BIOMARKER IDENTIFICATION IN HIV AND NON- HIV RELATED LYMPHOMAS

BY

PUMZA SAMANTHA MAGANGANE

PHLPHU001

Thesis Presented to the Faculty of Health Science of

The University of Cape Town

In Fulfilment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

FEBRUARY 2016

Supervisor: Prof Richard Naidoo

Co-supervisor: Prof Dhiren Govender

Declaration

I, Pumza Samantha Maganganе..., hereby declare that the work on which this thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: …Signed by candidate……

Date: ……20/02/2016…….

DEDICATION

Tothree generations of tenacious women: my grandmother, Angelina Eland;my mother, Juliet Phillips and my sister, Pumeza Phillips.

ACKNOWLEDGEMENTS

I hereby acknowledge my supervisor Professor Richard Naidoo for his guidance, encouragement and support. I also acknowledge Dr RaveendraSookhayi for his assistance from 2012-2014, and following his resignation Professor Dhiren Govender as my co-supervisors.In addition I acknowledge Tony Hu, SubashGovender,AnelisaJaca, Andrew Ndlovu,Raymond Kriel, Emily McGrath, Zainab Mohammed and Charlie Wessels for the various ways they have assisted me.

The laboratory centres and departments listed below played anenormous role in the generation of my results:the Division of Anatomical Pathology at the University of Cape Town (UCT), The International Centre for Genetic Engineering and Biotechnology (ICGEB) (Luiz Zerbini and Durairaj Sekar), Department of Radiation Oncology (UCT), the Proteomics Research and Services Unit at University of the Western Cape (UWC), the Peptide and Catalysis Research Unit at Westville Campus, University of KwaZulu-Natal (UKZN), and finally Receptor Biology Group(UCT).

I will be forever indebted to the Carnegie Corporation and the Department of Science and Technology for funding my undergraduate studies. I am thankful for the funding received from the following organizations:The Medical Research Council (MRC), The National Research Foundation (NRF), The National Health Laboratory Services (NHLS) and the University of Cape Town (UCT).

I am grateful to all my family members for all their prayers, support and encouragement. Finally, I thank the Lord for Hisguidance, and providence.

ABSTRACT

DLBCL is the most common lymphoma subtype occurring in older populations as well as in younger HIV infected patients. The current treatment options for DLBCL are effective for most patients yet the relapse rate is high. While many biomarkers for DLBCL exist, they are not in clinical use due to low sensitivity and specificity. In addition, these biomarkers have not been studied in the HIV context. Therefore, the identification of new biomarkers for HIV negative and HIV positive DLBCL, may lead to a better understanding of the disease pathology and better therapeutic design.

Initially differences in the clinicopathological features between HIV negative and HIV positive DLBCL patients were determined by conducting a retrospective study of patients treated at GSH. Subsequent to this, potential protein biomarkers for DLBCL were determined using MALDI imaging mass spectrometry (IMS) and characterised using LC-MS. The expression of one of the biomarkers, heat shock protein (Hsp) 70, was confirmed on a separate cohort of samples using immunohistochemistry.

Our results indicate that the clinicopathological features for HIV negative and HIV positive DLBCL are similar except for median age, and frequency of elevated LDH levels. Several clinicopathological factors were prognostic for all DLBCL cases including age, gender, stage and bone marrow involvement. In addition, tumour extranodal site was also a prognostic indicator for the HIV negative cohort. The

biomarkers identified in the study consisted of four protein clusters including glycolytic enzymes, ribosomal proteins, histones and collagen. These proteins could differentiate between control and tumour tissue, and the DLBCL subtypes in both cohorts. The majority (41/52) of samples in the confirmation cohort were negative for Hsp70 expression. The HIV positive DLBCL cases had a higher percentage of cases expressing Hsp70 than their HIV negative counterparts. The non-GC subtype also frequently overexpressed Hsp70, confirming MALDI IMS data. Expression of Hsp70 correlated with poor outcome in the HIV negative cohort.

In conclusion, this study identified potential biomarkers for HIV negative and HIV positive DLBCL from both clinical and molecular sources. These may be used as diagnostic and prognostic markers complementary to current clinical management for DLBCL.

