway in chronic viral infection.

7.7 Future areas of work

There are a number of key areas of future work. Firstly, the identification of novel subclasses of HCV-specific $CD8^+$ T cell response (Figure 3.1) would need to be validated through comparison with other sample sets, or further experiments with larger cohorts. In addition, single cell analysis with the same samples used for gene expression profiling would be more comparable when looking for novel immunological subsets. Secondly, functional in vitro experiments with PBMCs from HCV infected donors could be performed. Samples with and without addition of the CD39 inhibitor POM-1 can be tested by flow cytomtery analysis to evaluate changes in markers of activation, exhaustion and cytokine profiles. Carboxyfluorescein succinimidyl ester (CFSE) assays can be used to determine changes in proliferation in virus-specific cells. Thirdly, functional genomic experiments could be performed to identify downstream regulators of T cell inhibition from the Adenosine A2A receptor. This could be done by engaging the adenosine A2A receptor on $CD8^+$ T cells using adenosine itself or a recognised agonist. Cells could then be processed for gene expression profiling and compared with unperturbed samples. Fourthly, additional experiments on CD39 null mice would be helpful to identify the functional purpose of CD39 in chronic viral infection. Histological examination of tissues after CD39 null mice that die from chronic clone 13 LCMV infection can help identify whether death occurs from an uncontrolled immune-pathological response or not. In addition functional assessment of $CD8^+$ T cells in terms of proliferative capacity and cytokine secretion could give insight into CD39's role

during chronic infection. Also, correlation of viral load and liver function in the mouse models would be helpful in consolidating human experiments where CD39 is thought to increase in response to a pro-inflammatory environment with high viral titers.

Another area of research that is highly applicable clinically is CD39's involvement with tumour growth. There is growing evidence that the malignant cells interact with the immune system in creating an inhibitory microenvironment through adenosine production (197), which therefore suggests a role for CD39 in tumour progression. Being able to counteract this by CD39 blockade could be of immense value in chemotherapeutics.

Chapter 8

References

8. References

1. Lauer GM & Walker BD (2001) Hepatitis C virus infection. *The New England journal of medicine* 345(1):41-52.
2. Feinstone SM, Kapikian AZ, Purcell RH, Alter HJ, & Holland PV (1975) Transfusion-associated hepatitis not due to viral hepatitis type A or B. *The New England journal of medicine* 292(15):767-770.
3. Choo QL, *et al.* (1989) Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* 244(4902):359-362.
4. Cooper S, *et al.* (1999) Analysis of a successful immune response against hepatitis C virus. *Immunity* 10(4):439-449.
5. Hatzakis A, *et al.* (2011) The state of hepatitis B and C in Europe: report from the hepatitis B and C summit conference*. *Journal of viral hepatitis* 18 Suppl 1:1-16.
6. Tai AW, *et al.* (2009) A functional genomic screen identifies cellular cofactors of hepatitis C virus replication. *Cell host & microbe* 5(3):298-307.
7. Lohmann V (2009) HCV replicons: overview and basic protocols. *Methods Mol Biol* 510:145-163.
8. Wakita T, *et al.* (2005) Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. *Nature medicine* 11(7):791-796.
9. Pybus OG, *et al.* (2009) Genetic history of hepatitis C virus in East Asia. *Journal of virology* 83(2):1071-1082.
10. Pybus OG, *et al.* (2001) The epidemic behavior of the hepatitis C virus. *Science* 292(5525):2323-2325.
11. Simmonds P (2004) Genetic diversity and evolution of hepatitis C virus--15 years on. *The Journal of general virology* 85(Pt 11):3173-3188.
12. Kwo PY, *et al.* (2010) Efficacy of boceprevir, an NS3 protease inhibitor, in combination with peginterferon alfa-2b and ribavirin in treatment-naive patients with genotype 1 hepatitis C infection (SPRINT-1): an open-label, randomised, multicentre phase 2 trial. *Lancet* 376(9742):705-716.
13. McHutchison JG, *et al.* (2009) Telaprevir with peginterferon and ribavirin for chronic HCV genotype 1 infection. *The New England journal of medicine* 360(18):1827-1838.
14. Jacobson IM, *et al.* (2014) Simeprevir with pegylated interferon alfa 2a plus ribavirin in treatment-naive patients with chronic hepatitis C virus genotype 1 infection (QUEST-1): a phase 3, randomised, double-blind, placebo-controlled trial. *Lancet* 384(9941):403-413.
15. Sulkowski MS, *et al.* (2013) Faldaprevir combined with pegylated interferon alfa-2a and ribavirin in treatment-naive patients with chronic genotype 1 HCV: SILEN-C1 trial. *Hepatology* 57(6):2143-2154.
16. Gentile I, *et al.* (2014) Daclatasvir: the first of a new class of drugs targeted against hepatitis C virus NS5A. *Current medicinal chemistry* 21(12):1391-1404.

17. Kowdley KV, *et al.* (2013) Sofosbuvir with pegylated interferon alfa-2a and ribavirin for treatment-naive patients with hepatitis C genotype-1 infection (ATOMIC): an open-label, randomised, multicentre phase 2 trial. *Lancet* 381(9883):2100-2107.
18. Barnes E, *et al.* (2012) Novel adenovirus-based vaccines induce broad and sustained T cell responses to HCV in man. *Science translational medicine* 4(115):115ra111.
19. Blouin A, Bolender RP, & Weibel ER (1977) Distribution of organelles and membranes between hepatocytes and nonhepatocytes in the rat liver parenchyma. A stereological study. *The Journal of cell biology* 72(2):441-455.
20. Doherty DG, *et al.* (1999) The human liver contains multiple populations of NK cells, T cells, and CD3+CD56+ natural T cells with distinct cytotoxic activities and Th1, Th2, and Th0 cytokine secretion patterns. *Journal of immunology* 163(4):2314-2321.
21. Crispe IN (2011) Liver antigen-presenting cells. *Journal of hepatology* 54(2):357-365.
22. Qian S, *et al.* (1994) Murine liver allograft transplantation: tolerance and donor cell chimerism. *Hepatology* 19(4):916-924.
23. Li W, *et al.* (2008) New insights into mechanisms of spontaneous liver transplant tolerance: the role of Foxp3-expressing CD25+CD4+ regulatory T cells. *American journal of transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons* 8(8):1639-1651.
24. Grakoui A, *et al.* (2003) HCV persistence and immune evasion in the absence of memory T cell help. *Science* 302(5645):659-662.
25. Shoukry NH, *et al.* (2003) Memory CD8+ T cells are required for protection from persistent hepatitis C virus infection. *The Journal of experimental medicine* 197(12):1645-1655.
26. Abdel-Hakeem MS, Bedard N, Murphy D, Bruneau J, & Shoukry NH (2014) Signatures of protective memory immune responses during hepatitis C virus reinfection. *Gastroenterology* 147(4):870-881 e878.
27. Kim AY, *et al.* (2011) Spontaneous control of HCV is associated with expression of HLA-B 57 and preservation of targeted epitopes. *Gastroenterology* 140(2):686-696 e681.
28. McKiernan SM, *et al.* (2004) Distinct MHC class I and II alleles are associated with hepatitis C viral clearance, originating from a single source. *Hepatology* 40(1):108-114.
29. Thimme R, *et al.* (2001) Determinants of viral clearance and persistence during acute hepatitis C virus infection. *The Journal of experimental medicine* 194(10):1395-1406.
30. Ge D, *et al.* (2009) Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* 461(7262):399-401.
31. Tillmann HL, *et al.* (2010) A polymorphism near IL28B is associated with spontaneous clearance of acute hepatitis C virus and jaundice. *Gastroenterology* 139(5):1586-1592, 1592 e1581.

32. Tanaka Y, *et al.* (2009) Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. *Nature genetics* 41(10):1105-1109.
33. Thompson AJ, *et al.* (2010) Interleukin-28B polymorphism improves viral kinetics and is the strongest pretreatment predictor of sustained virologic response in genotype 1 hepatitis C virus. *Gastroenterology* 139(1):120-129 e118.
34. Lauer GM (2013) Immune responses to hepatitis C virus (HCV) infection and the prospects for an effective HCV vaccine or immunotherapies. *The Journal of infectious diseases* 207 Suppl 1:S7-S12.
35. Thimme R, *et al.* (2002) Viral and immunological determinants of hepatitis C virus clearance, persistence, and disease. *Proceedings of the National Academy of Sciences of the United States of America* 99(24):15661-15668.
36. Cox AL, *et al.* (2005) Prospective evaluation of community-acquired acute-phase hepatitis C virus infection. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 40(7):951-958.
37. McGovern BH, *et al.* (2009) Improving the diagnosis of acute hepatitis C virus infection with expanded viral load criteria. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 49(7):1051-1060.
38. Cox AL, *et al.* (2005) Comprehensive analyses of CD8+ T cell responses during longitudinal study of acute human hepatitis C. *Hepatology* 42(1):104-112.
39. Wong DK, *et al.* (1998) Liver-derived CTL in hepatitis C virus infection: breadth and specificity of responses in a cohort of persons with chronic infection. *Journal of immunology* 160(3):1479-1488.
40. Germain RN, Miller MJ, Dustin ML, & Nussenzweig MC (2006) Dynamic imaging of the immune system: progress, pitfalls and promise. *Nature reviews. Immunology* 6(7):497-507.
41. Kaech SM, Wherry EJ, & Ahmed R (2002) Effector and memory T-cell differentiation: implications for vaccine development. *Nature reviews. Immunology* 2(4):251-262.
42. Williams MA & Bevan MJ (2007) Effector and memory CTL differentiation. *Annual review of immunology* 25:171-192.
43. Kaech SM & Wherry EJ (2007) Heterogeneity and cell-fate decisions in effector and memory CD8+ T cell differentiation during viral infection. *Immunity* 27(3):393-405.
44. Harty JT & Badovinac VP (2008) Shaping and reshaping CD8+ T-cell memory. *Nature reviews. Immunology* 8(2):107-119.
45. Joshi NS & Kaech SM (2008) Effector CD8 T cell development: a balancing act between memory cell potential and terminal differentiation. *Journal of immunology* 180(3):1309-1315.
46. Rutishauser RL, *et al.* (2009) Transcriptional repressor Blimp-1 promotes CD8(+) T cell terminal differentiation and represses the acquisition of central memory T cell properties. *Immunity* 31(2):296-308.

47. Stemberger C*, et al.* (2007) A single naive CD8+ T cell precursor can develop into diverse effector and memory subsets. *Immunity* 27(6):985-997.
48. Rubinstein MP*, et al.* (2008) IL-7 and IL-15 differentially regulate CD8+ T-cell subsets during contraction of the immune response. *Blood* 112(9):3704-3712.
49. Cui W & Kaech SM (2010) Generation of effector CD8+ T cells and their conversion to memory T cells. *Immunological reviews* 236:151-166.
50. Akbar AN, Terry L, Timms A, Beverley PC, & Janossy G (1988) Loss of CD45R and gain of UCHL1 reactivity is a feature of primed T cells. *Journal of immunology* 140(7):2171-2178.
51. Clement LT, Yamashita N, & Martin AM (1988) The functionally distinct subpopulations of human CD4+ helper/inducer T lymphocytes defined by anti-CD45R antibodies derive sequentially from a differentiation pathway that is regulated by activation-dependent post-thymic differentiation. *Journal of immunology* 141(5):1464-1470.
52. Masopust D, Vezys V, Marzo AL, & Lefrancois L (2014) Pillars article: preferential localization of effector memory cells in nonlymphoid tissue. Science. 2001. 291: 2413-2417. *Journal of immunology* 192(3):845-849.
53. Woodland DL & Kohlmeier JE (2009) Migration, maintenance and recall of memory T cells in peripheral tissues. *Nature reviews. Immunology* 9(3):153-161.
54. Barber DL, Wherry EJ, & Ahmed R (2003) Cutting edge: rapid in vivo killing by memory CD8 T cells. *Journal of immunology* 171(1):27-31.
55. Wherry EJ & Ahmed R (2004) Memory CD8 T-cell differentiation during viral infection. *Journal of virology* 78(11):5535-5545.
56. Marzo AL, Yagita H, & Lefrancois L (2007) Cutting edge: migration to nonlymphoid tissues results in functional conversion of central to effector memory CD8 T cells. *Journal of immunology* 179(1):36-40.
57. Hikono H*, et al.* (2007) Activation phenotype, rather than central- or effector-memory phenotype, predicts the recall efficacy of memory CD8+ T cells. *The Journal of experimental medicine* 204(7):1625-1636.
58. Darrah PA*, et al.* (2007) Multifunctional TH1 cells define a correlate of vaccine-mediated protection against Leishmania major. *Nature medicine* 13(7):843-850.
59. Altman JD*, et al.* (2011) Phenotypic analysis of antigen-specific T lymphocytes. Science. 1996. 274: 94-96. *Journal of immunology* 187(1):7-9.
60. Davis MM, Altman JD, & Newell EW (2011) Interrogating the repertoire: broadening the scope of peptide-MHC multimer analysis. *Nature reviews. Immunology* 11(8):551-558.
61. Yamamoto T*, et al.* (2011) Surface expression patterns of negative regulatory molecules identify determinants of virus-specific CD8+ T-cell exhaustion in HIV infection. *Blood* 117(18):4805-4815.
62. Lee PP*, et al.* (1999) Characterization of circulating T cells specific for tumor-associated antigens in melanoma patients. *Nature medicine* 5(6):677-685.
63. Gallimore A*, et al.* (1998) Induction and exhaustion of lymphocytic choriomeningitis virus-specific cytotoxic T lymphocytes visualized using

soluble tetrameric major histocompatibility complex class I-peptide complexes. *The Journal of experimental medicine* 187(9):1383-1393.
64. Zajac AJ, *et al.* (1998) Viral immune evasion due to persistence of activated T cells without effector function. *The Journal of experimental medicine* 188(12):2205-2213.
65. Virgin HW, Wherry EJ, & Ahmed R (2009) Redefining chronic viral infection. *Cell* 138(1):30-50.
66. Wherry EJ, Blattman JN, Murali-Krishna K, van der Most R, & Ahmed R (2003) Viral persistence alters CD8 T-cell immunodominance and tissue distribution and results in distinct stages of functional impairment. *Journal of virology* 77(8):4911-4927.
67. Moskophidis D, Lechner F, Pircher H, & Zinkernagel RM (1993) Virus persistence in acutely infected immunocompetent mice by exhaustion of antiviral cytotoxic effector T cells. *Nature* 362(6422):758-761.
68. Zhang JY, *et al.* (2007) PD-1 up-regulation is correlated with HIV-specific memory CD8+ T-cell exhaustion in typical progressors but not in long-term nonprogressors. *Blood* 109(11):4671-4678.
69. Matloubian M, Concepcion RJ, & Ahmed R (1994) CD4+ T cells are required to sustain CD8+ cytotoxic T-cell responses during chronic viral infection. *Journal of virology* 68(12):8056-8063.
70. Blackburn SD, *et al.* (2009) Coregulation of CD8+ T cell exhaustion by multiple inhibitory receptors during chronic viral infection. *Nature immunology* 10(1):29-37.
71. Wherry EJ (2011) T cell exhaustion. in *Nature Publishing Group* (Nature Publishing Group), pp 492-499.
72. Lechner F, *et al.* (2000) Analysis of successful immune responses in persons infected with hepatitis C virus. *The Journal of experimental medicine* 191(9):1499-1512.
73. Shin H, Blackburn SD, Blattman JN, & Wherry EJ (2007) Viral antigen and extensive division maintain virus-specific CD8 T cells during chronic infection. *The Journal of experimental medicine* 204(4):941-949.
74. Shin H & Wherry EJ (2007) CD8 T cell dysfunction during chronic viral infection. *Current opinion in immunology* 19(4):408-415.
75. Bengsch B, *et al.* (2007) Analysis of CD127 and KLRG1 expression on hepatitis C virus-specific CD8+ T cells reveals the existence of different memory T-cell subsets in the peripheral blood and liver. *Journal of virology* 81(2):945-953.
76. Paiardini M, *et al.* (2005) Loss of CD127 expression defines an expansion of effector CD8+ T cells in HIV-infected individuals. *Journal of immunology* 174(5):2900-2909.
77. Bengsch B, *et al.* (2010) Coexpression of PD-1, 2B4, CD160 and KLRG1 on exhausted HCV-specific CD8+ T cells is linked to antigen recognition and T cell differentiation. *PLoS pathogens* 6(6):e1000947.
78. Nitschke K, *et al.* (2015) Tetramer Enrichment Reveals the Presence of Phenotypically Diverse Hepatitis C Virus-Specific CD8+ T Cells in Chronic Infection. *Journal of virology* 89(1):25-34.

79. von Hahn T, *et al.* (2007) Hepatitis C virus continuously escapes from neutralizing antibody and T-cell responses during chronic infection in vivo. *Gastroenterology* 132(2):667-678.
80. Wedemeyer H, *et al.* (2002) Impaired effector function of hepatitis C virus-specific CD8+ T cells in chronic hepatitis C virus infection. *Journal of immunology* 169(6):3447-3458.
81. McMahan RH, *et al.* (2010) Tim-3 expression on PD-1+ HCV-specific human CTLs is associated with viral persistence, and its blockade restores hepatocyte-directed in vitro cytotoxicity. *The Journal of clinical investigation* 120(12):4546-4557.
82. Nakamoto N, *et al.* (2008) Functional restoration of HCV-specific CD8 T cells by PD-1 blockade is defined by PD-1 expression and compartmentalization. *Gastroenterology* 134(7):1927-1937, 1937 e1921-1922.
83. Schlaphoff V, *et al.* (2011) Dual function of the NK cell receptor 2B4 (CD244) in the regulation of HCV-specific CD8+ T cells. *PLoS pathogens* 7(5):e1002045.
84. Sangro B, *et al.* (2013) A clinical trial of CTLA-4 blockade with tremelimumab in patients with hepatocellular carcinoma and chronic hepatitis C. *Journal of hepatology* 59(1):81-88.
85. Klenerman P & Thimme R (2012) T cell responses in hepatitis C: the good, the bad and the unconventional. *Gut* 61(8):1226-1234.
86. Fuller MJ, *et al.* (2013) Immunotherapy of chronic hepatitis C virus infection with antibodies against programmed cell death-1 (PD-1). *Proceedings of the National Academy of Sciences of the United States of America* 110(37):15001-15006.
87. Nebbia G, *et al.* (2012) Upregulation of the Tim-3/galectin-9 pathway of T cell exhaustion in chronic hepatitis B virus infection. *PloS one* 7(10):e47648.
88. Freeman GJ, Wherry EJ, Ahmed R, & Sharpe AH (2006) Reinvigorating exhausted HIV-specific T cells via PD-1-PD-1 ligand blockade. *The Journal of experimental medicine* 203(10):2223-2227.
89. Petrovas C, *et al.* (2006) PD-1 is a regulator of virus-specific CD8+ T cell survival in HIV infection. *The Journal of experimental medicine* 203(10):2281-2292.
90. Velu V, *et al.* (2009) Enhancing SIV-specific immunity in vivo by PD-1 blockade. *Nature* 458(7235):206-210.
91. Brahmer JR, *et al.* (2010) Phase I study of single-agent anti-programmed death-1 (MDX-1106) in refractory solid tumors: safety, clinical activity, pharmacodynamics, and immunologic correlates. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 28(19):3167-3175.
92. Blackburn SD, *et al.* (2010) Tissue-specific differences in PD-1 and PD-L1 expression during chronic viral infection: implications for CD8 T-cell exhaustion. *Journal of virology* 84(4):2078-2089.
93. Workman CJ, *et al.* (2004) Lymphocyte activation gene-3 (CD223) regulates the size of the expanding T cell population following antigen activation in vivo. *Journal of immunology* 172(9):5450-5455.

94. Pentcheva-Hoang T, Egen JG, Wojnoonski K, & Allison JP (2004) B7-1 and B7-2 selectively recruit CTLA-4 and CD28 to the immunological synapse. *Immunity* 21(3):401-413.
95. Okazaki T, Maeda A, Nishimura H, Kurosaki T, & Honjo T (2001) PD-1 immunoreceptor inhibits B cell receptor-mediated signaling by recruiting src homology 2-domain-containing tyrosine phosphatase 2 to phosphotyrosine. *Proceedings of the National Academy of Sciences of the United States of America* 98(24):13866-13871.
96. Quigley M*, et al.* (2010) Transcriptional analysis of HIV-specific CD8+ T cells shows that PD-1 inhibits T cell function by upregulating BATF. in *Nature medicine* (Nature Publishing Group), pp 1147-1151.
97. Hasko G, Linden J, Cronstein B, & Pacher P (2008) Adenosine receptors: therapeutic aspects for inflammatory and immune diseases. *Nature reviews. Drug discovery* 7(9):759-770.
98. Yegutkin GG (2008) Nucleotide- and nucleoside-converting ectoenzymes: Important modulators of purinergic signalling cascade. *Biochimica et biophysica acta* 1783(5):673-694.
99. Vignali DA, Collison LW, & Workman CJ (2008) How regulatory T cells work. *Nature reviews. Immunology* 8(7):523-532.
100. Mandapathil M, Lang S, Gorelik E, & Whiteside TL (2009) Isolation of functional human regulatory T cells (Treg) from the peripheral blood based on the CD39 expression. *Journal of immunological methods* 346(1-2):55-63.
101. Deaglio S*, et al.* (2007) Adenosine generation catalyzed by CD39 and CD73 expressed on regulatory T cells mediates immune suppression. *The Journal of experimental medicine* 204(6):1257-1265.
102. Boer MC, van Meijgaarden KE, Bastid J, Ottenhoff TH, & Joosten SA (2013) CD39 is involved in mediating suppression by Mycobacterium bovis BCG-activated human CD8(+) CD39(+) regulatory T cells. *European journal of immunology* 43(7):1925-1932.
103. Alizadeh A, Eisen M, Botstein D, Brown PO, & Staudt LM (1998) Probing lymphocyte biology by genomic-scale gene expression analysis. *Journal of clinical immunology* 18(6):373-379.
104. Schena M, Shalon D, Davis RW, & Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270(5235):467-470.
105. Kaech SM, Hemby S, Kersh E, & Ahmed R (2002) Molecular and functional profiling of memory CD8 T cell differentiation. *Cell* 111(6):837-851.
106. Kaech SM*, et al.* (2003) Selective expression of the interleukin 7 receptor identifies effector CD8 T cells that give rise to long-lived memory cells. *Nature immunology* 4(12):1191-1198.
107. Barber DL*, et al.* (2006) Restoring function in exhausted CD8 T cells during chronic viral infection. *Nature* 439(7077):682-687.
108. Kallies A, Xin A, Belz GT, & Nutt SL (2009) Blimp-1 transcription factor is required for the differentiation of effector CD8(+) T cells and memory responses. *Immunity* 31(2):283-295.

109. Quigley M, *et al.* (2010) Transcriptional analysis of HIV-specific CD8+ T cells shows that PD-1 inhibits T cell function by upregulating BATF. *Nature medicine* 16(10):1147-1151.
110. Lopes AR, *et al.* (2008) Bim-mediated deletion of antigen-specific CD8 T cells in patients unable to control HBV infection. *The Journal of clinical investigation* 118(5):1835-1845.
111. Day CL, *et al.* (2006) PD-1 expression on HIV-specific T cells is associated with T-cell exhaustion and disease progression. *Nature* 443(7109):350-354.
112. Huang da W, Sherman BT, & Lempicki RA (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic acids research* 37(1):1-13.
113. Ashburner M, *et al.* (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics* 25(1):25-29.
114. Ogata H, *et al.* (1999) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic acids research* 27(1):29-34.
115. Matys V, *et al.* (2003) TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic acids research* 31(1):374-378.
116. Liberzon A (2014) A description of the Molecular Signatures Database (MSigDB) Web site. *Methods Mol Biol* 1150:153-160.
117. Novershtern N, *et al.* (2011) Densely interconnected transcriptional circuits control cell states in human hematopoiesis. *Cell* 144(2):296-309.
118. Golub TR, *et al.* (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286(5439):531-537.
119. Guo G, *et al.* (2010) Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. *Developmental cell* 18(4):675-685.
120. Flatz L, *et al.* (2011) Single-cell gene-expression profiling reveals qualitatively distinct CD8 T cells elicited by different gene-based vaccines. *Proceedings of the National Academy of Sciences of the United States of America* 108(14):5724-5729.
121. Leisner C, *et al.* (2008) One-pot, mix-and-read peptide-MHC tetramers. *PloS one* 3(2):e1678.
122. Murali-Krishna K, *et al.* (1998) Counting antigen-specific CD8 T cells: a reevaluation of bystander activation during viral infection. *Immunity* 8(2):177-187.
123. Lazarowski ER, *et al.* (2004) Nucleotide release provides a mechanism for airway surface liquid homeostasis. *The Journal of biological chemistry* 279(35):36855-36864.
124. Chen Y, *et al.* (2006) ATP release guides neutrophil chemotaxis via P2Y2 and A3 receptors. *Science* 314(5806):1792-1795.
125. Haining WN, *et al.* (2008) Identification of an evolutionarily conserved transcriptional signature of CD8 memory differentiation that is shared by T and B cells. *Journal of immunology* 181(3):1859-1868.
126. Doering TA, *et al.* (2012) Network analysis reveals centrally connected genes and pathways involved in CD8+ T cell exhaustion versus memory. *Immunity* 37(6):1130-1144.

127. Subramanian A, *et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* 102(43):15545-15550.
128. Ahmed R, Salmi A, Butler LD, Chiller JM, & Oldstone MB (1984) Selection of genetic variants of lymphocytic choriomeningitis virus in spleens of persistently infected mice. Role in suppression of cytotoxic T lymphocyte response and viral persistence. *The Journal of experimental medicine* 160(2):521-540.
129. Parish IA & Kaech SM (2009) Diversity in CD8(+) T cell differentiation. *Current opinion in immunology* 21(3):291-297.
130. Rappuoli R & Aderem A (2011) A 2020 vision for vaccines against HIV, tuberculosis and malaria. *Nature* 473(7348):463-469.
131. Armstrong SA, *et al.* (2002) MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nature genetics* 30(1):41-47.
132. O'Donnell EA, Ernst DN, & Hingorani R (2013) Multiparameter flow cytometry: advances in high resolution analysis. *Immune network* 13(2):43-54.
133. Chen G & Weng NP (2012) Analyzing the phenotypic and functional complexity of lymphocytes using CyTOF (cytometry by time-of-flight). *Cellular & molecular immunology* 9(4):322-323.
134. Newell EW, Sigal N, Bendall SC, Nolan GP, & Davis MM (2012) Cytometry by time-of-flight shows combinatorial cytokine expression and virus-specific cell niches within a continuum of CD8+ T cell phenotypes. *Immunity* 36(1):142-152.
135. Hertoghs KM, *et al.* (2010) Molecular profiling of cytomegalovirus-induced human CD8+ T cell differentiation. *The Journal of clinical investigation* 120(11):4077-4090.
136. Wang Y, Miller DJ, & Clarke R (2008) Approaches to working in high-dimensional data spaces: gene expression microarrays. *British journal of cancer* 98(6):1023-1028.
137. Sachs K, Perez O, Pe'er D, Lauffenburger DA, & Nolan GP (2005) Causal protein-signaling networks derived from multiparameter single-cell data. *Science* 308(5721):523-529.
138. Arsenio J, *et al.* (2014) Early specification of CD8+ T lymphocyte fates during adaptive immunity revealed by single-cell gene-expression analyses. *Nature immunology* 15(4):365-372.
139. Boon WC, *et al.* (2011) Acoustic microstreaming increases the efficiency of reverse transcription reactions comprising single-cell quantities of RNA. *BioTechniques* 50(2):116-119.
140. Sun X, *et al.* (2010) CD39/ENTPD1 expression by CD4+Foxp3+ regulatory T cells promotes hepatic metastatic tumor growth in mice. *Gastroenterology* 139(3):1030-1040.
141. Nikolova M, *et al.* (2011) CD39/Adenosine Pathway Is Involved in AIDS Progression. in *PLoS pathogens*, p e1002110.

142. Chattopadhyay PK, Hogerkorp CM, & Roederer M (2008) A chromatic explosion: the development and future of multiparameter flow cytometry. *Immunology* 125(4):441-449.
143. Nevins JR & Potti A (2007) Mining gene expression profiles: expression signatures as cancer phenotypes. *Nature reviews. Genetics* 8(8):601-609.
144. Hyatt G, *et al.* (2006) Gene expression microarrays: glimpses of the immunological genome. *Nature immunology* 7(7):686-691.
145. Querec TD, *et al.* (2009) Systems biology approach predicts immunogenicity of the yellow fever vaccine in humans. *Nature immunology* 10(1):116-125.
146. Ramilo O, *et al.* (2007) Gene expression patterns in blood leukocytes discriminate patients with acute infections. *Blood* 109(5):2066-2077.
147. Zarek PE, *et al.* (2008) A2A receptor signaling promotes peripheral tolerance by inducing T-cell anergy and the generation of adaptive regulatory T cells. *Blood* 111(1):251-259.
148. Sauer AV, *et al.* (2012) Alterations in the adenosine metabolism and CD39/CD73 adenosinergic machinery cause loss of Treg cell function and autoimmunity in ADA-deficient SCID. *Blood* 119(6):1428-1439.
149. Reutershan J, *et al.* (2009) Adenosine and inflammation: CD39 and CD73 are critical mediators in LPS-induced PMN trafficking into the lungs. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 23(2):473-482.
150. Dwyer KM, *et al.* (2007) CD39 and control of cellular immune responses. *Purinergic signalling* 3(1-2):171-180.
151. Kaech SM & Cui W (2012) Transcriptional control of effector and memory CD8+ T cell differentiation. *Nature reviews. Immunology* 12(11):749-761.
152. Wherry EJ (2011) T cell exhaustion. *Nature immunology* 12(6):492-499.
153. Wherry EJ, *et al.* (2007) Molecular signature of CD8+ T cell exhaustion during chronic viral infection. *Immunity* 27(4):670-684.
154. Kroy DC, *et al.* (2014) Liver environment and HCV replication affect human T-cell phenotype and expression of inhibitory receptors. *Gastroenterology* 146(2):550-561.
155. Kasprowicz V, *et al.* (2008) High level of PD-1 expression on hepatitis C virus (HCV)-specific CD8+ and CD4+ T cells during acute HCV infection, irrespective of clinical outcome. *Journal of virology* 82(6):3154-3160.
156. Raziorrouh B, *et al.* (2010) The immunoregulatory role of CD244 in chronic hepatitis B infection and its inhibitory potential on virus-specific CD8+ T-cell function. *Hepatology* 52(6):1934-1947.
157. Duraiswamy J, *et al.* (2011) Phenotype, function, and gene expression profiles of programmed death-1(hi) CD8 T cells in healthy human adults. *Journal of immunology* 186(7):4200-4212.
158. Kansas GS, Wood GS, & Tedder TF (1991) Expression, distribution, and biochemistry of human CD39. Role in activation-associated homotypic adhesion of lymphocytes. *Journal of immunology* 146(7):2235-2244.
159. Borsellino G, *et al.* (2007) Expression of ectonucleotidase CD39 by Foxp3+ Treg cells: hydrolysis of extracellular ATP and immune suppression. *Blood* 110(4):1225-1232.

160. Heine P, *et al.* (2001) The C-terminal cysteine-rich region dictates specific catalytic properties in chimeras of the ectonucleotidases NTPDase1 and NTPDase2. *European journal of biochemistry / FEBS* 268(2):364-373.
161. Lokshin A, *et al.* (2006) Adenosine-mediated inhibition of the cytotoxic activity and cytokine production by activated natural killer cells. *Cancer research* 66(15):7758-7765.
162. Moncrieffe H, *et al.* (2010) High Expression of the Ectonucleotidase CD39 on T Cells from the Inflamed Site Identifies Two Distinct Populations, One Regulatory and One Memory T Cell Population. in *The Journal of Immunology*, pp 134-143.
163. Schulze Zur Wiesch J, *et al.* (2011) Comprehensive analysis of frequency and phenotype of T regulatory cells in HIV infection: CD39 expression of FoxP3+ T regulatory cells correlates with progressive disease. *Journal of virology* 85(3):1287-1297.
164. Dwyer KM, *et al.* (2010) Expression of CD39 by human peripheral blood CD4+ CD25+ T cells denotes a regulatory memory phenotype. *American journal of transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons* 10(11):2410-2420.
165. Chen Y, *et al.* (2006) ATP Release Guides Neutrophil Chemotaxis via P2Y2 and A3 Receptors. in *Science*, pp 1792-1795.
166. Sumi Y, *et al.* (2014) Plasma ATP is Required for Neutrophil Activation in a Mouse Sepsis Model. *Shock* 42(2):142-147.
167. Antonioli L, Pacher P, Vizi ES, & Haskó G (2013) CD39 and CD73 in immunity and inflammation. in *Trends in Molecular Medicine* (Elsevier Ltd), pp 355-367.
168. Filippini A, Taffs RE, & Sitkovsky MV (1990) Extracellular ATP in T-lymphocyte activation: possible role in effector functions. *Proceedings of the National Academy of Sciences of the United States of America* 87(21):8267-8271.
169. Schenk U, *et al.* (2008) Purinergic control of T cell activation by ATP released through pannexin-1 hemichannels. *Science signaling* 1(39):ra6.
170. Yip L, *et al.* (2009) Autocrine regulation of T-cell activation by ATP release and P2X7 receptors. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 23(6):1685-1693.
171. Mizumoto N, *et al.* (2002) CD39 is the dominant Langerhans cell-associated ecto-NTPDase: modulatory roles in inflammation and immune responsiveness. *Nature medicine* 8(4):358-365.
172. Ohta A, *et al.* (2006) A2A adenosine receptor protects tumors from antitumor T cells. *Proceedings of the National Academy of Sciences of the United States of America* 103(35):13132-13137.
173. Rutebemberwa A, *et al.* (2008) High-programmed death-1 levels on hepatitis C virus-specific T cells during acute infection are associated with viral persistence and require preservation of cognate antigen during chronic infection. *Journal of immunology* 181(12):8215-8225.

174. Urbani S, *et al.* (2006) PD-1 expression in acute hepatitis C virus (HCV) infection is associated with HCV-specific CD8 exhaustion. *Journal of virology* 80(22):11398-11403.
175. Nikolova M, *et al.* (2011) CD39/adenosine pathway is involved in AIDS progression. *PLoS pathogens* 7(7):e1002110.
176. Nigam P, *et al.* (2010) Expansion of FOXP3+ CD8 T cells with suppressive potential in colorectal mucosa following a pathogenic simian immunodeficiency virus infection correlates with diminished antiviral T cell response and viral control. *Journal of immunology* 184(4):1690-1701.
177. Maliszewski CR, *et al.* (1994) The CD39 lymphoid cell activation antigen. Molecular cloning and structural characterization. *Journal of immunology* 153(8):3574-3583.
178. Junger WG (2011) Immune cell regulation by autocrine purinergic signalling. in *Nature Publishing Group* (Nature Publishing Group), pp 201-212.
179. Radziewicz H, *et al.* (2007) Liver-infiltrating lymphocytes in chronic human hepatitis C virus infection display an exhausted phenotype with high levels of PD-1 and low levels of CD127 expression. *Journal of virology* 81(6):2545-2553.
180. Trautmann L, *et al.* (2006) Upregulation of PD-1 expression on HIV-specific CD8+ T cells leads to reversible immune dysfunction. *Nature medicine* 12(10):1198-1202.
181. Merico D, Isserlin R, Stueker O, Emili A, & Bader GD (2010) Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PloS one* 5(11):e13984.
182. Pulte D, *et al.* (2011) CD39 expression on T lymphocytes correlates with severity of disease in patients with chronic lymphocytic leukemia. *Clinical lymphoma, myeloma & leukemia* 11(4):367-372.
183. Heng TS & Painter MW (2008) The Immunological Genome Project: networks of gene expression in immune cells. *Nature immunology* 9(10):1091-1094.
184. Isserlin R, Merico D, Voisin V, & Bader GD (2014) Enrichment Map - a Cytoscape app to visualize and explore OMICs pathway enrichment results. *F1000Research* 3:141.
185. Huttenhower C, *et al.* (2007) Nearest Neighbor Networks: clustering expression data based on gene neighborhoods. *BMC bioinformatics* 8:250.
186. Baitsch L, *et al.* (2011) Exhaustion of tumor-specific CD8(+) T cells in metastases from melanoma patients. *The Journal of clinical investigation* 121(6):2350-2360.
187. Shin H, *et al.* (2009) A role for the transcriptional repressor Blimp-1 in CD8(+) T cell exhaustion during chronic viral infection. *Immunity* 31(2):309-320.
188. Kukulski F, *et al.* (2011) NTPDase1 controls IL-8 production by human neutrophils. *Journal of immunology* 187(2):644-653.
189. Kohler D, *et al.* (2007) CD39/ectonucleoside triphosphate diphosphohydrolase 1 provides myocardial protection during cardiac ischemia/reperfusion injury. *Circulation* 116(16):1784-1794.

190. Fuller MJ & Zajac AJ (2003) Ablation of CD8 and CD4 T cell responses by high viral loads. *Journal of immunology* 170(1):477-486.
191. van der Most RG, *et al.* (2003) Changing immunodominance patterns in antiviral CD8 T-cell responses after loss of epitope presentation or chronic antigenic stimulation. *Virology* 315(1):93-102.
192. Barber DL, *et al.* (2005) Restoring function in exhausted CD8 T cells during chronic viral infection. in *Nature*, pp 682-687.
193. Boni C, *et al.* (2007) Characterization of hepatitis B virus (HBV)-specific T-cell dysfunction in chronic HBV infection. *Journal of virology* 81(8):4215-4225.
194. Levesque SA, Lavoie EG, Lecka J, Bigonnesse F, & Sevigny J (2007) Specificity of the ecto-ATPase inhibitor ARL 67156 on human and mouse ectonucleotidases. *British journal of pharmacology* 152(1):141-150.
195. Migueles SA, *et al.* (2002) HIV-specific CD8+ T cell proliferation is coupled to perforin expression and is maintained in nonprogressors. *Nature immunology* 3(11):1061-1068.
196. Paley MA, *et al.* (2012) Progenitor and terminal subsets of CD8+ T cells cooperate to contain chronic viral infection. *Science* 338(6111):1220-1225.
197. Bastid J, *et al.* (2013) ENTPD1/CD39 is a promising therapeutic target in oncology. *Oncogene* 32(14):1743-1751.

Chapter 9

Appendices

9. Appendices

Appendix Table 9.1. HCV clinical data used in chapter 4

Patient ID	Gender	Chronic/Resolver	Viral Load/iu/L	Virus Genotype	ALT/iu/L
00-23 P11 C63B	F	Chronic	<300	1a	18
00-23 P27 C63B	F	Chronic		1b	24
06-42 P3	F	Chronic	632972	1a	280
06-42 P5 143D	F	Chronic	<615	1a	
06K P3 143D	F	Resolver	<600	1a	124
06K P6 143D	F	Resolver	undetected	1a	
06L	M	Resolver		3	
06L P7 4H	M	Resolver	undetected	3	82
07-32 P2 4H	M	Chronic		2b	565
07-32 P5 4H	M	Chronic	89200	2b	263
07-39 P18	M	Chronic	1170	1a	33
07I P3	M	Chronic	1162	4a	34
07P P4	M	Chronic		1a	1379
08-024 P1 250A	F	Chronic	>700000	1	411
08-024 P14 A3Pool	F	Chronic	<43	1	11
08-024 P4 250A	F	Chronic	7540000	1	259
08-024 P5 A3Pool	F	Chronic	<600	1	35
08-024 P6 A3Pool	F	Chronic		1	11
08-027 P1 A2-198	M	Chronic	3838	no test	58
08-027 P5 A2-198	M	Chronic	1021	no test	75
08-23 P13	M	Resolver		1	
08-27 P2	M	Resolver		no test	
09-31 P3	M	Chronic	detected	1a	128
09-33 P3	F	Resolver		1a	20
09-33 P5 143D	F	Resolver	undetected	1a	15
09-37 P3 c63b	M	Chronic	<600	1a	32
09B P1 143D	M	Chronic	217000	1a	354
09B P5	M	Chronic	223000	1a	73
10-048 P2 143D	F	Chronic		1a	179
10-054 P1 143D	F	Chronic	1130	1a	209
10-078 P1 A2 226D	M	Chronic	89200	3a	875
10-19 P3	F	Resolver	<615		19
11-014 P1 143D	M	Resolver	3150	2a	129
11-017 P4 140G/259F	F	Chronic	25431	1a	481
12-043 P2 143D	M	Resolver	61602		692
12-103 P1 4H	F	Chronic	432	3a	44
12-181 P1 4H	F	Chronic		3a	
13-024 P1 140G	M	Chronic	147	1a	205
BR-3000 P12 A2 Mix2	M	Resolver	undetected	1a	24
BR-3000 P2 A2 Mix2	M	Resolver	47272	1a	36
BR-554 P13 C63B	F	Chronic	2038	1a	9
BR-554 P17 C63B	F	Chronic	6463017	1a	45
BR-554 P3 C63B	F	Chronic	64497	1a	39
BR1036 P13 C63B	F	Resolver	<1000	undetectable	28
BR1036 P9 C63B	F	Resolver	<1000	undetectable	4
BR1144 P10 C63B	F	Resolver	<1000	undetectable	13
BR1144 P5 C63B	F	Resolver	<1000	undetectable	2
BR554 P13 C63B	F	Chronic	2038	1a	9
BR554 P17 C63B	F	Chronic	6463017	1a	45
BR949 P5 C63B	F	Chronic	70047	1	36
CR54 P2 4H	F	Chronic	detected	1	237
CR54 P3 4H	F	Chronic	detected	1	103

Appendix Table 9. 2. HIV clinical data used in chapter 4

Patient ID	Gender	Progressor/Controller	ON/OFF Rx	Viral Load/cpm	CD4 Count
254567	M	Chronic	OFF	1823	606
350103	F	Chronic	OFF	431	625
350534	M	Chronic	OFF	24500	154
359260	M	Chronic	OFF	10322	541
384682	M	Chronic	ON	147	510
387879	M	Chronic	OFF	14600	677
403998	F	Chronic	OFF	2100	877
128019	M	Viraemic Controllers	OFF		
186089	M	Viraemic Controllers	OFF	82	740
237983	F	Viraemic Controllers	OFF	189	1232
270245	M	Viraemic Controllers	OFF	15	
302225	M	Viraemic Controllers	OFF	65	484
711950	M	Viraemic Controllers	OFF	300	700
732751	M	Viraemic Controllers	OFF	1860	1550
255675	M	Elite Controllers	OFF	103	963
269198	M	Elite Controllers	OFF		
285297	F	Elite Controllers	OFF	118	1246
321797	M	Elite Controllers	OFF		
831969	F	Elite Controllers	OFF		
R060	M	Chronic	OFF	117934	480
R086	M	Chronic	OFF	172886	410
R089	M	Chronic	OFF	44000	680
R046	M	Chronic	OFF	28445	910
R050	M	Chronic	OFF	20210	440
R041	M	Chronic	OFF	8435	320
R017	M	Chronic	OFF	172886	410
N034	M	Chronic	OFF	44000	680
R134	M	Chronic	OFF	500000	430
N012	M	Chronic	OFF	36695	
N090	F	Chronic	OFF	3362	490
N104	M	Chronic	OFF	4533	390
OX019	F	Chronic	OFF	42912	740
R051	M	Chronic	OFF	500000	560
R069	M	Chronic	OFF	63257	450
N004	M	Chronic	OFF	500000	430
N093	F	Chronic	OFF	2216	700
OX034	M	Chronic	OFF	124153	430
H005	M	Chronic	OFF	747	640
H033	M	Chronic	OFF	8036	430
R103	F	Chronic	OFF	8435	320