ABBREVIATIONS

Regions of Interests	ROIs
Imaging Mass Spectrometry	IMS
α-cyano-4-hydroxycinnamic acid	HCCA
Indium tin oxide	ITO
Principal component analysis	PCA
Room Temperature	RT
Mass Spectrometry	MS
Tandem mass spectrometry	MS/MS
Hematoxyin & eosin	H&E
Hodgkin's lymphoma	HL
non- Hodgkins lymphoma	NHL
Acquired Immune Deficiency Syndrome	AIDS
Minutes	Min
Seconds	Sec
Matrix assisted laser desorption/ionisation	MALDI
Surface enhanced laser desorption/ionisation	SELDI
Time of flight	TOF
Electrospray ionization	ESI
Liquid chromatography	LC
Formalin fixed paraffin embedded	FFPE
Anti retroviral therapy	ART
Cell of origin	COO

Kaposi's Sarcoma	KS
Aids related lymphoma	ARL
Mass-to-charge	m/z
Immunohistochemistry	IHC
Total ion count	TIC
University of Cape Town	UCT
University of Western Cape	UWC
University of KwaZulu-Natal	UKZN
Diffuse large B cell lymphoma	DLBCL
Human Immunodefeciency virus	HIV
Follicular dendritic cells	FDC
Germinal centre B cell subtype	GCB
Activated B cell subtype	ABC
Epstein Barr virus	EBV
Burkitt lymphoma	BL
Working formulation	WF
Revised American European lymphoma	REAL
World health organization	WHO
Not otherwise specified	NOS
Latent membrane protein	LMP
Gene expression profiling	GEP
Highly active antiretroviral therapy	HAART
Post translational modifications	PTMs

Electrospray ionization	ESI
Tuberculosis bacterium	TB
Reactive Node	RN
Glyceraldehyde-3-Phosphate Dehydrogenase	GAPDH
Retrieval of interacting genes/proteins	STRING
Protein analysis through evolutionary relationships	PANTHER
Heat shock protein	Hsp
Trifluoroacetic acid	TFA
Overall survival	OS
Complete response rate	CR
International prognostic index	IPI
Stable isotope labeling by amino acids in cell culture	SILAC
Forkhead box protein P1	FOXP1
B-cell leukemia/lymphoma-2	BCL-2
Cyclophosphamide, doxorubicin, vincristine, prednisolone	CHOP
Rituximab	R
Primary central nervous system lymphoma	PCNSL
2D polyacrylamide gel electrophoresis	2D-PAGE
Fourier transform ion cyclotron resonance	FITCR
Performance status	PS
Lactate dehydrogenase	LDH
Event free survival	EFS
Anaplastic large cell lymphoma	ALCL

Common lymphocyte progenitor	CLP
Surrogate L chains	SLC
B- cell receptor	BCR
Class switch recombination	CSR
Somatic hypermutation	SHM
Constitutive heat shock protein	HSC
Inducable heat shock protein	HSPi
Ubiquitin	Ub
Ubiquitin like	UbL
Histone deacetyltransferases	HDAC
Histone acetyletransferases	HATs
Extranodal	EN
Bone marrow involvement	BMi
Groote Schuur Hospital	GSH

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: LITERATURE REVIEW

1.1 BACKGROUND

1.1.1 B lymphocyte development

There are two types of lymphocytes include T lymphocytes and B lymphocytes. The former is derived from the thymus while the latter is derived from the bone marrow. The development of B lymphocytes begins in the bone marrow followed by maturation in secondary lymphoid tissue (Lebien & Tedder 2008). Development begins with the common lymphocyte progenitor (CLP) differentiating into pre B- cells expressing surrogate L chains (SLC)- a pre B- cell receptor (BCR) (Bankovich et al. 2007). Maturation into BCR occurs through signalling induced by ligand independent oligomerization of two pre-BCR (Bankovich et al. 2007). The differentiation stages occurring in secondary lymphoid tissue are antigen dependent; giving rise to either antibody producing plasma cells or memory B- cells (Lebien & Tedder 2008).