Appendix Table 9. 3. HCV clinical data used in chapter 3

Patient ID	Genotype	Age/years	Gender	Chro/Res	SC	WPI	Viral Load/iu/L	ALT	IL28B	Multimer	Peptide used	Antigen derived from
01-21 HCV	1a	28	F	Chronic	1	16	<600	311	CC	4H	GPRLGVRAT	HCV core 41-49
04-08 HCV	NA	29	F	Resolver	1	10	<600	13	CC	143D	ATDALMTGY	HCV NS3 1435-1443
05Y HCV	NA	21	F	Resolver	1	11	<600	24	TT	c63b	CINGVCWTV	HCV NS3 1073-1081
06K HCV	1a	22	F	Resolver	1	28	<600	159	CC	143D	ATDALMTGY	HCV NS3 1435-1443
06L HCV	3	28	M	Resolver	1	11	1960	463	CC	4H	GPRLGVRAT	HCV core 41-49
07-32 HCV	2b	27	M	Chronic	1	20	12833	51	CT	4H	GPRLGVRAT	HCV core 41-49
07-39 HCV	1	31	M	Chronic	1	12	80100	718	TT	c63b	CINGVCWTV	HCV NS3 1073-1081
07S HCV	NA	31	F	Resolver	1	24	29327	55	CC	127D	GIDPNIRTGV	HCV NS3 1273-1082
10974 HCV	1a	27	M	Resolver	1	9	<600	31	CC	c63b	CINGVCWTV	HCV NS3 1073-1081
50026 HCV	1a	26	F	Chronic	1	10	97400	267	CC	A2-61	YPYRLWHYPC	HCV E2 610-619
50068 HCV	1a	27	F	Chronic	1	17	642	42	CC	143D	ATDALMTGY	HCV NS3 1435-1443
CR1036 HCV	NA	24	F	Resolver	1	19	<1000	1756	CC	c63b	CINGVCWTV	HCV NS3 1073-1081
CR1144 HCV	NA	25	F	Resolver	1	10	<1000	173	CC	c63b	CINGVCWTV	HCV NS3 1073-1081
CR1816 HCV	NA	36	M	Resolver	1	8	<1000	291	CT	B27	ARMILMTHF	HCV core 470-478
CR554 HCV	1a	20	F	Chronic	1	20	2745	no data	TT	c63b	CINGVCWTV	HCV NS3 1073-1081
02-03 HCV	1a	21	F	Chronic	2	28	<600	No data	CT	c63b	CINGVCWTV	HCV NS3 1073-1081
03H HCV	1a	21	F	Chronic	2	28	3020	54	CC	c63b	CINGVCWTV	HCV NS3 1073-1081
04-11 HCV	1a	22	F	Chronic	2	20	<600	261	CC	c63b	CINGVCWTV	HCV NS3 1073-1081
06-42 HCV	1a	39	M	Chronic	2	12	242000	394	CC	143D	ATDALMTGY	HCV NS3 1435-1443
07-34 HCV	3a	24	F	Chronic	2	12	55684	269	TT	226D	REISVPAEIL	HCV NS5a 2266-2275
07-47 HCV	1a	32	M	Chronic	2	10	22635	no data	CT	140G	KLVALGINAV	HCV NS3 1406-1415
07-52 HCV	1a	36	F	Chronic	2	32	458000	no data	CT	143D	ATDALMTGY	HCV NS3 1435-1443
07M HCV	NA	19	F	Resolver	2	14	<615	12	CC	A2-198	VLSDFKTWL	HCV NS5a 1987-1995
07P HCV	1a	33	M	Chronic	2	16	2290	65	CT	c63b	CINGVCWTV	HCV NS3 1073-1081
07Z HCV	NA	21	F	Resolver	2	9	<615	no data	CC	143D	ATDALMTGY	HCV NS3 1435-1443
CR949 HCV	1a	37	F	Chronic	2	26	59421	18	CC	c63b	CINGVCWTV	HCV NS3 1073-1081
00-23 HCV	1b	14	F	Chronic	2	16	444000	94	CC	c63b	CINGVCWTV	HCV NS3 1073-1081

M = Male, F = Female
WPI = Weeks Post Infection when sample taken
NA=Not Available
Viral Load, ALT samples taken at same time as samples for multimer staining.
Chro/Res = Chronic, Resolver status
SC = Subclass 1 or 2

Appendix Table 9. 4. Complete list of multimers used.

Virus	Multimer type	HLA	Peptide used	Antigen derived from	Supplier
HCV	Pentamer	A*02:01	GIDPNIRTGV	HCV NS3 1273-1082	Proimmune
HCV	Pentamer	A*02:01	KLVALGINAV	HCV NS3 1406-1415	Proimmune
HCV	Pentamer	A*01:01	ATDALMTGY	HCV NS3 1435-1443	Proimmune
HCV	Pentamer	B*40:01	REISVPAEIL	HCV NS5a 2266-2275	Proimmune
HCV	Pentamer	B*07:02	GPRLGVRAT	HCV core 41-49	Proimmune
HCV	Pentamer	A*02:01	VLSDFKTWL	HCV NS5a 1987-1995	Proimmune
HCV	Pentamer	A*02:01	YPYRLWHYPC	HCV E2 610-619	Proimmune
HCV	Pentamer	B*27:01	ARMILMTHF	HCV core 470-478	Proimmune
HCV	Pentamer	A*02:01	CINGVCWTV	HCV NS3 1073-1081	Proimmune
CMV	Dextramer	A*02:01	NLVPMVATC	HCMV pp65	Immudex
EBV	Dextramer	A*02:01	GLCTLVAML	EBV BMLF-1	Immudex
HIV	Tetramer	A*24:02	RYPLTFGW	Nef RW8	Custom made
HIV	Tetramer	B*57:01	KAFSPEVIPMF	Gag KF11	Custom made
HIV	Tetramer	B*14:02	DRFYKTLRA	Gag DA9	Custom made
HIV	Tetramer	B*35:01	HPVHAGPIA	Gag HA9	Custom made
HIV	Tetramer	B*14:02	DRFYKTLRA	Gag DA9	Custom made
HIV	Dextramer	A*02:01	SLYNTVATL	Gag SL9	Immudex
HIV	Pentamer	B*07:02	TPQDLNTML	Gag TL9	Proimmune
HIV	Dextramer	A*02:01	SLYNTVATL	Gag SL9	Immudex
HIV	Dextramer	B*57:01	KAFSPEVIPMF	Gag KF11	Immudex
HIV	Tetramer	B*08:01	EIYKRWII	Gag EI8	Custom made
HIV	Tetramer	B*35:01	VPLRPMTY	Nef VY8	Beckman
HIV	Dextramer	B*07:02	GPGHKARVL	Gag GL9	Immudex
LCMV	Tetramer	H-2Db	SGVENPGGYCL	GP276-286	Dr. E. John Wherry
LCMV	Tetramer	H-2Db	KAVYNFATM	GP33-41	Dr. E. John Wherry
YFV-17D	Tetramer	B*3502	IPVIVADDL	E	Emory University

Appendix Table 9.5. RT-STA master mix for single cell sorts

Component	Vol to add per reaction/plate (ul)		Final conc.
	1x	120x	
2x CellsDirect reaction mix	5	600	1x
0.2x Primer/probe mix	2.5	300	0.2x
Superscript III/PlatinumTamix	0.2	24	
Superase-In	0.1	12	20U/ul
TE Buffer	2.3	276	
Total	10.1	1212	

Appendix Table 9.6. RT protocol

	Process	Temp/◦c	Time/mins
	RT	50	60
	Taq activation	95	2
20 to 24 cycles	Denature	95	15sec
	Anneal/extension	60	4mins
		Hold at 4 ◦c	

Appendix Table 9.7. Taqman probe assay mix

Component	Volume per inlet(ul)
20x Taqman assay	3.75
2x Assay loading reagent	3.75
Total Volume	7.5

Appendix Table 9.8. cDNA sample mix

Component	Volume per inlet(ul)
cDNA	3.38
Taqman Universal PCR Master Mix(2x) Applied bio – PN4304437	3.75
20x gene sample loading reagent	0.375
Total volume	7.505

Appendix Table 9.9. List of gene clusters from functional network analysis used in single cell analysis.

Control	Interesting	HCV Subclass 1				HCV Subclass 2	
B2M	ADA	ABHD3	ACAA2	ATF7IP	C9orf95	CD83	AUTS2
CD4	ADORA2A	ACYP1	ACSL5	C11orf73	CROT	CREM	BCOR
CD8A	BATF	APAF1	ARFIP1	CAT	EIF1AY	FOSB	CCR7
GAPDH	JUNB	CCDC91	ARH6AP15	COPZ1	ENTPD1	GADD4A	FOSL2
	NT5E	CCNG2	ARHGEF6	DCK	FAM8A1	GOS2	IL4R
		DCP2	CRYZ	DZIP3	FLI1	NR4A2	PGAP1
		HSD17B4	GBP1	LEF1	GMPR2	REL	PPP1R16B
		POLA1	GZMA	MAD2L1	MDFIC	SC5DL	RBM38
		PRDX3	LY75	MRPL48	NOLC1	SKIL	SAMSN1
		PTPN6	M6PR	PMS1	PAICS	SLC7A5	SLC39A8
		RAC2	NDUFB5	PRKACB	PKP4	SMAD7	TMEM57
		RRM2	PRIM1	RPS6KA1	STAT6	SOD2	TNFRSF4
		SLC35D2	PTGER2	SAC3D1	TRIM27	THBS1	ZNF222
		VCL	U2AF1	TRIB2		USP36	
		ZFP106	ZWINT	USP33		ZNF331	

Cluster A	Cluster B	Cluster C	Cluster D	Cluster E	Cluster F

Appendix Table 9.10. Samples used for cDNA microarray of $CD39^+$ versus $CD39^-$ $CD8^+$ T cells in HCV infection – Chapter 5

SampleID	PatientID	CD39	Chronic/Re	Cell Numbe	Conc (ng/u	Vol cDNA	Vol H2O	Total cDNA
02-58	02-58	+	Chronic	96177	322.37	15.51	9.49	5
0258	02-58	-	Chronic	165812	351.4	14.23	10.77	5
0327	03-27	-	Resolver	250000	308.72	16.2	8.8	5
0327	03-27	+	Resolver	65000	308.31	16.22	8.78	5
0329	03-29	-	Resolver	171000	275.8	18.13	6.87	5
0329	03-29	+	Resolver	12000	233.48	21.42	3.58	5
111	BR111	-	Chronic	250000	325.85	15.34	9.66	5
111	BR111	+	Chronic	12358	231.58	21.59	3.41	5
1524	BR1524	-	Resolver	250000	335.43	14.91	10.09	5
1927	BR1927	-	Chronic	139000	360.32	13.88	11.12	5
1927	BR1927	+	Chronic	433	312.73	15.99	9.01	5
3000	BR3000	-	Resolver	47246	308.69	16.2	8.8	5
3000 (Failed QC)	*BR3000*	+	*Resolver*	*4303*	*167.52*	*25*	*0*	*4.2*
515	BR515	-	Resolver	250000	353.85	14.13	10.87	5
515	BR515	+	Resolver	14312	214.37	23.32	1.68	5
941	BR941	+	Chronic	54969	333.32	15	10	5

Appendix Table 9.11. List of differentially expressed genes in CD39^{+} versus CD39^{-} transcription profiling

Gene	Signal to noise	Rank	p-value	FDR(BH)	Fold Change
RCAN3	1.7913	1	0.00031	0.07173	2.1042
CYP2A13	1.7671	2	0.00031	0.07173	1.186
ENTPD1	1.6166	3	0.00187	0.1043	5.7818
HSF1	1.5899	4	0.00031	0.07173	1.4454
INTS1	1.581	5	0.00031	0.07173	1.6224
POU3F2	1.5791	6	0.00062	0.08198	1.3309
GYG2	1.5733	7	0.00031	0.07173	1.2974
CTLA4	1.5189	8	0.00062	0.08198	3.8931
DNASE1L3	1.4971	9	0.00031	0.07173	1.204
WNT8B	1.4476	10	0.00031	0.07173	1.2508
ZNRF4	1.4331	11	0.00062	0.08198	1.278
GABRA2	1.4211	12	0.00031	0.07173	1.2344
SNED1	1.4009	13	0.00031	0.07173	2.1733
RHBG	1.3928	14	0.00031	0.07173	1.2811
GJB5	1.379	15	0.00031	0.07173	1.2133
GYPC	1.3787	16	0.00093	0.0867	1.5515
DNMT3L	1.3772	17	0.00062	0.08198	1.362
TLR8	1.3757	18	0.00155	0.1014	1.3698
CHN1	1.3749	19	0.00093	0.0867	3.6246
DSG3	1.3706	20	0.00062	0.08198	1.2011
NPM1	1.3622	21	0.00031	0.07173	1.1079
EPCAM	1.3614	22	0.00031	0.07173	1.1867
CGB	1.3582	23	0.00031	0.07173	1.2567
CX3CL1	1.3554	24	0.00031	0.07173	1.2618
AP2M1	1.3502	25	0.00093	0.0867	1.221
CCIN	1.3469	26	0.00062	0.08198	1.2621
BAI3	1.3393	27	0.00031	0.07173	1.3617
GH1	1.3316	28	0.00062	0.08198	1.1055
CD74	1.317	29	0.00124	0.09507	1.6103
PRODH2	1.3045	30	0.00062	0.08198	1.3053
TCEB3	1.2999	31	0.00093	0.0867	1.386
PITX3	1.2969	32	0.00093	0.0867	1.316
SH3GL1	1.2892	33	0.00031	0.07173	1.29
LMCD1	1.2828	34	0.00093	0.0867	1.6858
MYO9B	1.2785	35	0.00093	0.0867	1.4598
C7orf26	1.2783	36	0.00062	0.08198	1.2691
CC2D2B	1.277	37	0.00062	0.08198	2.0535
SPOCK3	1.2757	38	0.00093	0.0867	1.2352
FABP4	1.2746	39	0.00093	0.0867	1.1985
SLC6A14	1.2723	40	0.00031	0.07173	1.2065
TRIB1	1.2677	41	0.00062	0.08198	2.2411
AZGP1P1	1.2577	42	0.00093	0.0867	1.2371
GTPBP10	1.256	43	0.00093	0.0867	1.3162
CD79A	1.2543	44	0.00031	0.07173	1.5641
MRAS	1.2478	45	0.00155	0.1014	1.2022
FOLR2	1.2393	46	0.00062	0.08198	1.3092
POLR1E	1.238	47	0.00124	0.09507	1.6397
TIAM1	1.2308	48	0.00124	0.09507	2.339
SLITRK5	1.2308	49	0.00124	0.09507	1.2273
AGPAT3	1.2296	50	0.000932	0.0867	1.3958
DEAF1	1.2261	51	0.000932	0.0867	1.3357
HAPLN2	1.2256	52	0.000311	0.07173	1.3387
MYO7B	1.2227	53	0.000622	0.08198	1.3584
AGPAT2	1.2154	54	0.000622	0.08198	1.2025
SEC13	1.2138	55	0.001865	0.1043	1.3041
ADRA1D	1.2113	56	0.001554	0.1014	1.216
MLLT1	1.2092	57	0.000932	0.0867	1.5011
C19orf6	1.2051	58	0.000932	0.0867	1.4635
PRKAR1B	1.2049	59	0.000932	0.0867	1.5641
LPAR1	1.2036	60	0.000311	0.07173	1.2707
TRPV6	1.2019	61	0.001554	0.1014	1.1896
BUB1	1.2001	62	0.000311	0.07173	1.9324
PCSK2	1.1988	63	0.000622	0.08198	1.2676
SYNJ2	1.1955	64	0.000932	0.0867	1.6214
DLGAP2	1.1925	65	0.000311	0.07173	1.2154
FLJ13224	1.1922	66	0.001243	0.09507	1.6371
SBNO2	1.1859	67	0.000622	0.08198	1.2249
DSC2	1.1826	68	0.002176	0.1073	1.1357
HLX	1.1812	69	0.002176	0.1073	1.2686
TNNT1	1.1787	70	0.000622	0.08198	1.262
TNFRSF25	1.1711	71	0.001554	0.1014	2.4302
NOX1	1.1696	72	0.001865	0.1043	1.2805
RRP1B	1.164	73	0.000311	0.07173	1.3039
CNOT3	1.1635	74	0.002176	0.1073	1.479
ETV1	1.1624	75	0.001554	0.1014	1.4196
RFX2	1.1614	76	0.002797	0.1126	1.358
DLG1	1.1583	77	0.001865	0.1043	1.2407
AP3D1	1.1529	78	0.001554	0.1014	1.3716
F3	1.1498	79	0.000311	0.07173	1.3511

Gene	Signal to noise	Rank	p-value	FDR(BH)	Fold Change
WSB2	-1.5008	12922	0.00031	0.07173	-1.3756
MEAF6	-1.4912	12921	0.0001	0	-1.4765
SLCO4C1	-1.4875	12920	0.00031	0.07173	-2.6663
MST4	-1.4622	12919	0.0001	0	-1.6189
NPTN	-1.4245	12918	0.0001	0	-1.6221
SLC20A1	-1.4102	12917	0.00031	0.07173	-1.659
C1D	-1.4076	12916	0.00187	0.1043	-1.5637
TMEM59	-1.3999	12915	0.0001	0	-1.2436
CACNB3	-1.3554	12914	0.00062	0.08198	-1.6523
UBE2E3	-1.3387	12913	0.0001	0	-2.0872
C6orf62	-1.3338	12912	0.0001	0	-1.2748
QKI	-1.332	12911	0.00031	0.07173	-1.5651
KIAA0485	-1.3268	12910	0.00124	0.09507	-1.3636
SPINLW1	-1.3118	12909	0.00124	0.09507	-1.5986
ZNF137P	-1.3036	12908	0.00062	0.08198	-1.8262
VDAC3	-1.3	12907	0.00062	0.08198	-1.2718
BMI1	-1.2995	12906	0.00031	0.07173	-2.0904
ATP6AP2	-1.2974	12905	0.00093	0.0867	-1.3542
TGOLN2	-1.2854	12904	0.00031	0.07173	-1.3608
MYL12B	-1.2719	12903	0.0001	0	-1.3278
SLK	-1.2644	12902	0.00062	0.08198	-1.4204
ADRB2	-1.2637	12901	0.00093	0.0867	-3.1899
GBE1	-1.2602	12900	0.0001	0	-2.1332
ATF2	-1.2594	12899	0.00031	0.07173	-1.5058
NAA40	-1.2516	12898	0.00062	0.08198	-1.9045
GPR56	-1.2429	12897	0.00124	0.09507	-4.074
LOC1005067	-1.2372	12896	0.00062	0.08198	-1.1827
IGF2R	-1.2358	12895	0.00031	0.07173	-2.0166
CPEB3	-1.221	12894	0.00031	0.07173	-1.7952
AGL	-1.2146	12893	0.0001	0	-1.515
ARHGEF12	-1.21	12892	0.00062	0.08198	-2.1042
ARAP2	-1.199	12891	0.00062	0.08198	-1.3739
KIAA0494	-1.1918	12890	0.00093	0.0867	-1.3387
S1PR5	-1.1848	12889	0.00062	0.08198	-1.6397
GNPTAB	-1.1841	12888	0.00031	0.07173	-1.9914
XBP1	-1.1825	12887	0.00155	0.1014	-1.9028
FGR	-1.1813	12886	0.00155	0.1014	-3.8525
ELF4	-1.1764	12885	0.0001	0	-1.4442
RNF5	-1.1707	12884	0.00062	0.08198	-1.4264
MAP4K5	-1.1676	12883	0.00311	0.1141	-1.5815
KLRAP1	-1.1575	12882	0.0001	0	-2.9092
PSTPIP2	-1.1531	12881	0.00031	0.07173	-3.3844
DMTF1	-1.1502	12880	0.00031	0.07173	-1.2465
C14orf169	-1.1472	12879	0.00124	0.09507	-1.3764
SPRYD7	-1.1444	12878	0.00062	0.08198	-1.6958
YWHAZ	-1.1416	12877	0.00187	0.1043	-1.2776
GATAD2A	-1.1395	12876	0.00062	0.08198	-1.2893
YTHDF2	-1.1362	12875	0.00249	0.1089	-1.1723
EFHD2	-1.135	12874	0.00155	0.1014	-2.2398
ERBB2IP	-1.1207	12873	0.000933	0.0867	-1.4399
SSX2IP	-1.1155	12872	0.001865	0.1043	-1.7475
RAB6A	-1.1114	12871	0.000311	0.07173	-1.2912
PRSS23	-1.1073	12870	0.002486	0.1089	-4.6908
KLRD1	-1.103	12869	0.002486	0.1089	-3.2576
KDM3B	-1.1015	12868	0.002176	0.1073	-1.191
SLTM	-1.0985	12867	0.002486	0.1089	-1.2656
RAB9A	-1.0957	12866	0.001243	0.09507	-1.4902
PANK3	-1.0948	12865	0.000622	0.08198	-1.6069
ENPP4	-1.0943	12864	0.001865	0.1043	-2.1683
TTC38	-1.0917	12863	0.002486	0.1089	-2.0559
GALNT10	-1.0852	12862	0.005594	0.1372	-1.8054
ZEB2	-1.0848	12861	0.000622	0.08198	-2.4852
SLC4A4	-1.0845	12860	0.001865	0.1043	-3.5152
OTUD3	-1.0795	12859	0.001554	0.1014	-1.3159
PCNP	-1.0766	12858	0.003108	0.1141	-1.4412
TDRD7	-1.0745	12857	0.000311	0.07173	-1.6619
GOLGA8IP	-1.0709	12856	0.002486	0.1089	-1.752
PPP1R11	-1.0625	12855	0.002797	0.1126	-1.2753
VPS13A	-1.0576	12854	0.000622	0.08198	-1.4501
RHOC	-1.0559	12853	0.003108	0.1141	-1.7934
PSMD7	-1.0548	12852	0.002486	0.1089	-1.2665
CRBN	-1.0508	12851	0.001865	0.1043	-1.2858
ANKRD49	-1.0451	12850	0.002486	0.1089	-1.5034
HNRNPU	-1.0436	12849	0.000933	0.0867	-1.2573
VAV3	-1.0424	12848	0.002797	0.1126	-2.187
ARL5A	-1.0387	12847	0.001243	0.09507	-1.5866
TOMM70A	-1.0354	12846	0.000311	0.07173	-1.3367
PPP3CA	-1.0337	12845	0.002176	0.1073	-1.8039
RPS6	-1.0334	12844	0.001243	0.09507	-1.0821

Gene	Signal to noise	Rank	p-value	FDR(BH)	Fold Change
ZEB1	1.1401	80	0.000622	0.08198	1.5586
CAMK4	1.1362	81	0.001554	0.1014	1.6948
LMTK2	1.1306	82	0.002486	0.1089	1.3387
MAP2K2	1.129	83	0.001243	0.09507	1.3106
MOBP	1.1274	84	0.001554	0.1014	1.1864
SPANXA1	1.1266	85	0.000932	0.0867	1.3228
SLC22A14	1.1229	86	0.002176	0.1073	1.2741
CNTFR	1.1227	87	0.001554	0.1014	1.2208
CD28	1.1224	88	0.001243	0.09507	1.8519
CDH7	1.122	89	0.003108	0.1141	1.1919
NGFRAP1	1.1219	90	0.000622	0.08198	2.2196
NAB2	1.1187	91	0.000622	0.08198	1.37
CDKN1C	1.1186	92	0.001865	0.1043	1.1926
OBSL1	1.1151	93	0.002486	0.1089	1.2483
FLG	1.1062	94	0.000622	0.08198	1.274
ASNA1	1.1055	95	0.001865	0.1043	1.374
UBE2S	1.1038	96	0.000622	0.08198	1.5641
EPB41	1.0982	97	0.001554	0.1014	1.3944
OGDH	1.0981	98	0.002486	0.1089	1.3711
KCNJ16	1.098	99	0.000932	0.0867	1.1944
TRAF3IP1	1.098	100	0.002797	0.1126	1.2112
IQSEC3	1.0976	101	0.000932	0.0867	1.1835
GAST	1.095	102	0.002176	0.1073	1.235
PASK	1.0923	103	0.002176	0.1073	3.0861
MED15	1.0916	104	0.000622	0.08198	1.4552
A4GALT	1.091	105	0.001243	0.09507	1.2427
DIO3	1.0903	106	0.002797	0.1126	1.2144
GREB1	1.089	107	0.001243	0.09507	1.3115
C9orf53	1.0879	108	0.001865	0.1043	1.3055
MMP26	1.0863	109	0.004662	0.1285	1.1701
POU2AF1	1.0855	110	0.001243	0.09507	2.2527
HSPB3	1.0851	111	0.001554	0.1014	1.1716
HOXD3	1.0828	112	0.003108	0.1141	1.285
NOS2	1.0802	113	0.002176	0.1073	1.2863
ZC2HC1A	1.0797	114	0.000932	0.0867	1.952
CYP4F11	1.0791	115	0.003108	0.1141	1.4053
NR2F2	1.0777	116	0.000622	0.08198	1.3708
LIMD2	1.0731	117	0.001554	0.1014	1.5658
AKAP6	1.0721	118	0.001554	0.1014	1.2803
GADD45GIP	1.0705	119	0.000311	0.07173	1.2829
LILRB4	1.0689	120	0.002176	0.1073	1.2383
LYL1	1.0671	121	0.001865	0.1043	1.3297
PAX3	1.0641	122	0.002486	0.1089	1.2659
GUCY1A2	1.064	123	0.003419	0.1181	1.2075
ANKRD2	1.0629	124	0.001243	0.09507	1.3197
PDIA5	1.059	125	0.002486	0.1089	1.4642
HLA-DOA	1.0575	126	0.001554	0.1014	1.672
ESPN	1.057	127	0.002176	0.1073	1.9585
CLK3	1.0569	128	0.000622	0.08198	1.3383
USE1	1.0559	129	0.002797	0.1126	1.3419
RPL4	1.0497	130	0.003108	0.1141	1.1454
RFPL2	1.0473	131	0.000932	0.0867	1.3008
ARNT2	1.0455	132	0.000311	0.07173	1.1687
CDX1	1.0453	133	0.001554	0.1014	1.2712
MID1IP1	1.0447	134	0.001554	0.1014	1.654
CCDC40	1.0436	135	0.002797	0.1126	1.2232
CPNE7	1.043	136	0.001243	0.09507	1.3022
CCNL2	1.0409	137	0.001243	0.09507	1.3391
CCDC94	1.0401	138	0.000932	0.0867	1.4143
EDF1	1.039	139	0.00404	0.1192	1.1932
GDF10	1.0383	140	0.002486	0.1089	1.2596
ATXN3L	1.0379	141	0.001243	0.09507	1.2475
GPRC5A	1.037	142	0.000932	0.0867	1.2553
GNA15	1.0357	143	0.003108	0.1141	1.6005
RFX1	1.0357	144	0.002176	0.1073	1.3427
PARD6A	1.0353	145	0.001243	0.09507	1.332
CSF1	1.0351	146	0.002176	0.1073	1.2618
UXS1	1.0347	147	0.001243	0.09507	1.5698
BRD4	1.0326	148	0.00404	0.1192	1.2785
KRT18	1.0322	149	0.002486	0.1089	1.296
SLC15A1	1.0307	150	0.001865	0.1043	1.201
CYP1A1	1.0302	151	0.001865	0.1043	1.1418
MIR1204	1.029	152	0.003108	0.1141	1.8742
MAGEB1	1.0278	153	0.002176	0.1073	1.2611
KRT20	1.0261	154	0.002176	0.1073	1.1827
B3GNT3	1.0254	155	0.002797	0.1126	1.2043
FGF3	1.0251	156	0.002176	0.1073	1.2446
TACC3	1.0244	157	0.003419	0.1181	1.8655
UBXN1	1.0217	158	0.001865	0.1043	1.2283

Gene	Signal to noise	Rank	p-value	FDR(BH)	Fold Change
RPS27A	-1.0319	12843	0.002797	0.1126	-1.1217
RFC5	-1.0312	12842	0.002797	0.1126	-1.2716
CAMK2B	-1.0289	12841	0.001865	0.1043	-1.1364
DICER1	-1.0285	12840	0.002486	0.1089	-1.2502
EIF4E	-1.0221	12839	0.002486	0.1089	-1.3652
CD164	-1.0194	12838	0.001865	0.1043	-1.3125
FAM49A	-1.0164	12837	0.001865	0.1043	-2.3099
RAB11FIP5	-1.0158	12836	0.003108	0.1141	-1.8307
SRSF11	-1.0158	12835	0.003419	0.1181	-1.4789
MAF	-1.0125	12834	0.002797	0.1126	-2.0716
ARF6	-1.0124	12833	0.002176	0.1073	-1.2949
SESN1	-1.0119	12832	0.002176	0.1073	-2.314
UFM1	-1.0119	12831	0.001554	0.1014	-1.347
CCL5	-1.009	12830	0.00373	0.1192	-1.1487
SEC22B	-1.0085	12829	0.003108	0.1141	-1.1817
ABI1	-1.0074	12828	0.003419	0.1181	-1.2485
MAPK1	-1.0021	12827	0.007148	0.1493	-1.2513
ZNF562	-1.0021	12826	0.002486	0.1089	-1.359
XPNPEP2	-0.9975	12825	0.006527	0.1462	-1.621
CLCF1	-0.9969	12824	0.003108	0.1141	-1.7058
TGFBR3	-0.9961	12823	0.003108	0.1141	-1.7999
TM7SF3	-0.9943	12822	0.001554	0.1014	-1.3436
SNX24	-0.9861	12821	0.00373	0.1192	-1.8719
ROCK1	-0.9852	12820	0.00373	0.1192	-1.251
PTCH1	-0.9827	12819	0.003108	0.1141	-5.1609
GZMH	-0.9823	12818	0.00373	0.1192	-2.4241
DZANK1	-0.9801	12817	0.00404	0.1192	-1.5514
NONO	-0.9772	12816	0.002797	0.1126	-1.1576
CALR	-0.9771	12815	0.003419	0.1181	-1.2227
PRPF38B	-0.977	12814	0.00373	0.1192	-1.2534
FAM18B1	-0.9766	12813	0.001865	0.1043	-1.4417
RAP2B	-0.9754	12812	0.007148	0.1493	-1.2612
EEF1B2	-0.9744	12811	0.001554	0.1014	-1.1756
FAM8A1	-0.9731	12810	0.004973	0.1317	-1.4963
HUS1	-0.9728	12809	0.002486	0.1089	-1.3909
BICD2	-0.9685	12808	0.005594	0.1372	-1.5019
LOC100509755	-0.9648	12807	0.002486	0.1089	-1.6933
SETBP1	-0.9643	12806	0.00373	0.1192	-2.2191
SOX13	-0.9633	12805	0.005594	0.1372	-2.2081
PRICKLE4	-0.9609	12804	0.001865	0.1043	-1.3859
SLC35G2	-0.9607	12803	0.004973	0.1317	-2.0201
MGEA5	-0.9597	12802	0.004662	0.1285	-1.2435
MAN1A1	-0.9571	12801	0.00404	0.1192	-2.1058
UBQLN2	-0.957	12800	0.002486	0.1089	-1.5065
DSTN	-0.9564	12799	0.002486	0.1089	-1.5771
RBM26	-0.9556	12798	0.001865	0.1043	-1.4086
EPS15	-0.9554	12797	0.00373	0.1192	-1.3043
KLRG1	**-0.9548**	**12796**	**0.00404**	**0.1192**	**-2.5442**
DCUN1D1	-0.9546	12795	0.003419	0.1181	-1.2839
FUBP3	-0.9538	12794	0.00373	0.1192	-1.4808
RAP1A	-0.9502	12793	0.00404	0.1192	-1.3669
SPG11	-0.9481	12792	0.007459	0.1513	-2.0775
FBXO9	-0.9448	12791	0.001865	0.1043	-1.3228
PAK6	-0.9418	12790	0.00373	0.1192	-1.697
NKG7	-0.9405	12789	0.00373	0.1192	-1.5884
MED4	-0.9381	12788	0.006527	0.1462	-1.4327
CASP8	-0.9334	12787	0.003108	0.1141	-1.5232
CTAGE5	-0.933	12786	0.005594	0.1372	-1.2702
HIPK1	-0.9311	12785	0.00404	0.1192	-1.3468
FLJ11292	-0.93	12784	0.004973	0.1317	-1.204
RCOR3	-0.9298	12783	0.00404	0.1192	-1.3502
SRP9	-0.9278	12782	0.001243	0.09507	-1.3765
PXN	-0.9265	12781	0.004662	0.1285	-1.3009
TMED9	-0.9255	12780	0.001243	0.09507	-1.3187
DDR1	-0.9248	12779	0.002797	0.1126	-1.2692
NUP153	-0.9216	12778	0.01336	0.1857	-1.3117
IFNGR1	-0.9212	12777	0.001554	0.1014	-1.7701
CHSY1	-0.9212	12776	0.005905	0.1406	-1.6708
RAP1GAP2	-0.9187	12775	0.004973	0.1317	-1.603
CDK8	-0.9187	12774	0.004973	0.1317	-1.3357
CX3CR1	-0.9172	12773	0.005905	0.1406	-5.8299
PPP2R5E	-0.9132	12772	0.002797	0.1126	-1.3329
LPAL2	-0.9117	12771	0.005905	0.1406	-2.0607
FAM3C	-0.9075	12770	0.001243	0.09507	-1.6633
TAF7	-0.9073	12769	0.001865	0.1043	-1.3675
KDM2A	-0.9059	12768	0.005594	0.1372	-1.3309
MCTP2	-0.9057	12767	0.006527	0.1462	-1.9686
UBP1	-0.9056	12766	0.004351	0.1241	-1.2424
FAM65A	-0.9054	12765	0.002486	0.1089	-1.2839

Gene	Signal to noise	Rank	p-value	FDR(BH)	Fold Change
ADAM5P	1.0197	159	0.002176	0.1073	1.1995
TBXA2R	1.0191	160	0.002797	0.1126	1.5346
IL20RA	1.0189	161	0.002797	0.1126	1.2137
PDCD1	1.0186	162	0.002486	0.1089	1.3992
CENPO	1.0167	163	0.002176	0.1073	1.5639
ADRBK1	1.0157	164	0.000932	0.0867	1.4631
LMF2	1.0135	165	0.002486	0.1089	1.1616
GPR88	1.0131	166	0.001554	0.1014	1.1913
ADCY3	1.0122	167	0.001243	0.09507	1.4586
LOC284244	1.0118	168	0.00404	0.1192	1.2667
FKBP6	1.0112	169	0.00373	0.1192	1.2632
PFKL	1.011	170	0.001554	0.1014	1.16
SURF2	1.0104	171	0.008392	0.159	1.3388
CCDC64	1.01	172	0.002176	0.1073	1.7529
LMNB2	1.0084	173	0.000932	0.0867	1.3568
PYY	1.0067	174	0.00373	0.1192	1.2152
CLPB	1.0053	175	0.000932	0.0867	1.3478
RBM47	1.0046	176	0.003108	0.1141	1.3296
PRRG4	1.0034	177	0.003108	0.1141	1.4913
BST1	1.0023	178	0.000932	0.0867	1.1376
SERPINA7	0.9993	179	0.002176	0.1073	1.1228
ICOS	0.9978	180	0.00373	0.1192	2.0859
USP36	0.9968	181	0.000311	0.07173	1.8693
HTR5A	0.996	182	0.001243	0.09507	1.2411
ITM2A	0.9947	183	0.002176	0.1073	1.7762
SPRY1	0.9938	184	0.004662	0.1285	2.5022
NR1I2	0.9916	185	0.004662	0.1285	1.2197
FBRS	0.9899	186	0.002486	0.1089	1.2252
PDZRN3	0.9897	187	0.002797	0.1126	1.2172
CCNE1	0.9869	188	0.003108	0.1141	2.2869
FSD1	0.9858	189	0.003108	0.1141	1.2877
DTX2	0.984	190	0.003108	0.1141	1.3941
HGC6.3	0.9832	191	0.00373	0.1192	1.1836
C1QL1	0.9824	192	0.00404	0.1192	1.1573
HOXD9	0.9821	193	0.00404	0.1192	1.1854
SCNN1A	0.9811	194	0.002797	0.1126	1.2033
CDK5RAP2	0.981	195	0.004351	0.1241	1.2477
TM4SF5	0.9805	196	0.002486	0.1089	1.2603
RPL8	0.9799	197	0.00404	0.1192	1.1011
SHMT2	0.9797	198	0.000932	0.0867	1.5858
CEACAM7	0.9788	199	0.01088	0.1731	1.2182
CHAC1	0.9782	200	0.002176	0.1073	1.2186
KLF13	0.978	201	0.003108	0.1141	1.3877
PPP2R1A	0.976	202	0.00373	0.1192	1.2463
OR7A17	0.973	203	0.00373	0.1192	1.3091
SRM	0.9723	204	0.002486	0.1089	1.3561
EIF3C	0.9721	205	0.00404	0.1192	1.2533
SLC17A2	0.9699	206	0.00404	0.1192	1.1347
CNN1	0.9698	207	0.003419	0.1181	1.2449
LOC441666	0.9687	208	0.001865	0.1043	1.4797
ABO	0.9686	209	0.002176	0.1073	1.186
MRPL12	0.965	210	0.000622	0.08198	1.3494
RAD54L	0.9647	211	0.000932	0.0867	1.4297
HLA-DRA	0.9644	212	0.00404	0.1192	2.5181
OSGIN1	0.964	213	0.00404	0.1192	1.2327
TRIM29	0.9618	214	0.002486	0.1089	1.2643
TBX4	0.9603	215	0.002486	0.1089	1.2191
FOXL1	0.96	216	0.003419	0.1181	1.2611
P2RX1	0.9597	217	0.002176	0.1073	1.2307
SLC19A1	0.9586	218	0.000932	0.0867	1.4771
RBFOX1	0.9576	219	0.000622	0.08198	1.2159
EHMT1	0.9574	220	0.001554	0.1014	1.2788
PA2G4	0.956	221	0.00404	0.1192	1.2223
COL6A3	0.9549	222	0.002176	0.1073	1.4518
WIF1	0.9548	223	0.004973	0.1317	1.3382
DCLK1	0.9531	224	0.00373	0.1192	1.1766
CSF3R	0.9518	225	0.004351	0.1241	1.3572
IL13	0.9513	226	0.005594	0.1372	1.2194
MIR3917	0.9512	227	0.004351	0.1241	1.1955
MYF5	0.9508	228	0.000932	0.0867	1.1658
DGCR6	0.9507	229	0.002797	0.1126	1.265
GIF	0.9502	230	0.001865	0.1043	1.2618
GNB1L	0.949	231	0.005594	0.1372	1.2421
OC100131319	0.9486	232	0.002797	0.1126	1.1511
CATSPERG	0.9482	233	0.00404	0.1192	1.2364
PBK	0.9477	234	0.004662	0.1285	1.5127
PKDREJ	0.9458	235	0.006216	0.1445	1.1755
AMFR	0.9452	236	0.004973	0.1317	1.2653
MFGE8	0.9448	237	0.00373	0.1192	1.6769

Gene	Signal to noise	Rank	p-value	FDR(BH)	Fold Change
ADCY7	-0.9049	12764	0.006838	0.1475	-1.4457
ZNF358	-0.9047	12763	0.003108	0.1141	-1.2506
B3GAT1	-0.9029	12762	0.006527	0.1462	-2.1587
VPS13C	-0.9027	12761	0.006838	0.1475	-1.3402
KLRF1	-0.9015	12760	0.004351	0.1241	-4.9476
BCL2L13	-0.9	12759	0.005284	0.1352	-1.434
PDS5B	-0.8955	12758	0.001243	0.09507	-1.4323
FYCO1	-0.8949	12757	0.004973	0.1317	-1.3311
PPP2R2B	-0.894	12756	0.006838	0.1475	-1.9443
KTN1	-0.893	12755	0.007459	0.1513	-1.3391
PDE4DIP	-0.8921	12754	0.00404	0.1192	-1.4221
KIF1B	-0.8909	12753	0.004662	0.1285	-1.2349
MBNL1	-0.8885	12752	0.007459	0.1513	-1.5306
LTN1	-0.8846	12751	0.005905	0.1406	-1.3471
PLEK	-0.8829	12750	0.008081	0.1559	-2.7379
STX12	-0.8821	12749	0.01368	0.187	-1.2407
PPID	-0.8815	12748	0.005284	0.1352	-1.2824
HERC5	-0.8798	12747	0.004351	0.1241	-1.9733
WDR44	-0.8795	12746	0.005594	0.1372	-1.2969
VPS4B	-0.8774	12745	0.006216	0.1445	-1.3531
ADAM8	-0.8771	12744	0.00404	0.1192	-1.3077
TPP1	-0.8765	12743	0.00404	0.1192	-1.4228
GNAZ	-0.8741	12742	0.003108	0.1141	-1.4217
KANSL1L	-0.8712	12741	0.007459	0.1513	-1.5418
ZNHIT6	-0.8663	12740	0.009013	0.1634	-1.3287
DPY19L1	-0.8661	12739	0.00404	0.1192	-1.9074
ELF2	-0.8648	12738	0.00777	0.1531	-1.4808
TMA16	-0.8647	12737	0.005905	0.1406	-1.6161
UBE2G1	-0.8641	12736	0.005905	0.1406	-1.3074
CYP3A5	-0.863	12735	0.005594	0.1372	-1.367
RPL27	-0.8594	12734	0.005905	0.1406	-1.1147
SMAP1	-0.8588	12733	0.00404	0.1192	-1.2642
LOC10050639	-0.8585	12732	0.007459	0.1513	-1.2565
MICB	-0.8585	12731	0.007148	0.1493	-1.622
VPS26A	-0.8539	12730	0.006527	0.1462	-1.4133
CDC42SE1	-0.8521	12729	0.009324	0.1635	-1.4556
PIGF	-0.8521	12728	0.004973	0.1317	-1.3481
GOLGA6L3	-0.8514	12727	0.00777	0.1531	-1.8726
PLOD1	-0.8497	12726	0.009635	0.1645	-1.6111
RP2	-0.8496	12725	0.005905	0.1406	-1.4911
DSTNP2	-0.8463	12724	0.000933	0.0867	-1.2528
TRIM3	-0.8459	12723	0.005905	0.1406	-1.2204
AP5M1	-0.8456	12722	0.006838	0.1475	-1.3615
SLC9A3R1	-0.8452	12721	0.00777	0.1531	-1.4196
PTPN12	-0.844	12720	0.01026	0.1686	-1.9017
C17orf101	-0.8429	12719	0.006527	0.1462	-1.3705
TMEM48	-0.8399	12718	0.00777	0.1531	-1.2618
AMIGO2	-0.8396	12717	0.007459	0.1513	-1.5347
LYN	-0.8382	12716	0.008081	0.1559	-2.0724
PHF8	-0.838	12715	0.008392	0.159	-1.1918
ARCN1	-0.8374	12714	0.008392	0.159	-1.1725
NAA38	-0.8369	12713	0.01212	0.179	-1.2074
UEVLD	-0.8363	12712	0.00373	0.1192	-1.5562
RGS3	-0.8339	12711	0.009013	0.1634	-1.9742
CRIPT	-0.8298	12710	0.01958	0.2064	-1.4187
FZD4	-0.8291	12709	0.01399	0.1875	-1.2741
ARGLU1	-0.8291	12708	0.009635	0.1645	-1.3932
SLC17A5	-0.8289	12707	0.00777	0.1531	-1.4157
CRY1	-0.8279	12706	0.0115	0.1753	-1.525
RNF44	-0.8275	12705	0.009946	0.1667	-1.3292
TMED10	-0.827	12704	0.001865	0.1043	-1.4075
ACO1	-0.8245	12703	0.008081	0.1559	-1.6125
GGNBP2	-0.8231	12702	0.005284	0.1352	-1.2297
RCAN1	-0.8229	12701	0.00373	0.1192	-1.8819
EMC2	-0.8228	12700	0.009946	0.1667	-1.337
PPM1B	-0.8225	12699	0.008392	0.159	-1.342
SET	-0.8222	12698	0.01026	0.1686	-1.259
HNRNPA3	-0.821	12697	0.00404	0.1192	-1.1585
RPL7	-0.8186	12696	0.002486	0.1089	-1.0672
NHLRC2	-0.8185	12695	0.01119	0.1746	-1.5975
NUCKS1	-0.8172	12694	0.007459	0.1513	-1.424
LRIF1	-0.8168	12693	0.009013	0.1634	-1.527
TARP	-0.8164	12692	0.009635	0.1645	-1.5205
FBXL5	-0.8157	12691	0.01212	0.179	-1.3921
BCL10	-0.8137	12690	0.005284	0.1352	-1.5731
TGFBR1	-0.8136	12689	0.009946	0.1667	-2.2977
ZNF480	-0.8133	12688	0.01305	0.1838	-1.3875
PAPOLG	-0.8117	12687	0.009324	0.1635	-1.2316
TSPYL1	-0.8108	12686	0.01616	0.195	-1.3056