1.1.2 B- cell maturation

Germinal centres develop in lymph nodes in response to antigen binding to the BCR expressed by B lymphocytes (Gatto & Brink 2012). Germinal centres contain the B- cell rich dark zone and the follicular dendritic cells (FDC) and follicular helper T- cells (T_{FH}) cell rich light zone (Figure 1.1). Germinal centres form when naive B- cells, upon antigen binding, are co-activated outside follicles in the T- cell rich zone (MacLennan

1994)(Figure 1.1). The movement of B- cells to the T cell rich zone is mediated by the CCR7 chemokine receptor which is upregulated upon antigen binding (Reif et al. 2002).

.

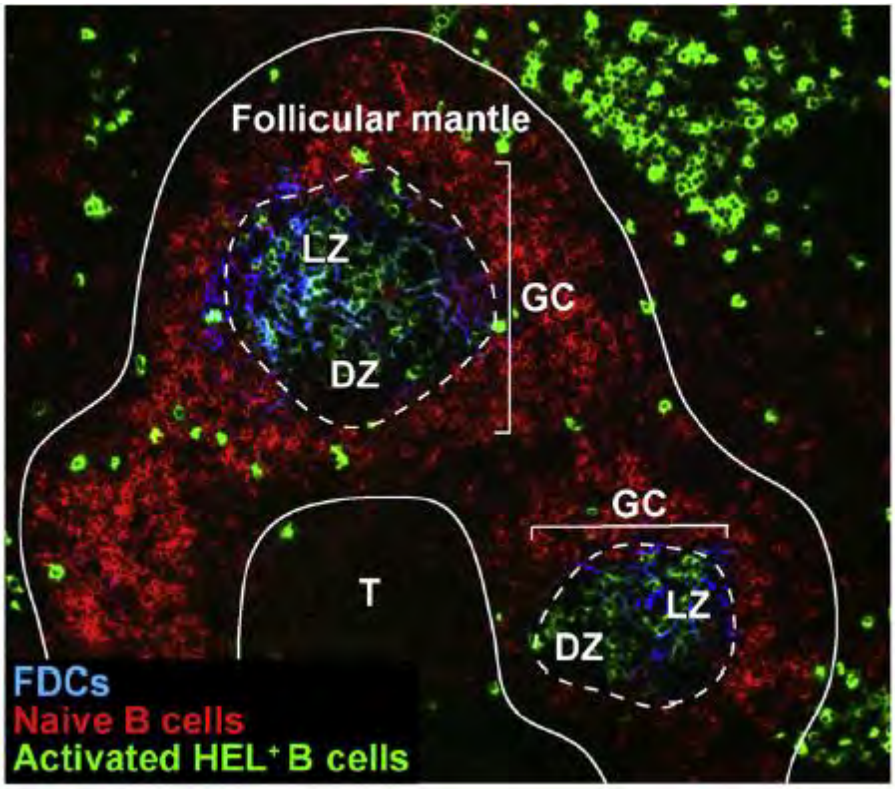


Figure 1.1: The structure of the germinal centre in lymph nodes. The dark zone (DZ) is made up of a high density of B- cells while the light zone (LZ) has fewer numbers with a meshwork of follicular dendritic cells (FDC) (Gatto & Brink 2012).

Following activation by T_{FH} cells, B-cells undergo rapid expansion; somatic hypermutation (SHM) of the IG genes; and class switch recombination (CSR) at the IgH locus in the dark zone (Lebien & Tedder 2008). They are now known as centroblasts.

Upon antigen stimulation, the centroblasts migrate to the light zone and become centrocytes. They are now in activated apoptosis and have high levels of BCL6. They compete for the CD 40L expressed on T_{FH} cells and FDCs in a process called 'selection' enabling them to survive (Vinuesa 2013). A subset of B- cells resume the dark zone process in order to improve antigen selection and this may be mediated by the MYC protein (Karube & Campo 2015). Upon stimulation the centrocytes exit the germinal centre as an antibody producing plasma cell or a memory B- cell.