Gene	Signal to noise	Rank	p-value	FDR(BH)	Fold Change
GPR162	0.943	238	0.002797	0.1126	1.178
SLC6A1	0.9404	239	0.004973	0.1317	1.1848
ADRBK2	0.9401	240	0.002486	0.1089	1.5054
FAM108B1	0.9388	241	0.004662	0.1285	1.4267
AIF1	0.9386	242	0.00404	0.1192	2.0438
PNMA2	0.9384	243	0.003108	0.1141	1.2169
MYB	0.9376	244	0.003419	0.1181	2.084
TOP2A	0.9373	245	0.00404	0.1192	3.4747
CCR4	0.9371	246	0.003419	0.1181	2.0074
SLC25A6	0.9357	247	0.004351	0.1241	1.1914
GEMIN7	0.9356	248	0.004351	0.1241	1.2278
PXDC1	0.9352	249	0.004973	0.1317	1.2061
MYH6	0.9332	250	0.005284	0.1352	1.2161
MSX1	0.9326	251	0.001865	0.1043	1.26
SLC13A3	0.9324	252	0.00404	0.1192	1.1351
CDC37	0.9299	253	0.004662	0.1285	1.3827
SERPINA3	0.9299	254	0.00404	0.1192	1.2458
CNN2	0.9295	255	0.004351	0.1241	1.2425
FMO3	0.9275	256	0.00373	0.1192	1.2835
EEF2	0.9253	257	0.002486	0.1089	1.1396
PHLDA3	0.9252	258	0.008702	0.1604	1.159
EIF3H	0.925	259	0.007459	0.1513	1.1926
ATXN7	0.924	260	0.00373	0.1192	1.2966
PHOX2B	0.9238	261	0.002797	0.1126	1.1766
MGAT4B	0.9225	262	0.005284	0.1352	1.3074
CWH43	0.922	263	0.006527	0.1462	1.2196
SEC14L2	0.9212	264	0.00373	0.1192	1.4728
ZNF592	0.9194	265	0.002797	0.1126	1.2472
TRPC4	0.9186	266	0.001865	0.1043	1.1755
ACOT1	0.918	267	0.004662	0.1285	1.2194
RPS6KB1	0.9169	268	0.003108	0.1141	1.2664
ADAMTS13	0.9169	269	0.004662	0.1285	1.2145
CD7	0.9165	270	0.004662	0.1285	1.6204
HNF4A	0.9141	271	0.00373	0.1192	1.1812
CNTLN	0.9138	272	0.005284	0.1352	1.2451
NDST4	0.9123	273	0.007148	0.1493	1.1649
TPX2	0.9121	274	0.00777	0.1531	3.4816
PNPLA6	0.9118	275	0.006838	0.1475	1.204
POLDIP3	0.9099	276	0.005594	0.1372	1.3296
TWF2	0.9091	277	0.003108	0.1141	1.2327
CFHR5	0.9088	278	0.006527	0.1462	1.24
DZIP1	0.9079	279	0.008392	0.159	1.2301
CDT1	0.9076	280	0.01119	0.1746	1.5118
EDAR	0.9075	281	0.006216	0.1445	1.3232
RBFOX2	0.9073	282	0.003419	0.1181	1.3649
PHEX	0.9058	283	0.00404	0.1192	2.061
CYP4B1	0.905	284	0.007459	0.1513	1.1752
RNF126P1	0.9047	285	0.008081	0.1559	1.323
PON3	0.9047	286	0.001865	0.1043	1.169
GRIK2	0.9041	287	0.005594	0.1372	1.2133
C17orf48	0.9041	288	0.006216	0.1445	1.5007
R3HDM4	0.9036	289	0.005284	0.1352	1.3674
ZNF335	0.9035	290	0.006838	0.1475	1.5271
TBXAS1	0.9024	291	0.007148	0.1493	1.3498
ZBTB48	0.9019	292	0.006216	0.1445	1.4256
MCHR1	0.9018	293	0.005594	0.1372	1.2567
PRAME	0.9011	294	0.004662	0.1285	1.2288
MED8	0.9005	295	0.003419	0.1181	1.3126
MMP7	0.9002	296	0.002176	0.1073	1.1748
SMOX	0.8996	297	0.003419	0.1181	1.3654
AKT2	0.8991	298	0.005594	0.1372	1.3288
KRT6A	0.899	299	0.005284	0.1352	1.1826
ASCL2	0.8975	300	0.003419	0.1181	1.3546
OC100131822	0.8953	301	0.00777	0.1531	1.2836
OTUB1	0.8949	302	0.006527	0.1462	1.3654
MRC1	0.8948	303	0.00404	0.1192	1.2301
SLC5A12	0.8936	304	0.00777	0.1531	1.1974
RAB9BP1	0.8935	305	0.004662	0.1285	1.3454
PEG3	0.8924	306	0.006216	0.1445	1.2696
AKAP8L	0.8915	307	0.007148	0.1493	1.2535
ETV4	0.8911	308	0.004973	0.1317	1.2162
GCKR	0.8889	309	0.003419	0.1181	1.1514
CSPG5	0.888	310	0.004351	0.1241	1.2179
IL23A	0.8877	311	0.00404	0.1192	1.2228
FAM198B	0.8873	312	0.006216	0.1445	1.2013
LOC254896	0.8867	313	0.008702	0.1604	1.2099
RND3	0.8857	314	0.008081	0.1559	1.182
LPPR4	0.8855	315	0.004351	0.1241	1.1492
DHRS9	0.8851	316	0.000932	0.0867	1.192

Gene	Signal to noise	Rank	p-value	FDR(BH)	Fold Change
C1orf21	-0.8105	12685	0.01181	0.1777	-2.6254
CD47	-0.81	12684	0.00777	0.1531	-1.2815
OSTM1	-0.8099	12683	0.008702	0.1604	-1.4013
YTHDF1	-0.8088	12682	0.008702	0.1604	-1.1851
ACTL6A	-0.8075	12681	0.009324	0.1635	-1.4554
UBA3	-0.8047	12680	0.008392	0.159	-1.2181
WWP2	-0.8035	12679	0.006838	0.1475	-1.2622
HERC3	-0.8033	12678	0.009946	0.1667	-1.2637
CELSR2	-0.803	12677	0.007459	0.1513	-1.6014
PLCB1	-0.8008	12676	0.01181	0.1777	-3.1538
RCBTB2	-0.8007	12675	0.009324	0.1635	-1.8871
CD81	-0.7996	12674	0.01243	0.1803	-1.2465
RPS11	-0.7987	12673	0.009635	0.1645	-1.0302
TMEM126B	-0.7979	12672	0.01181	0.1777	-1.3792
EFR3A	-0.7979	12671	0.0115	0.1753	-1.2354
TROVE2	-0.7955	12670	0.008081	0.1559	-1.1987
CERS2	-0.7951	12669	0.01368	0.187	-1.2355
NABP1	-0.7931	12668	0.01026	0.1686	-1.4232
A2M	-0.7923	12667	0.01119	0.1746	-2.4403
EPB41L4A	-0.7909	12666	0.009635	0.1645	-1.8784
TRIM5	-0.7907	12665	0.009013	0.1634	-1.6138
CMPK1	-0.7904	12664	0.009324	0.1635	-1.3349
AGPAT4	-0.7896	12663	0.01647	0.1959	-2.0978
ZBTB11	-0.7887	12662	0.006838	0.1475	-1.3935
LOC10013087	-0.7878	12661	0.01399	0.1875	-1.8434
AZI2	-0.7871	12660	0.01523	0.1909	-1.3705
LEPROT	-0.7852	12659	0.01057	0.1701	-1.4715
PSMG1	-0.7851	12658	0.01368	0.187	-1.3542
LINC00597	-0.7841	12657	0.009013	0.1634	-1.6666
ATAD2B	-0.7828	12656	0.01305	0.1838	-1.2015
H3F3A	-0.7826	12655	0.0143	0.1887	-1.1135
TUBGCP3	-0.782	12654	0.009635	0.1645	-1.2812
CRIM1	-0.7814	12653	0.01368	0.187	-1.6442
WDR19	-0.781	12652	0.01057	0.1701	-1.5521
PALLD	-0.781	12651	0.01399	0.1875	-2.1123
STUB1	-0.7806	12650	0.01368	0.187	-1.1269
MXI1	-0.7804	12649	0.01399	0.1875	-1.4047
FBXO3	-0.7789	12648	0.01274	0.1814	-1.3996
C4orf34	-0.7789	12647	0.0258	0.2276	-1.454
BZW1	-0.7785	12646	0.01709	0.198	-1.2064
FRYL	-0.7768	12645	0.008081	0.1559	-1.2554
TBX21	-0.7759	12644	0.01243	0.1803	-1.8154
ZDHHC14	-0.7757	12643	0.01088	0.1731	-1.4872
VAMP4	-0.7749	12642	0.008392	0.159	-1.3896
SMPD3	-0.7747	12641	0.01212	0.179	-1.1833
PZP	-0.7744	12640	0.01274	0.1814	-2.8802
BTBD1	-0.774	12639	0.01181	0.1777	-1.2069
TBKBP1	-0.7733	12638	0.004973	0.1317	-1.4075
CUL4A	-0.773	12637	0.009946	0.1667	-1.3598
ZMIZ2	-0.772	12636	0.01243	0.1803	-1.4863
C2orf27A	-0.7709	12635	0.008702	0.1604	-1.1943
CELF1	-0.7688	12634	0.007459	0.1513	-1.2048
SAR1A	-0.7685	12633	0.006527	0.1462	-1.2319
MSN	-0.7682	12632	0.004973	0.1317	-1.2308
CUTC	-0.767	12631	0.005284	0.1352	-1.2251
NFATC2IP	-0.7667	12630	0.01274	0.1814	-1.1715
MKNK1	-0.7664	12629	0.0143	0.1887	-1.429
GZMB	-0.7663	12628	0.0115	0.1753	-3.2611
ORC4	-0.7655	12627	0.01274	0.1814	-1.3269
THAP1	-0.7649	12626	0.01554	0.192	-1.3123
MORF4L1	-0.7647	12625	0.01523	0.1909	-1.1365
STYK1	-0.7641	12624	0.01119	0.1746	-2.3311
SERINC1	-0.7638	12623	0.02269	0.2183	-1.2989
DOCK5	-0.7631	12622	0.01523	0.1909	-2.0831
SLC31A2	-0.763	12621	0.01523	0.1909	-1.3529
PLEKHA1	-0.7626	12620	0.01305	0.1838	-1.3743
LILRB1	-0.7621	12619	0.01678	0.1965	-1.6011
C8orf44-SGK	-0.761	12618	0.01647	0.1959	-1.7672
EZH1	-0.7601	12617	0.01305	0.1838	-1.156
ZNF395	-0.7599	12616	0.01523	0.1909	-1.8539
RNF11	-0.7599	12615	0.01616	0.195	-1.4829
LAMP2	-0.7596	12614	0.008392	0.159	-1.4389
AGAP1	-0.7586	12613	0.01492	0.1909	-3.9747
B2M	-0.7586	12612	0.01212	0.179	-1.1184
CDC73	-0.7577	12611	0.01927	0.2046	-1.3323
IQGAP2	-0.7574	12610	0.01305	0.1838	-1.3628
C1orf56	-0.7572	12609	0.006216	0.1445	-1.3958
ALMS1	-0.7571	12608	0.01554	0.192	-1.224
KDELR2	-0.7559	12607	0.01336	0.1857	-1.3606

Gene	Signal to noise	Rank	p-value	FDR(BH)	Fold Change
MLNR	0.8848	317	0.003419	0.1181	1.3299
LUC7L	0.8845	318	0.00777	0.1531	1.3052
ARHGDIA	0.8844	319	0.00404	0.1192	1.39
CHMP7	0.8836	320	0.006527	0.1462	1.4306
HMGCS2	0.8836	321	0.006216	0.1445	1.1126
DUSP4	0.8836	322	0.01026	0.1686	5.8167
TFDP3	0.8832	323	0.003419	0.1181	1.1783
BTBD2	0.883	324	0.002797	0.1126	1.3731
BSG	0.8824	325	0.004351	0.1241	1.3327
SCGN	0.8806	326	0.00404	0.1192	1.158
NAT9	0.8804	327	0.005594	0.1372	1.2376
PPP1R7	0.8799	328	0.006838	0.1475	1.3831
MMP13	0.8786	329	0.007459	0.1513	1.1986
DDAH1	0.8767	330	0.005905	0.1406	1.2569
OR2J2	0.876	331	0.004351	0.1241	1.2009
CLN6	0.8746	332	0.005284	0.1352	1.3338
STMN3	0.8746	333	0.006527	0.1462	1.3154
MYCNOS	0.8742	334	0.006527	0.1462	1.2588
GDNF	0.8731	335	0.003419	0.1181	1.1968
NRXN3	0.8723	336	0.005905	0.1406	1.2132
CCNJL	0.8708	337	0.003108	0.1141	1.1757
BCAT1	0.8707	338	0.00404	0.1192	3.7141
GPR85	0.8696	339	0.003419	0.1181	1.1834
DKKL1	0.8683	340	0.006527	0.1462	1.2551
KRT15	0.8683	341	0.006838	0.1475	1.213
GATA4	0.867	342	0.005284	0.1352	1.1507
ISG20	0.8665	343	0.00777	0.1531	1.3241
HRH1	0.8662	344	0.00404	0.1192	1.1649
VAV2	0.8658	345	0.002176	0.1073	1.1995
EFNA3	0.8646	346	0.00404	0.1192	1.2264
AURKAIP1	0.8636	347	0.005594	0.1372	1.3874
DLGAP1	0.8623	348	0.006216	0.1445	1.277
CYP2A6	0.8621	349	0.00777	0.1531	1.2106
FAM134B	0.8621	350	0.006838	0.1475	1.7984
TNP1	0.8619	351	0.01057	0.1701	1.1851
FKBP5	0.8617	352	0.005905	0.1406	1.8418
C16orf45	0.8615	353	0.005594	0.1372	1.497
DAAM2	0.8614	354	0.009013	0.1634	1.2679
ASAP1-IT1	0.8612	355	0.005284	0.1352	1.6268
EHD2	0.86	356	0.000932	0.0867	1.2525
MECP2	0.8589	357	0.006216	0.1445	1.4147
HMG20B	0.8576	358	0.008081	0.1559	1.5495
MYO1E	0.8573	359	0.006838	0.1475	1.2328
ASCL1	0.8535	360	0.007459	0.1513	1.214
JAG1	0.8535	361	0.001554	0.1014	2.0002
SRGAP3	0.8534	362	0.000311	0.07173	1.2567
EIF3A	0.8531	363	0.002797	0.1126	1.3818
FFAR3	0.8524	364	0.009324	0.1635	1.2036
CORO2B	0.8521	365	0.007148	0.1493	1.2505
HAO1	0.852	366	0.006527	0.1462	1.2452
CHAF1A	0.8502	367	0.006527	0.1462	1.3576
RALY	0.8502	368	0.006527	0.1462	1.2415
KRT24	0.8495	369	0.008081	0.1559	1.2136
SMARCB1	0.849	370	0.00777	0.1531	1.3684
STRN4	0.8489	371	0.008392	0.159	1.31
ALOXE3	0.8484	372	0.006527	0.1462	1.2621
MAD1L1	0.8481	373	0.00777	0.1531	1.4975
MEOX1	0.8474	374	0.007148	0.1493	1.9779
CD4	0.8471	375	0.006527	0.1462	1.2885
RS1	0.8467	376	0.006838	0.1475	1.2789
LGALS2	0.8456	377	0.007148	0.1493	1.49
PIK3CA	0.8455	378	0.005905	0.1406	1.1449
ADTRP	0.8452	379	0.003108	0.1141	1.1975
MAPK10	0.8441	380	0.003108	0.1141	1.1572
R3HCC1L	0.844	381	0.009324	0.1635	1.2942
ODAM	0.843	382	0.008702	0.1604	1.1917
MAPKBP1	0.8429	383	0.01057	0.1701	1.2888
NFRKB	0.8421	384	0.005905	0.1406	1.3413
SEC14L4	0.8419	385	0.008392	0.159	1.253
CTCF	0.8419	386	0.009635	0.1645	1.406
SIRPA	0.8401	387	0.007148	0.1493	1.2055
FEZF2	0.839	388	0.008702	0.1604	1.2669
GTF2F1	0.8388	389	0.01057	0.1701	1.3072
RAPGEF1	0.8387	390	0.008702	0.1604	1.2953
HPCAL1	0.8366	391	0.009324	0.1635	1.5645
CD248	0.8363	392	0.009324	0.1635	1.5606
NYX	0.8357	393	0.004662	0.1285	1.2242
CTSZ	0.8353	394	0.01026	0.1686	1.2386
HSPB7	0.8349	395	0.008392	0.159	1.1952

Gene	Signal to noise	Rank	p-value	FDR(BH)	Fold Change
STOM	-0.7556	12606	0.009324	0.1635	-1.78
MYBL1	-0.7556	12605	0.01368	0.187	-6.5054
FCGR3A	-0.7542	12604	0.01616	0.195	-3.0245
GTPBP3	-0.7538	12603	0.0143	0.1887	-1.22
DNAJB14	-0.7533	12602	0.00404	0.1192	-1.4024
C12orf29	-0.7533	12601	0.01647	0.1959	-1.3856
SAP30	-0.7529	12600	0.007148	0.1493	-1.2844
RTN4	-0.7528	12599	0.01119	0.1746	-1.2312
ATP8B1	-0.7528	12598	0.01368	0.187	-1.341
BMPR1A	-0.7523	12597	0.009324	0.1635	-1.7933
CDC40	-0.7518	12596	0.01243	0.1803	-1.2163
RHOQ	-0.7503	12595	0.01119	0.1746	-1.2945
YES1	-0.7501	12594	0.01492	0.1909	-2.1321
RAB7A	-0.7499	12593	0.01585	0.1936	-1.4333
ZNF506	-0.7481	12592	0.01803	0.2007	-1.2746
PRF1	-0.7476	12591	0.01399	0.1875	-1.6173
GTF2H2	-0.7475	12590	0.01274	0.1814	-1.482
TBC1D19	-0.7474	12589	0.01492	0.1909	-1.8045
ANKRA2	-0.7473	12588	0.02269	0.2183	-1.3798
CFH	-0.7471	12587	0.0202	0.209	-3.3511
COPZ2	-0.7465	12586	0.01585	0.1936	-1.4781
PLEKHA5	-0.7451	12585	0.01399	0.1875	-1.7494
FAM106A	-0.7447	12584	0.0143	0.1887	-1.2041
NIP7	-0.7444	12583	0.01616	0.195	-1.3086
GOLPH3	-0.7444	12582	0.01865	0.2027	-1.3688
ZFAND6	-0.7436	12581	0.01616	0.195	-1.2591
ATP10D	-0.7426	12580	0.0143	0.1887	-1.5936
KIR3DL2	-0.7424	12579	0.01088	0.1731	-2.3605
GOLGA8C	-0.7424	12578	0.01647	0.1959	-1.2227
MIR3656	-0.7416	12577	0.0174	0.1984	-1.2456
ZNF814	-0.741	12576	0.009946	0.1667	-1.3297
FNTA	-0.7402	12575	0.01399	0.1875	-1.1774
GNLY	-0.7399	12574	0.0202	0.209	-2.1133
MTSS1	-0.7397	12573	0.03357	0.2497	-1.728
LOC10050606	-0.7397	12572	0.00373	0.1192	-1.3067
COPS5	-0.7394	12571	0.01554	0.192	-1.2238
CD244	-0.7384	12570	0.01803	0.2007	-2.0376
BHLHE40	-0.7379	12569	0.01399	0.1875	-1.6567
NOL10	-0.7378	12568	0.01461	0.1903	-1.2145
SNX29P2	-0.7362	12567	0.01585	0.1936	-1.2768
ME3	-0.7361	12566	0.01896	0.2042	-1.9696
FKBP14	-0.7354	12565	0.01181	0.1777	-1.3177
CMKLR1	-0.7349	12564	0.01772	0.1998	-1.9028
SYF2	-0.7338	12563	0.01865	0.2027	-1.2406
LANCL1	-0.7333	12562	0.01461	0.1903	-1.3876
NCBP2	-0.7333	12561	0.01647	0.1959	-1.3043
OGFOD1	-0.7328	12560	0.01057	0.1701	-1.373
LSS	-0.7319	12559	0.01896	0.2042	-1.4379
NRAS	-0.7309	12558	0.01709	0.198	-1.2622
CYP20A1	-0.7302	12557	0.0202	0.209	-1.2403
LTBP4	-0.73	12556	0.01709	0.198	-1.2331
ERI2	-0.7297	12555	0.01865	0.2027	-1.4939
PCYOX1	-0.7296	12554	0.01585	0.1936	-1.4607
UGDH	-0.729	12553	0.01772	0.1998	-1.7916
NXPE3	-0.7283	12552	0.0174	0.1984	-1.6649
TNKS2	-0.7273	12551	0.02393	0.2235	-1.2545
EFHC1	-0.7271	12550	0.01647	0.1959	-1.6376
PDCD1LG2	-0.7265	12549	0.01492	0.1909	-1.282
ZNF24	-0.7264	12548	0.01647	0.1959	-1.1653
XYLT1	-0.7262	12547	0.01523	0.1909	-1.7513
DUSP14	-0.7258	12546	0.01274	0.1814	-1.302
COX7A2	-0.7238	12545	0.02145	0.214	-1.2455
LOC10065282	-0.723	12544	0.01523	0.1909	-1.0808
JTB	-0.7229	12543	0.01896	0.2042	-1.1645
RPL13A	-0.7218	12542	0.02051	0.2104	-1.0271
CASP1	-0.7216	12541	0.0115	0.1753	-1.5257
DAB2	-0.7215	12540	0.02051	0.2104	-1.5532
REEP4	-0.7213	12539	0.01989	0.2082	-1.3987
ZFR2	-0.7212	12538	0.01057	0.1701	-1.2261
RNF144A	-0.7204	12537	0.02113	0.2134	-3.2767
WNT10B	-0.7193	12536	0.0143	0.1887	-1.5686
RHOA	-0.7192	12535	0.01678	0.1965	-1.2955
SOAT1	-0.7186	12534	0.01399	0.1875	-1.2957
PRPF39	-0.7181	12533	0.02362	0.2232	-1.2542
LOC92249	-0.7176	12532	0.01709	0.198	-1.873
TFCP2L1	-0.7172	12531	0.02517	0.227	-1.8505
EIF2AK3	-0.717	12530	0.01772	0.1998	-1.4656
AGTPBP1	-0.7156	12529	0.01989	0.2082	-1.5281
SPCS3	-0.7145	12528	0.01803	0.2007	-1.2812

Gene	Signal to noise	Rank	p-value	FDR(BH)	Fold Change
MYL7	0.8333	396	0.009635	0.1645	1.1658
TNFAIP6	0.8331	397	0.008702	0.1604	1.3091
ATP1B4	0.8329	398	0.008081	0.1559	1.1594
GUCA2B	0.8325	399	0.006838	0.1475	1.2098
AAK1	0.8322	400	0.005594	0.1372	1.2225
SFI1	0.8321	401	0.008702	0.1604	1.2445
PRPH	0.8315	402	0.009635	0.1645	1.2165
CDKN2B	0.8313	403	0.006838	0.1475	1.2109
CLDN5	0.8311	404	0.009946	0.1667	1.1904
PHGDH	0.8309	405	0.01523	0.1909	1.5025
GNB2L1	0.8308	406	0.007459	0.1513	1.1588
PSMD8	0.8306	407	0.007148	0.1493	1.2759
FAM155A	0.8304	408	0.0115	0.1753	1.1924
PDK1	0.8299	409	0.007459	0.1513	1.2806
TNNT2	0.8289	410	0.005284	0.1352	1.1563
CWF19L1	0.8273	411	0.00777	0.1531	1.3043
INPP5K	0.8269	412	0.007148	0.1493	1.2989
HMHA1	0.8269	413	0.007148	0.1493	1.1395
NTRK2	0.8266	414	0.006216	0.1445	1.223
PNPLA3	0.8261	415	0.004973	0.1317	1.2792
IGK@	0.8259	416	0.009635	0.1645	1.4051
DRG2	0.8252	417	0.009013	0.1634	1.2005
DGKA	0.8252	418	0.01368	0.187	1.5667
FABP2	0.825	419	0.01181	0.1777	1.1185
GNB1	0.8249	420	0.01026	0.1686	1.1512
TMEM223	0.8248	421	0.009635	0.1645	1.3671
ORC1	0.8246	422	0.008081	0.1559	1.7036
KCTD12	0.8235	423	0.01243	0.1803	1.5446
FXYD5	0.8234	424	0.01057	0.1701	1.2811
AVPR1B	0.8231	425	0.009013	0.1634	1.1264
ACADL	0.8231	426	0.008702	0.1604	1.1358
FAM176B	0.8229	427	0.007148	0.1493	1.2334
LOC79999	0.8229	428	0.0115	0.1753	1.2976
LGR5	0.8229	429	0.004973	0.1317	1.1369
LRRC49	0.8218	430	0.005594	0.1372	1.1279
CHST15	0.8217	431	0.006838	0.1475	1.2886
JUP	0.8215	432	0.003419	0.1181	1.208
ZYX	0.821	433	0.009013	0.1634	1.3218
OR1A1	0.8208	434	0.004662	0.1285	1.1556
SOX14	0.8207	435	0.001865	0.1043	1.2838
CBX8	0.8205	436	0.007459	0.1513	1.2858
HSPA12A	0.8204	437	0.009635	0.1645	1.3008
RBM17	0.8202	438	0.01026	0.1686	1.195
PPP4C	0.8189	439	0.005594	0.1372	1.3227
PODXL2	0.8186	440	0.01119	0.1746	1.2607
PSMB7	0.8184	441	0.01119	0.1746	1.1813
LRRC61	0.8183	442	0.004973	0.1317	1.3211
PIWIL1	0.8176	443	0.008702	0.1604	1.2544
LINC00312	0.8172	444	0.007148	0.1493	1.4928
FOXE3	0.8159	445	0.01243	0.1803	1.288
RANBP9	0.8144	446	0.006216	0.1445	1.2622
CUL9	0.8141	447	0.01119	0.1746	1.131
MICALL1	0.8136	448	0.005284	0.1352	1.3435
DNAJB12	0.8124	449	0.006838	0.1475	1.2967
AK1	0.8122	450	0.01243	0.1803	1.2855
SPIB	0.8114	451	0.005284	0.1352	1.2771
FGF4	0.8114	452	0.006838	0.1475	1.2259
C16orf59	0.8107	453	0.004973	0.1317	1.2029
MIS18BP1	0.8105	454	0.008392	0.159	1.2636
TSSK2	0.8103	455	0.005905	0.1406	1.1851
LOC642838	0.8101	456	0.008702	0.1604	1.2178
TSPY1	0.8098	457	0.009635	0.1645	1.1684
FPR2	0.8094	458	0.01119	0.1746	1.191
CPLX2	0.8092	459	0.01088	0.1731	1.1444
GNG7	0.8092	460	0.01026	0.1686	1.3916
CSTF1	0.8084	461	0.01026	0.1686	1.5281
HIST1H3A	0.8081	462	0.0115	0.1753	1.7202
KLHL1	0.8076	463	0.01368	0.187	1.1915
WT1-AS	0.8073	464	0.01119	0.1746	1.3663
KDM5C	0.8071	465	0.009324	0.1635	1.3253
AP3B1	0.8066	466	0.008702	0.1604	1.2101
TEX13A	0.8066	467	0.009324	0.1635	1.1842
ANGPTL4	0.8061	468	0.0115	0.1753	1.248
PFDN6	0.8061	469	0.007148	0.1493	1.2493
LRRC36	0.8056	470	0.01088	0.1731	1.2087
MSMB	0.8055	471	0.008702	0.1604	1.169
NET1	0.8053	472	0.01274	0.1814	1.6685
COG7	0.8052	473	0.009946	0.1667	1.2924
COTL1	0.805	474	0.008702	0.1604	1.7772

Gene	Signal to noise	Rank	p-value	FDR(BH)	Fold Change
SYNE2	-0.7141	12527	0.01896	0.2042	-1.347
SERTAD3	-0.7119	12526	0.03947	0.2657	-1.7599
NUP43	-0.7118	12525	0.02269	0.2183	-1.3768
LRRC1	-0.7112	12524	0.01927	0.2046	-1.7489
RPL23	-0.7109	12523	0.01927	0.2046	-1.0593
NBPF10	-0.7102	12522	0.02859	0.2378	-1.6101
CTBP1	-0.7099	12521	0.01927	0.2046	-1.1544
RASSF1	-0.7097	12520	0.02051	0.2104	-1.4203
CTSC	-0.7095	12519	0.03015	0.239	-1.6906
OPHN1	-0.7091	12518	0.02797	0.236	-1.0862
SYNGR1	-0.7089	12517	0.02145	0.214	-1.9961
ELOVL6	-0.7086	12516	0.01896	0.2042	-2.7408
YWHAQ	-0.7083	12515	0.01523	0.1909	-1.1787
ADAM10	-0.7083	12514	0.02145	0.214	-1.4295
UTP14C	-0.708	12513	0.0174	0.1984	-1.3013
PCOLCE	-0.7079	12512	0.02269	0.2183	-1.3419
RECK	-0.7077	12511	0.01523	0.1909	-1.1808
CDC42EP3	-0.7074	12510	0.01927	0.2046	-1.61
PNISR	-0.7072	12509	0.01989	0.2082	-1.2933
SEC24D	-0.7069	12508	0.02331	0.2215	-1.2578
CA11	-0.7067	12507	0.007148	0.1493	-1.5954
PLAC8	-0.7063	12506	0.01678	0.1965	-2.0045
RNF38	-0.7059	12505	0.01554	0.192	-1.2152
ZNF253	-0.7057	12504	0.01274	0.1814	-1.1913
NEK4	-0.7056	12503	0.02486	0.2266	-1.849
VEZF1	-0.7053	12502	0.01554	0.192	-1.2673
HSP90B1	-0.7051	12501	0.02207	0.2162	-1.2001
TSR1	-0.7049	12500	0.02113	0.2134	-1.1818
COMMD8	-0.7044	12499	0.0202	0.209	-1.4987
PGAP3	-0.704	12498	0.0202	0.209	-1.2637
PTGDR	-0.7038	12497	0.01865	0.2027	-3.8306
DNAJA3	-0.7026	12496	0.006838	0.1475	-1.3327
ARPC5L	-0.7024	12495	0.01057	0.1701	-1.3093
ZMPSTE24	-0.7017	12494	0.0202	0.209	-1.3295
RAB36	-0.7014	12493	0.03263	0.2475	-1.1189
RSBN1	-0.7009	12492	0.01989	0.2082	-1.3414
SLMO2	-0.7008	12491	0.0202	0.209	-1.3904
DENND4A	-0.7007	12490	0.01336	0.1857	-1.3597
ALG6	-0.7005	12489	0.02486	0.2266	-1.3098
ELMO2	-0.6992	12488	0.02238	0.2179	-1.4241
PTGDS	-0.6991	12487	0.02362	0.2232	-2.6557
TRAPPC10	-0.6987	12486	0.02176	0.2155	-1.2534
PPIP5K2	-0.6981	12485	0.0143	0.1887	-1.2087
KIAA0182	-0.6981	12484	0.02176	0.2155	-1.3706
DDX5	-0.6977	12483	0.02393	0.2235	-1.1728
BTN3A1	-0.6968	12482	0.02207	0.2162	-1.3198
FBXL14	-0.6967	12481	0.02207	0.2162	-1.1764
ZBTB38	-0.6967	12480	0.03294	0.2484	-1.367
PRKRIR	-0.6965	12479	0.02113	0.2134	-1.2795
FAM46A	-0.6961	12478	0.02113	0.2134	-1.7883
RPL21	-0.696	12477	0.0174	0.1984	-1.0567
RNF130	-0.6957	12476	0.02113	0.2134	-1.9435
TSHB	-0.6949	12475	0.02486	0.2266	-1.225
FAM224A	-0.6946	12474	0.02766	0.235	-1.0766
ZNF611	-0.6945	12473	0.02176	0.2155	-1.1559
SNX3	-0.6938	12472	0.02113	0.2134	-1.2006
OR7E2P	-0.6936	12471	0.01461	0.1903	-1.6689
C1orf174	-0.6935	12470	0.02082	0.2117	-1.4351
SMAD5	-0.6932	12469	0.01492	0.1909	-1.494
CBFB	-0.6931	12468	0.02207	0.2162	-1.1799
MOB1A	-0.6931	12467	0.01803	0.2007	-1.2546
CSAD	-0.693	12466	0.01772	0.1998	-1.4807
FAM188A	-0.6926	12465	0.0174	0.1984	-1.248
TMEM131	-0.6924	12464	0.02517	0.227	-1.2666
CD302	-0.6918	12463	0.02082	0.2117	-1.4638
SLC25A16	-0.6908	12462	0.02984	0.239	-1.0932
TMEM127	-0.6907	12461	0.02486	0.2266	-1.166
GMPR2	-0.6903	12460	0.009013	0.1634	-1.278
FNDC3B	-0.6902	12459	0.02549	0.2276	-2.0339
TCF7L2	-0.6897	12458	0.02331	0.2215	-1.6084
LY75	-0.6894	12457	0.0202	0.209	-1.9663
NR3C2	-0.6888	12456	0.02797	0.236	-1.791
BAG5	-0.6884	12455	0.02176	0.2155	-1.3442
ABR	-0.6879	12454	0.01709	0.198	-1.1675
PTH	-0.686	12453	0.0115	0.1753	-1.1073
APLP2	-0.6856	12452	0.02859	0.2378	-1.3711
CHUK	-0.6851	12451	0.02145	0.214	-1.4395
PYROXD1	-0.6846	12450	0.01927	0.2046	-1.4471
ATP11B	-0.6823	12449	0.02331	0.2215	-1.2085

Gene	Signal to noise	Rank	p-value	FDR(BH)	Fold Change
TXNRD3	0.8045	475	0.01274	0.1814	1.1823
PTGER1	0.8039	476	0.005594	0.1372	1.107
KLF15	0.803	477	0.01523	0.1909	1.1831
TFAP4	0.8029	478	0.009324	0.1635	1.183
IL18BP	0.8028	479	0.01026	0.1686	1.3631
KANSL3	0.8026	480	0.009324	0.1635	1.3418
UPK3B	0.8025	481	0.009324	0.1635	1.2378
ACRV1	0.8024	482	0.0115	0.1753	1.1843
SEC23B	0.8019	483	0.009013	0.1634	1.3042
ACHE	0.7998	484	0.009946	0.1667	1.2539
ZGPAT	0.7998	485	0.01026	0.1686	1.3313
KL	0.7996	486	0.007148	0.1493	1.1271
STXBP2	0.7996	487	0.01243	0.1803	1.1781
TAPT1	0.7984	488	0.008702	0.1604	1.3684
KHSRP	0.7978	489	0.003419	0.1181	1.2341
HUNK	0.7978	490	0.01336	0.1857	1.2651
YLPM1	0.7961	491	0.01212	0.179	1.228
ENOX1	0.7948	492	0.009013	0.1634	1.2023
NAB1	0.7948	493	0.009635	0.1645	2.3442
SART1	0.7947	494	0.005905	0.1406	1.1888
HIST1H4G	0.7939	495	0.006527	0.1462	1.2163
EYA1	0.7929	496	0.008702	0.1604	1.1658
IL17RA	0.7924	497	0.01181	0.1777	1.3346
RBMY1A1	0.7923	498	0.01057	0.1701	1.1727
CD27	0.7921	499	0.01212	0.179	2.2026

Gene	Signal to noise	Rank	p-value	FDR(BH)	Fold Change
SEC63	-0.682	12448	0.0258	0.2276	-1.1998
VAPA	-0.6811	12447	0.02828	0.237	-1.2326
IFI44	-0.6808	12446	0.02207	0.2162	-2.1266
MAGEF1	-0.6808	12445	0.02207	0.2162	-1.3981
KIR3DX1	-0.6792	12444	0.01678	0.1965	-1.5423
NCKAP1	-0.6789	12443	0.02176	0.2155	-1.8561
NPRL2	-0.6783	12442	0.02922	0.2384	-1.4223
LOC10050582	-0.678	12441	0.01678	0.1965	-1.5942
TCF12	-0.6779	12440	0.02207	0.2162	-1.7687
FXC1	-0.6761	12439	0.02082	0.2117	-1.3101
NAE1	-0.6759	12438	0.023	0.2203	-1.2246
PLEKHG3	-0.6752	12437	0.02859	0.2378	-1.7316
EXTL2	-0.6748	12436	0.02238	0.2179	-1.7575
CXCR2	-0.6735	12435	0.02362	0.2232	-1.7779
GLOD4	-0.6734	12434	0.02455	0.2257	-1.2177
STXBP3	-0.673	12433	0.03419	0.2513	-1.3933
RANBP6	-0.6728	12432	0.02828	0.237	-1.2847
HNRNPK	-0.6725	12431	0.02424	0.2241	-1.0917
NACC2	-0.6724	12430	0.02517	0.227	-1.9837
MRPL44	-0.6724	12429	0.04755	0.2872	-1.1593
TERF1	-0.6723	12428	0.02269	0.2183	-1.2001
G6PC2	-0.6718	12427	0.02517	0.227	-1.2183
SCP2	-0.6711	12426	0.023	0.2203	-1.4285
RCAN2	-0.671	12425	0.01927	0.2046	-2.4783
USP24	-0.6707	12424	0.0258	0.2276	-1.3613

Apppendix Table 9.12.Genes positively and negatively correlated with CD39

Positively correlated with CD39 (ENTPD1)				Negatively correlated with CD39 (ENTPD1)		
Gene	Pearson Correlation	p-value		Gene	Pearson Correlation	p-value
ENTPD1	1	0.001996		AGL	-0.8166	0.001996
HLA-DRA	0.943	0.001996		ATAD2B	-0.8116	0.003992
DIO3	0.9317	0.001996		WSB2	-0.8024	0.001996
DSC2	0.9174	0.001996		LOC100506732	-0.8023	0.001996
TOP2A	0.9003	0.001996		DZANK1	-0.7983	0.003992
CD74	0.8941	0.001996		HERC3	-0.798	0.001996
CTLA4	0.8854	0.001996		CACNA2D2	-0.7898	0.001996
DRG2	0.8814	0.001996		RAB6A	-0.7838	0.001996
MAP2K2	0.8803	0.001996		CHSY1	-0.7751	0.001996
NPM1	0.8793	0.001996		CLCF1	-0.7719	0.001996
AURKA	0.8784	0.001996		ENPP4	-0.7657	0.003992
GYPC	0.8695	0.001996		CA11	-0.7649	0.001996
MIR3917	0.8622	0.001996		ADRB2	-0.7643	0.001996
PRKAR1B	0.8597	0.001996		CD164	-0.7636	0.001996
MRPL12	0.8589	0.001996		RPS27A	-0.7613	0.001996
COTL1	0.858	0.001996		RPS6	-0.7547	0.001996
OTUB1	0.8559	0.001996		KIAA0494	-0.7528	0.001996
CDT1	0.8471	0.001996		RNF5	-0.7525	0.001996
CDH7	0.8459	0.001996		TTC38	-0.7519	0.001996
GSTA3	0.8415	0.001996		MST4	-0.7516	0.001996
PDCD1	0.8368	0.003992		SEC63	-0.7482	0.005988
CYP1A1	0.8342	0.001996		RPL23	-0.7417	0.00998
NOS2	0.8341	0.001996		ADAM10	-0.7396	0.003992
RALY	0.8339	0.001996		SNX24	-0.7395	0.001996
CC2D2B	0.8309	0.001996		QKI	-0.7377	0.003992
ASPM	0.8281	0.001996		KIAA0182	-0.7376	0.003992
C6orf108	0.8278	0.001996		GALNT10	-0.7374	0.005988
BUB1	0.8268	0.001996		ABR	-0.7352	0.001996
MGAT4B	0.8246	0.001996		RFC5	-0.735	0.001996
CWH43	0.8246	0.003992		CELF1	-0.733	0.001996
PDZRN3	0.8235	0.001996		PTP4A2	-0.7298	0.001996
BUB1B	0.8219	0.001996		RPL30	-0.7242	0.003992
TLR4	0.8208	0.001996		TARP	-0.7235	0.001996
PSMD8	0.8181	0.001996		PLEKHA1	-0.7233	0.005988
ACP2	0.8166	0.001996		BMI1	-0.7228	0.001996
SLC6A14	0.8148	0.001996		DHRS3	-0.7225	0.001996
CORT	0.8142	0.003992		GOLGA8IP	-0.7212	0.003992
CDC25A	0.8132	0.003992		ZNF562	-0.7197	0.003992
NUSAP1	0.8121	0.001996		PLEKHA5	-0.7168	0.007984
CENPE	0.8116	0.001996		SLC9A3R1	-0.7153	0.001996
MYO9B	0.8089	0.003992		PRX	-0.7151	0.001996
CCDC64	0.8083	0.001996		NBPF10	-0.7144	0.005988
TRIB1	0.806	0.001996		EEF1B2	-0.7126	0.001996
TCEB3	0.8031	0.003992		PIGF	-0.7126	0.003992
GPR162	0.8016	0.001996		ZNF506	-0.7124	0.005988
BIRC5	0.8002	0.001996		ZNF253	-0.7115	0.001996
MYCNOS	0.7999	0.001996		ENO2	-0.7107	0.001996
RANBP9	0.7997	0.003992		HSP90B1	-0.7068	0.007984
PSMA1	0.7984	0.003992		NPTN	-0.7019	0.003992
OR2B2	0.7968	0.005988		PPIH	-0.7014	0.007984
RBBP8	0.7963	0.001996		TDRD7	-0.7008	0.001996
PPP1R7	0.7963	0.001996		RNF130	-0.7006	0.001996
ADAM5P	0.7958	0.001996		MXI1	-0.7	0.003992
INPP5K	0.7938	0.001996		NOL10	-0.6942	0.005988
PLK4	0.7928	0.001996		GOLGA6L3	-0.6935	0.001996
INTS1	0.792	0.001996		CAPN7	-0.6934	0.01198
NBL1	0.7919	0.003992		PURA	-0.691	0.005988
POLR1E	0.7918	0.001996		FAM3C	-0.6908	0.001996
SPAG6	0.7918	0.001996		GATAD2A	-0.6907	0.005988
RHBDD3	0.7917	0.001996		LOC100506390	-0.689	0.005988
ORC1	0.7917	0.001996		XBP1	-0.6888	0.007984
RAPGEF1	0.7906	0.001996		DSTN	-0.6877	0.007984
AKAP5	0.7844	0.001996		SRP9	-0.687	0.005988
PPP4C	0.784	0.001996		RPL27	-0.686	0.001996
ASXL1	0.784	0.005988		TMED10	-0.6855	0.005988
SSTR1	0.7838	0.003992		SIDT1	-0.684	0.01198
HOXC13	0.7838	0.001996		PSIP1	-0.6834	0.003992
PBK	0.7836	0.001996		ERBB2IP	-0.6825	0.001996
MYH6	0.7835	0.003992		GNPTAB	-0.6821	0.00998
HRH1	0.7833	0.001996		RPS14	-0.6801	0.005988
FAM155B	0.7829	0.003992		MAF	-0.6786	0.003992
YWHAE	0.7827	0.001996		UFM1	-0.678	0.003992
C1orf151-NBL1	0.7817	0.003992		SYNE2	-0.6772	0.003992
HSPB6	0.7809	0.001996		KLRD1	-0.677	0.003992
BRD4	0.7794	0.001996		RAB11FIP5	-0.6758	0.001996
DZIP1	0.7788	0.001996		GBE1	-0.6755	0.001996
LMCD1	0.7782	0.007984		GNLY	-0.6745	0.003992
BCAT1	0.7774	0.001996		WWP2	-0.6729	0.01198

Positively correlated with CD39 (ENTPD1)				Negatively correlated with CD39 (ENTPD1)		
Gene	Pearson Correlation	p-value		Gene	Pearson Correlation	p-value
LOC400084	0.7773	0.003992		MICB	-0.6728	0.00998
VCAM1	0.7759	0.005988		RHOQ	-0.6718	0.003992
LANCL2	0.7752	0.003992		TNF	-0.6709	0.003992
CRISP3	0.7747	0.001996		STOM	-0.6708	0.003992
GJB5	0.7735	0.001996		LACTB2	-0.6687	0.005988
IL25	0.7733	0.001996		ZMYM1	-0.6685	0.005988
GARS	0.7728	0.001996		NHLRC2	-0.6683	0.003992
HLA-DOA	0.7714	0.007984		OTUD3	-0.6681	0.005988
MYO1E	0.7713	0.005988		RCBTB2	-0.668	0.005988
JAG1	0.7708	0.001996		KLRG1	-0.6679	0.001996
CDC45	0.7698	0.003992		ELF4	-0.6672	0.003992
MAB21L1	0.7697	0.001996		RCOR3	-0.6653	0.005988
CENPO	0.7691	0.001996		REEP4	-0.6647	0.007984
ZNF335	0.7691	0.007984		SEC22B	-0.6644	0.007984
C3orf14	0.7683	0.003992		CDC42EP3	-0.6629	0.005988
WBSCR22	0.768	0.005988		TSR1	-0.6611	0.007984
SLC15A3	0.7676	0.005988		GLUL	-0.66	0.003992
CDKN2A	0.7674	0.003992		DPY19L1	-0.6595	0.001996
ADCY3	0.7669	0.001996		ATF2	-0.6592	0.01397
HLA-DRB1	0.7665	0.001996		GCA	-0.6579	0.01198
CLN6	0.7661	0.001996		TDRKH	-0.6564	0.00998
AKT2	0.7657	0.001996		NUCKS1	-0.6556	0.005988
CHEK1	0.7655	0.001996		KLRC3	-0.6553	0.003992
ZGPAT	0.7645	0.001996		CERS2	-0.6552	0.01796
ASF1B	0.7622	0.005988		METTL2A	-0.6542	0.01397
PNMA2	0.7613	0.003992		TOMM70A	-0.6542	0.005988
SMARCE1	0.761	0.005988		GANAB	-0.6523	0.01597
MIS18BP1	0.7609	0.003992		VPS13C	-0.6515	0.005988
FOLH1B	0.76	0.003992		UTP14C	-0.651	0.00998
SOD1	0.7597	0.001996		ACO1	-0.6508	0.005988
YKT6	0.7593	0.001996		RPL21	-0.6503	0.001996
GBP2	0.7583	0.001996		SLCO4C1	-0.6503	0.001996
NAB1	0.7579	0.001996		B2M	-0.6502	0.01397
PROZ	0.7571	0.01397		LOC100505828	-0.6496	0.005988
EXO1	0.7567	0.003992		SSBP2	-0.6494	0.003992
DCLK1	0.7563	0.005988		DHRS7	-0.6491	0.001996
CAMK4	0.7561	0.003992		C6orf62	-0.649	0.01597
MRPS7	0.7556	0.005988		TAF7	-0.649	0.005988
TNFRSF9	0.7553	0.005988		CD320	-0.6489	0.005988
FIP1L1	0.7544	0.007984		CFH	-0.6488	0.007984
LIMD2	0.7543	0.001996		MEAF6	-0.6482	0.007984
PRCC	0.7539	0.003992		TMEM59	-0.6482	0.00998
DTL	0.7539	0.001996		TANC2	-0.6478	0.001996
TACC3	0.7538	0.005988		MGA	-0.6472	0.005988
MAGI2	0.7533	0.001996		RNF220	-0.6462	0.01397
HDLBP	0.753	0.005988		ARCN1	-0.6444	0.003992
LGALS8	0.7525	0.00998		SEH1L	-0.6428	0.01397
S100G	0.7516	0.003992		TPP1	-0.6428	0.001996
PKMYT1	0.7515	0.00998		PDE4DIP	-0.6424	0.003992
ZNRF4	0.751	0.005988		DMXL2	-0.6424	0.01597
GAGE12B	0.7509	0.003992		RNF44	-0.6414	0.00998
HLA-DMA	0.7508	0.001996		CALR	-0.6414	0.007984
CD79A	0.7508	0.00998		PPP3CA	-0.6412	0.005988
SCGB2A1	0.7494	0.001996		ZNF395	-0.6406	0.005988
ACOT1	0.7491	0.001996		UBE2E3	-0.6406	0.007984
ETV5	0.749	0.005988		UBE2W	-0.6396	0.007984
INPPL1	0.7488	0.003992		BAG5	-0.6395	0.007984
NGFRAP1	0.7486	0.003992		CXorf57	-0.639	0.01397
ZWINT	0.7476	0.001996		LECT2	-0.6389	0.03393
DUSP4	0.7476	0.003992		TCF7L2	-0.6388	0.007984
ZEB1	0.7472	0.003992		CAMK2B	-0.6386	0.001996
OBSL1	0.7472	0.003992		LAIR1	-0.6386	0.01397
SFTPD	0.7471	0.001996		UBA3	-0.6377	0.01796
PAX3	0.7463	0.007984		ARFGEF2	-0.6368	0.01397
MIR4800	0.746	0.01198		MBNL1	-0.6367	0.007984
HIST1H1B	0.7459	0.001996		THAP11	-0.6361	0.00998
SOX14	0.7457	0.001996		TGOLN2	-0.6357	0.00998
GMIP	0.7451	0.001996		UBP1	-0.6336	0.007984
GYG2	0.745	0.001996		ARF6	-0.633	0.007984
PRODH2	0.7448	0.003992		KLRAP1	-0.633	0.003992
CCNB2	0.7448	0.003992		SESN1	-0.6315	0.001996
CSPG5	0.7431	0.005988		GTPBP3	-0.6311	0.01198
NIPAL3	0.7428	0.005988		ZNF137P	-0.63	0.003992
SLC27A2	0.7426	0.005988		UBQLN2	-0.6298	0.003992
PQLC1	0.7424	0.005988		PTPN12	-0.6297	0.001996
FAHD2A	0.7419	0.003992		SLC4A4	-0.6295	0.001996
E2F1	0.7411	0.005988		ADCY7	-0.6292	0.01397
PRODH	0.7407	0.001996		SLK	-0.629	0.007984

Positively correlated with CD39 (ENTPD1)				Negatively correlated with CD39 (ENTPD1)		
Gene	Pearson Correlation	p-value		Gene	Pearson Correlation	p-value
AHI1	0.7402	0.007984		RGS3	-0.6288	0.005988
HMGB3	0.7397	0.007984		DFFB	-0.6286	0.02994
PSORS1C1	0.7396	0.007984		FXC1	-0.6286	0.003992
DNAJB12	0.7394	0.001996		NCOA3	-0.6285	0.00998
FBRS	0.7394	0.001996		AP3S2	-0.6282	0.01597
DSCR4	0.7391	0.005988		GTF2H2	-0.6282	0.005988
SPOCK3	0.7388	0.007984		CELSR2	-0.6276	0.007984
CDC42BPA	0.7386	0.005988		PFDN5	-0.6276	0.01198
MAD1L1	0.7383	0.007984		ADAM8	-0.6268	0.01597
ZDHHC3	0.7381	0.00998		DNAJB9	-0.6266	0.01397
TYMS	0.7377	0.005988		FAM117A	-0.6262	0.01796
NRBP1	0.7373	0.007984		USP33	-0.6259	0.02395
SHMT2	0.7371	0.001996		KIF1B	-0.6257	0.007984
MRPS15	0.7369	0.007984		ISOC1	-0.6255	0.01597
GTF2F1	0.7368	0.005988		WNT10B	-0.6255	0.01397
CNTFR	0.7367	0.003992		APBA1	-0.6251	0.007984
HSF1	0.7351	0.003992		RPL23A	-0.625	0.00998
NAPA	0.7343	0.007984		PPP1R11	-0.6248	0.01397
KCNB1	0.7339	0.01198		CUL4A	-0.6244	0.001996
HLA-DMB	0.7336	0.001996		CRBN	-0.6243	0.01397
TBXA2R	0.7332	0.00998		TRAPPC10	-0.6241	0.01796
MUC7	0.733	0.001996		DPY19L4	-0.6239	0.01996
ORAI2	0.733	0.00998		IGF2R	-0.6234	0.001996
SH3GL1	0.7328	0.005988		NEO1	-0.6232	0.007984
NRXN3	0.7327	0.005988		FLJ11292	-0.6228	0.01796
C17orf80	0.7317	0.007984		YTHDF1	-0.6225	0.001996
MSX1	0.7316	0.007984		DYNLT3	-0.6211	0.01397
WDR25	0.7315	0.005988		FBXL5	-0.621	0.01198
HSD11B2	0.7315	0.005988		RPL7	-0.6208	0.00998
SMTN	0.7309	0.001996		SLC20A1	-0.6207	0.00998
SYNJ2	0.7305	0.001996		C14orf169	-0.6203	0.01198
ICOS	0.7305	0.005988		S1PR5	-0.62	0.007984
NDN	0.7303	0.01597		SPINLW1	-0.6197	0.003992
LRRC2	0.7303	0.00998		HIPK3	-0.6183	0.01996
FABP4	0.7302	0.007984		GPR56	-0.6183	0.001996
RAB3B	0.7301	0.001996		SRSF11	-0.6182	0.005988
PHB2	0.7291	0.007984		PANK3	-0.6177	0.005988
RCAN3	0.7289	0.001996		SMAD5	-0.6175	0.007984
LRRN3	0.7286	0.003992		COPZ2	-0.6173	0.005988
PITX3	0.7276	0.001996		PTCH1	-0.6172	0.001996
OR3A1	0.7275	0.001996		NAE1	-0.6169	0.005988
PTPRK	0.7272	0.003992		C8orf44-SGK3	-0.6168	0.005988
SBNO2	0.727	0.005988		C4orf46	-0.6162	0.00998
EPCAM	0.7267	0.003992		OSTM1	-0.6156	0.005988

Appendix Figure 9.1. Consent form for HCV samples

Partners HealthCare System
Research Consent Form

Subject Identification

General Template
Version Date: November 2005

Protocol Title: Cell Mediated Immunity in Hepatitis C Virus Infection

Principal Investigator: Georg M. Lauer, M.D.

Site Principal Investigator: Athe Tsibris,M.D.

Description of Subject Population: Hepatitis C+ or Healthy subjects

About this consent form

Please read this form carefully. It tells you important information about a research study. A member of our research team will also talk to you about taking part in this research study. People who agree to take part in research studies are called "subjects." This term will be used throughout this consent form. If you have any questions about the research or about this form, please ask us. If you decide to take part in this research study, you must sign this form to show that you want to take part. We will give you a copy of this form to keep.

Some of the people who are eligible to take part in this study may not be able to give consent because they are less than 18 years of age (a minor). Instead we will ask their parent(s) to give permission for them to take part in the study and will ask them to agree (give their assent) to take part. Throughout the consent form, "you" always refers to the person who takes part in the study.

Why is this research study being done?

The purpose of this research study is to examine the immune response during infection with the hepatitis C virus. (HCV) You have been asked to be in a research study because you have been exposed to the hepatitis C virus, or you are co-infected with hepatitis C and HIV. Or you have not been exposed to hepatitis C or HIV and will serve as healthy volunteers. You are between the ages of 18-75 or 14-18. We expect about 1200 people will take part in this study.

How long will I take part in this research study?

Depending on your immune response to hepatitis C (HCV), you may be asked to return to provide additional samples of blood. If so, you will be contacted by the study doctor or research staff and asked if you will be willing to donate another blood sample. You are free to refuse this request. Your participation is entirely voluntary and you can withdraw at any time.

Consent Form Title: 4.14.14 CellMed Consent CLEAN
IRB Protocol No: 1999P004983 **Sponsor Protocol No: AI066345; AI067926; AI76706;**
Consent Form Valid Date: 9/8/2014 **IRB Amendment No: N/A** **Sponsor Amendment No: N/A**
IRB Expiration Date: 9/25/2015 **IRB Amendment Approval Date: N/A**

The interval between additional blood samples would depend on the course of the HCV infection, for example every other week during acute HCV infection. If you have chronic Hepatitis C, intervals between blood draws would typically be between twelve and twenty-four weeks. These intervals may be shorter under special circumstances, for example a change in disease activity or scheduling of a liver biopsy. Usually the duration of the follow up will be no longer than 2 years, although you could be asked to continue in the study even longer. Examples are the rare patients displaying very unique or strong immune responses against HCV that were associated with a better outcome of the infection.

What will happen in this research study?

You will have your blood taken:

(For women)

You will be asked if you may be pregnant and if you are of childbearing age you will be asked about sexual activity, whether you are seeking to become pregnant, and about use of contraception. Your responses will be documented in your research record and a urine pregnancy test will be done before a blood sample is collected if you think you are pregnant or your responses to the above questions suggest pregnancy is possible.

(For all participants)

A blood sample will be drawn (approximately 60 ml of blood or 12 teaspoons). The skin overlying the large veins of your forearm is carefully cleansed with alcohol. A needle is then inserted into the vein and 60 ml of blood or 12 teaspoons is withdrawn using a special tube or syringe.

The total amount of blood to be drawn annually will not exceed 2 pints and will be in most cases less than ½ pint. A maximum of 550 mL of blood may be drawn in any 8 week period and blood will not drawn more than 4 times in any 8 week period. Your hemoglobin level will be checked to make sure you are not anemic. For subjects ages 14-17, blood volume drawn will always be less than 3cc/kg of body weight per 8 week period.

Subjects whose samples are collected when they are minors will be re-consented when they turn 18, and if those individuals cannot be located their samples will be anonymized.

We will put a number code on the tubes of blood and the following laboratory tests may be done on your blood sample:

1. CD 4 and CD8 responses: Measure how your cells provide help to your immune system
2. Cytotoxic T cell responses and NK cell responses: Measure the ability of your blood cells to kill virus bearing cells
3. HIV and/or HCV viral load: how much virus is in your blood

4. HCV sequencing, analysis of the HCV viral genome: This helps figure out the type of HCV you have
5. Genetic testing: to find out what genes are special to you and if these genes help you to control HCV better than other people.
6. CBC - complete blood count
7. Liver function tests to tell us how healthy your liver is
8. RNA transcription profiling: This is a test that will show whether certain genes are turned on or off in your blood cells

A. We will ask to do genetic tests on your blood. We will test your blood for a number of genes you have inherited from your parents, that we believe to be important in your immune response against viruses and bacteria. We may ask other laboratories outside of Partners to do this test for us. All blood and information will be coded with a number and your name will not be known. This test is for research purposes only and no individual results will be given back to you. These results will not go into your medical record. We will not be able to get information that could be clinically helpful to any care you may receive from your primary care physician.

I agree to have genetic tests done on my blood.

_____YES _____NO _____YOUR INITIALS

B. We may contact you and ask you to give more blood samples at a later time. This will depend on the course of your HCV infection. For example, if you have acute HCV infection, we may ask for a blood sample every 1-2 weeks for the first two months and then once a month thereafter. If you have chronic HCV infection, we will usually collect blood samples every 3 to 6 months. We may collect blood samples more often under certain circumstances, for example if you have a change in disease activity. We will not draw your blood more often than twice a week. We will not ask for more than 550 ml (less than 2 ½ cups) in an eight week period. You do not have to give more blood; you can say no any time we ask you.

Would you be willing to be asked to donate more blood in the future?

_____YES _____NO _____YOUR INITIALS

C. Liver tissue collection:
If during your routine clinical care, a liver biopsy is done by your physician, (a separate informed consent for that procedure will be obtained), we request a small part of that tissue for analysis. Tissue will only be requested if there are sufficient amounts beyond what is needed for the diagnostic tests. Additional blood samples may be requested the day of your liver biopsy. You are not obligated to give blood, and may say no at anytime.

Partners HealthCare System
Research Consent Form

Subject Identification

General Template
Version Date: November 2005

If a liver biopsy is done by your physician, do you agree to have a part of this tissue given to us for analysis as explained above?

_____YES _____NO _____YOUR INITIALS

D. We will take medical information from your medical record and share it with researchers associated with this study. This information will be coded with a number instead of your name in order to protect your privacy. Information collected as part of this study may include: lab tests, medications, history of infections.

Do you agree to have medical information taken from your record. This is all coded with a number

_____YES _____NO _____YOUR INITIALS

E. We will ask you to let us store/save your samples and make immortalized cell lines from your blood.
Samples of blood or liver tissue may be stored (frozen) in our laboratory for future research to learn about Hepatitis C. We would like to make immortalized cell lines, which means that cells from your blood sample will be made to grow in the laboratory many times or we will save them frozen for growing later. This gives us a source of your cells for research for many years. Your samples/data will be used only for research and will not be sold or used directly for the production of commercial products. If at a later time you decide that you want these cell lines destroyed you have the right to do so.

Your samples/data will be coded with a number so that your name cannot be readily identified. Reports about research done with your samples will not be put in your medical record and will be kept confidential to the best of our ability within state and federal law.

I give my permission to store/save my sample and make immortalized cell lines to be used in future research as described above.

_____YES _____NO _____YOUR INITIALS

Consent Form Title: 4.14.14 CellMed Consent CLEAN
IRB Protocol No: 1999P004983 Sponsor Protocol No: AI066345; AI067926; AI76706;
Consent Form Valid Date: 9/8/2014 IRB Amendment No: N/A Sponsor Amendment No: N/A
IRB Expiration Date: 9/25/2015 IRB Amendment Approval Date: N/A

F. We may send out some of your samples or information such as laboratory results to researchers outside of Partners who are also studying HCV infection and we will also send information collected from this study to the Los Alamos Database, however, a code instead of your name will always be used, and the outside researchers will not have access to the link between this code and your name. However, the principle investigator and clinical research coordinators at Partners are able to identify your samples through a confidential patient list should you wish to have them withdrawn from research in the future. Results from future research using your samples may be presented in publications and meetings, but your name will not be identified.

I give permission to send some of my samples or information outside of Partners.

_____ YES _____NO _____YOUR INITIALS

You can change your mind at any time about allowing your samples to be used for future research. If you do, contact the study doctor in writing and let them know. Then your samples will no longer be made available for research and will be destroyed.

What are the risks and possible discomforts from being in this research study?

Sampling blood may cause a small amount of pain, bleeding or temporary bruising. Rarely a person feels lightheaded or faint when their blood is drawn. In extremely rare instances an infection may develop at the needle insertion site.

Genetic information that results from this study does not have medical or treatment importance at this time. We will not report any genetic test results to you or your doctor. There is a risk that information about taking part in a genetic study may influence insurance companies or employers regarding your health. If you do not share information about taking part in this study, you will reduce this risk. We will not place information about the study or the results of genetic tests in your medical record.

What are the possible benefits from being in this research study?

Although there are no direct benefits to participating in this study, it is hoped that you and others may benefit if we can learn more about the immune responses to the hepatitis C virus.

Partners HealthCare System
Research Consent Form

Subject Identification

General Template
Version Date: November 2005

Can I still get medical care within Partners if I don't take part in this research study, or if I stop taking part?

Yes. Your decision won't change the medical care you get within Partners now or in the future. There will be no penalty, and you won't lose any benefits you receive now or have a right to receive.

Taking part in this research study is up to you. You can decide not to take part. If you decide to take part now, you can change your mind and drop out later. We will tell you if we learn new information that could make you change your mind about taking part in this research study.

If you take part in this research study, and want to drop out, you should tell us. We will make sure that you stop the study safely. We will also talk to you about follow-up care, if needed.

It is possible that we will have to ask you to drop out before you finish the study. If this happens, we will tell you why. We will also help arrange other care for you, if needed.

Will I be paid to take part in this research study?

You will be paid $10.00 each time you have your blood drawn, and if you parked in the MGH or BWH garages you will be given a parking voucher.

To be able to pay you, we will need to ask you for your name, address, and social security number. For tax reporting purposes, we will need to report this information with the payment amount to MGH Accounts Payable Department.

What will I have to pay for if I take part in this research study?

Hospitalization will not be necessary. Neither your insurance company nor you will be charged for the cost of blood tests involved with this study. The cost of your routine medical care will be billed to your or to your insurance company in the usual way.

What happens if I am injured as a result of taking part in this research study?

We will offer you the care needed to treat any injury that directly results from taking part in this research study. We reserve the right to bill your insurance company or other third parties, if appropriate, for the care you get for the injury. We will try to have these costs paid for, but you may be responsible for some of them.

Partners HealthCare System
Research Consent Form

Subject Identification

General Template
Version Date: November 2005

Giving you care does not mean that Partners hospitals or researchers are at fault, or that there was any wrongdoing. There are no plans for Partners to pay you or give you other compensation for the injury. However, you are not giving up any of your legal rights by signing this form.

If you think you have been injured or have experienced a medical problem as a result of taking part in this research study, tell the person in charge of this study as soon as possible. The researcher's name and phone number are listed in the next section of this consent form.

If I have questions or concerns about this research study, whom can I call?

You can call us with your questions or concerns. Our telephone numbers are listed below. Ask questions as often as you want.

Georg Lauer, MD is the person in charge of this research study. For questions about the study, you can call him at 617-724-7515. You can also call Dr. Raymond Chung at 617-724-6006, Dr. James Morrill at 617-726-2000, pager 38152, Dr. Athe Tsibris at (617) 768-8352 or Dr. Arthur Kim at 617-724-7514. You can call them Monday-Friday from 9:00 am to 5:00 pm or by the page operator at 617-726-2241.

Stephanie Kulaga and Joelle Brown are the research assistants. Stephanie can be reached at 617-726-1475 and Joelle can be reached at 617-643-7754.

If you want to speak with someone **not** directly involved in this research study, please contact the Partners Human Research Committee office. You can call them at 617-424-4100.

You can talk to them about:

- Your rights as a research subject
- Your concerns about the research
- A complaint about the research

Also, if you feel pressured to take part in this research study, or to continue with it, they want to know and can help.

If I take part in this research study, how will you protect my privacy?

Federal law requires Partners (Partners HealthCare System and its hospitals, health care providers and researchers) to protect the privacy of health information that identifies you. This

information is called Protected Health Information. In the rest of this section, we refer to this simply as "health information."

If you decide to take part in this research study, your health information may be used within Partners and may be shared with others outside of Partners, as explained below.

We have marked with a ☒ how we plan to use and share your health information. If a box is not checked ☐, it means that type of use or sharing is not planned for in this research study.

We will also give you the **Partners Notice for Use and Sharing of Protected Health Information.** The Notice gives more details about how we use and share your health information.

- **Health Information About You That Might be Used or Shared During This Research**
  - ☒ Information from your hospital or office health records within Partners or elsewhere, that may be reasonably related to the conduct and oversight of the research study. If health information is needed from your doctors or hospitals outside Partners, you will be asked to give permission for these records to be sent to researchers within Partners.
  - ☒ New health information from tests, procedures, visits, interviews, or forms filled out as part of this research study

- **Why Health Information About You Might be Used or Shared with Others**

  The reasons we might use or share your health information are:
  - To do the research described above
  - To make sure we do the research according to certain standards - standards set by ethics and law, and by quality groups
  - For public health and safety - for example, if we learn new health information that could mean harm to you or others, we may need to report this to a public health or a public safety authority
  - For treatment, payment, or health care operations

- **People and Groups That May Use or Share Your Health Information**

  1. **People or groups within Partners**
     - ☒ Researchers and the staff involved in this research study

☒ The Partners review board that oversees the research

☒ Staff within Partners who need the information to do their jobs (such as billing, or for overseeing quality of care or research)

2. **People or groups outside Partners**

☒ People or groups that we hire to do certain work for us, such as data storage companies, our insurers, or our lawyers

☒ Federal and state agencies (such as the U.S. Department of Health and Human Services, the Food and Drug Administration, the National Institutes of Health, and/or the Office for Human Research Protections) and other U.S. or foreign government bodies, if required by law or involved in overseeing the research

☒ Organizations that make sure hospital standards are met

☒ The sponsor(s) of the research study, and people or groups it hires to help perform this research study

☒ Other researchers and medical centers that are part of this research study

☐ A group that oversees the data (study information) and safety of this research study

☐ Other:

Some people or groups who get your health information might not have to follow the same privacy rules that we follow. We share your health information only when we must, and we ask anyone who receives it from us to protect your privacy. However, once your information is shared outside Partners, we cannot promise that it will remain private.

- **Time Period During Which Your Health Information Might be Used or Shared With Others**
  - Because research is an ongoing process, we cannot give you an exact date when we will either destroy or stop using or sharing your health information.

- **Your Privacy Rights**
  - You have the right **not** to sign this form permitting us to use and share your health information for research. If you don't sign this form, you can't take part in this research study. This is because we need to use the health information of everyone who takes part in this research study.

- You have the right to withdraw your permission for us to use or share your health information for this research study. If you want to withdraw your permission, you must notify the person in charge of this research study in writing.

  If you withdraw your permission, we will not be able to take back information that has already been used or shared with others. This includes information used or shared to carry out the research study or to be sure the research is safe and of high quality.

  If you withdraw your permission, you cannot continue to take part in this research study.
- You have the right to see and get a copy of your health information that is used or shared for treatment or for payment. To ask for this information, please contact the person in charge of this research study.

- **If Research Results Are Published or Used to Teach Others**

  The results of this research study may be published in a medical book or journal, or used to teach others. However, your name or other identifying information **will not** be used for these purposes without your specific permission.

Consent/Assent to take part in this research study, and authorization to use or share your health information for research

Statement of Subject or Person Giving Consent/Assent

- I have read this consent form.
- This research study has been explained to me, including risks and possible benefits (if any), other options for treatments or procedures, and other important things about the study.
- I have had the opportunity to ask questions.

If you understand the information we have given you, and would like to take part in this research study, and also agree to allow your health information to be used and shared as described above, then please sign below:

Signature of Subject:

_____ _____

Partners HealthCare System
Research Consent Form

Subject Identification

General Template
Version Date: November 2005

Adults or Minors, ages 14-17 Date/Time

OR

If you understand the information we have given you, and would like to give your permission for your child/the person you are authorized to represent to take part in this research study, and also agree to allow his/her health information to be used and shared as described above, then please sign below:

Signature of Parent(s)/Guardian or Authorized Representative:

Parent(s)/Guardian of Minor Date/Time

OR

Court-appointed Guardian or Health Care Proxy Date/Time

OR

Family Member/Next-of-Kin Date/Time

Relationship to Subject:

Signature of a Witness:

Witness (when required by the PHRC or sponsor) Date/Time

Partners HealthCare System
Research Consent Form

Subject Identification

General Template
Version Date: November 2005

Statement of Study Doctor or Person Obtaining Consent

- I have explained the research to the study subject, and
- I have answered all questions about this research study to the best of my ability.

Study Doctor or Person Obtaining Consent _____ Date/Time _____

In certain situations, the Partners Human Research Committee (PHRC) will require that a subject advocate also be involved in the consent process. The subject advocate is a person who looks out for the interests of the study subject. This person is not directly involved in carrying out the research. By signing below, the subject advocate represents (or "says") that the subject has given meaningful consent to take part in the research study.

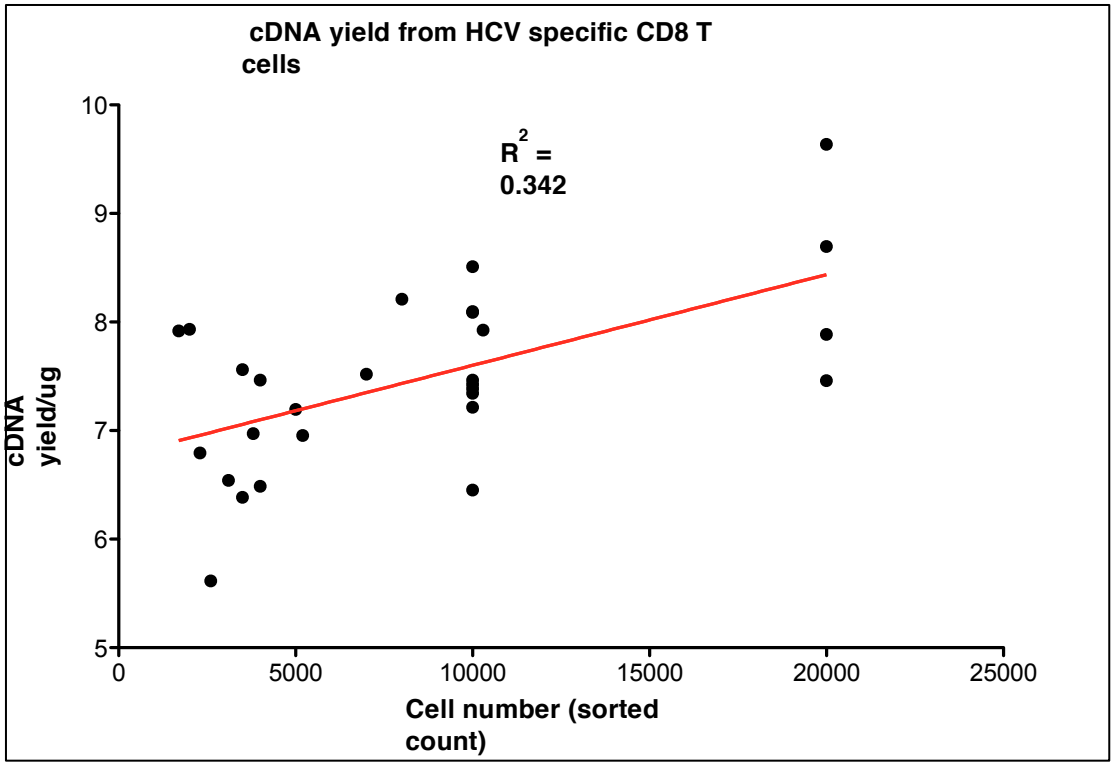
Statement of Subject Advocate Witnessing the Consent Process

- I represent that the subject or authorized individual signing above has given meaningful consent.

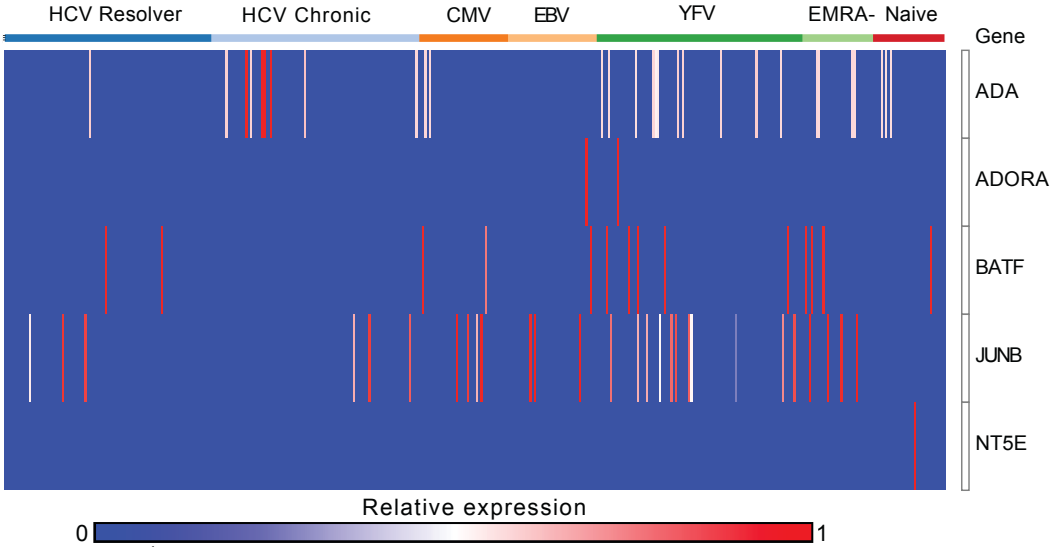
Subject Advocate (when required by the PHRC or sponsor) _____ Date/Time _____

Consent Form Version Date: 20 October 2010

Appendix Figure 9.2. cDNA yields of gene transcription profiling of HCV-specific CD8$^+$ T cells in chapter 3.

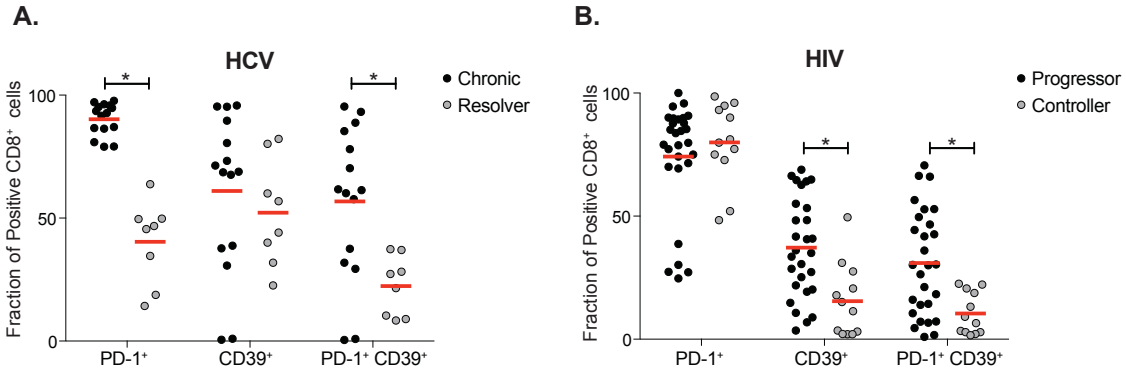


Appendix Figure 9.3. single cell analysis of interesting genes chosen in chapter 3.



Appendix Figure 9.3. Unclustered heat map demonstrating single cell analysis of 5 chosen genes of interest. Only 3.5% of data points were detected making analysis inaccurate.

Appendix Figure 9.4. CD39 and PD-1 expression in HCV, HIV-specific CD8$^+$ T cells



Appendix Figure 9.4. Fraction of HCV (**A**) or HIV (**B**) virus-specific CD8+ T cells expressing PD-1, CD39, or both in patients with chronic disease (black) or patients that resolve virus (grey). Statistical significance was assessed by one-way ANOVA. *P <0.05.

Appendix Figure 9.5. GSEA plot demonstrating cell cycle enrichment in CD39^{+} signature.

A

Table: GSEA Results Summary

Dataset	Samples_QC_OK.CD39_pos_chronic_vs_CD39_neg_chronic.cls #CD39_pos_chronic_versus_CD39_neg_chronic.CD39_pos_chronic_vs_CD39_neg_chronic.cls #CD39_pos_chronic_versus_CD39_neg_chronic_repos
Phenotype	CD39_pos_chronic_vs_CD39_neg_chronic.cls#CD39_pos_chronic_versus_CD39_neg_chronic_repos
Upregulated in class	CD39_pos_chronic
GeneSet	CELL_CYCLE_PHASE
Enrichment Score (ES)	0.46318358
Normalized Enrichment Score (NES)	1.8363216
Nominal p-value	0.0
FDR q-value	0.063467175
FWER p-Value	0.308

B

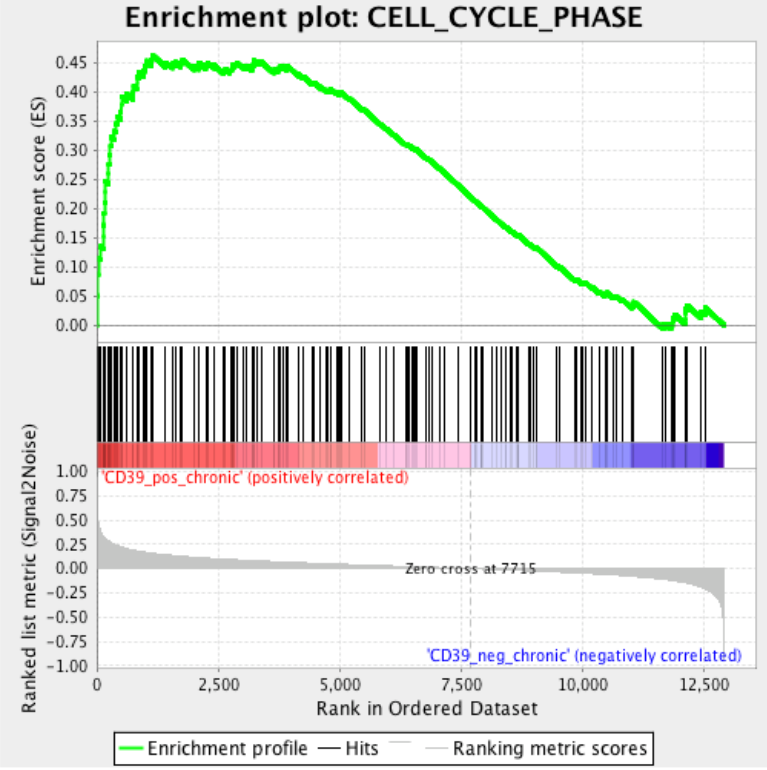


Fig 1: Enrichment plot: CELL_CYCLE_PHASE
Profile of the Running ES Score & Positions of GeneSet Members on the Rank Ordered List

Appendix Figure 9.5 A demonstrating statistics of GSEA analysis. Figure 9.5 B demonstrating enrichment plot of cell cycle signature enriched with CD39+ signature.

Appendix Figure 9.6. List of enriched genes in GSEA in Appendix Figure 9.6

Table: GSEA details [plain text format]

	PROBE	DESCRIPTION (from dataset)	GENE SYMBOL	GENE_TITLE	RANK IN GENE LIST	RANK METRIC SCORE	RUNNING ES	CORE ENRICHMENT
1	BCAT1	226517_at			3	0.833	0.0509	Yes
2	CENPE	205046_at			10	0.606	0.0876	Yes
3	TPX2	210052_s_at			30	0.470	0.1150	Yes
4	BUB1B	203755_at			75	0.386	0.1352	Yes
5	BUB1	216277_at			133	0.335	0.1513	Yes
6	CHEK1	238075_at			135	0.334	0.1718	Yes
7	NCAPH	212949_at			145	0.328	0.1912	Yes
8	KIF15	219306_at			156	0.320	0.2100	Yes
9	ZWINT	204026_s_at			173	0.310	0.2278	Yes
10	CENPF	209172_s_at			175	0.310	0.2468	Yes
11	AURKA	208080_at			241	0.283	0.2590	Yes
12	DLGAP5	203764_at			244	0.283	0.2762	Yes
13	CDKN3	209714_s_at			264	0.276	0.2917	Yes
14	NUSAP1	219978_s_at			280	0.272	0.3072	Yes
15	CDK6	243000_at			297	0.266	0.3223	Yes
16	BIRC5	210334_x_at			364	0.249	0.3324	Yes
17	RAD54L	204558_at			390	0.243	0.3454	Yes
18	POLA1	204835_at			421	0.238	0.3577	Yes
19	KIF11	204444_at			491	0.225	0.3661	Yes
20	SPO11	222259_s_at			502	0.223	0.3789	Yes
21	NDC80	204162_at			505	0.222	0.3924	Yes
22	SYCP1	216917_s_at			616	0.203	0.3963	Yes
23	EREG	205767_at			743	0.189	0.3981	Yes
24	CD28	211861_x_at			755	0.188	0.4088	Yes
25	NEK2	211080_s_at			824	0.182	0.4146	Yes
26	STAG3	219753_at			825	0.182	0.4257	Yes
27	CDKN2A	211156_at			868	0.178	0.4334	Yes
28	SMC4	215623_x_at			967	0.170	0.4361	Yes
29	POLD1	203422_at			991	0.169	0.4447	Yes
30	TTK	204822_at			1016	0.167	0.4530	Yes
31	CDC6	203968_s_at			1127	0.160	0.4542	Yes
32	MSH4	210533_at			1138	0.159	0.4632	Yes