1.1.3 B- cell abnormalities

Irregularities in B-cell development can give rise to immunodeficiencies, leukaemia, lymphomas and autoimmune diseases (Kuppers et al. 1999). Autoimmune diseases such as rheumatoid arthritis and systemic lupus erythematosus are a result of failure to tolerate self-antigens by B-cells (Hampe 2012). Approximately two thirds of B- cells are self-reactive and are subjected to three distinct tolerance mechanisms, including, clonal deletion, receptor editing and anergy (Tsubata 2005). Malignancies of B- cells arise at different stages of B- cell development resulting in a wide range of morphologic and clinical presentations (Longo 2010). They can present as either leukaemia or lymphoma (Longo 2010).

1.1.4 B-cell lymphomas form at different stages of B- cell development

The majority B- cell non-Hodgkin lymphomas, including Burkitt lymphoma, follicular lymphoma and diffuse large B- cell lymphoma (DLBCL) are derived from GC B- cells (Kuppers et al. 1999; Stevenson & et al 2001). Burkitt lymphomas are derived from the dark zone B- cells (Dave et al. 2006), whereas follicular lymphoma and DLBCLs correspond to B- cells arrested at various stages of GC development (Figure 1.2) (Alizadeh et al. 2000). In particular, follicular lymphoma and the GC B- cell (GCB)-like subtype of DLBCL resemble light zone B- cells, whereas activated B cell (ABC)-like DLBCLs derive from GC cells arrested during the early stages of plasma cell differentiation (Alizadeh et al. 2000).

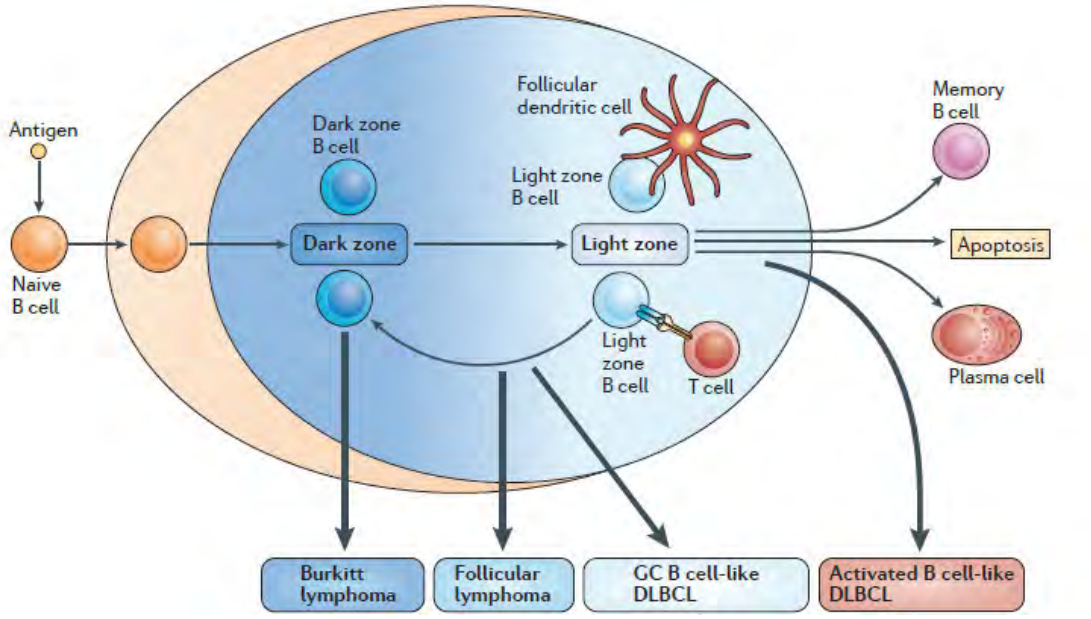


Figure 1.2: The origin of B- cell malignancies. Various types of NHL subtypes arise from B- cells arrested at different stages of B-cell development. Burkitt lymphoma arises from the dark zone while follicular lymphoma and the germinal centre subtype of diffuse large B cell arise from the light zone (Basso & Dalla-favera 2015).

1.2 OVERVIEW

Lymphoma is an uncontrollable growth of lymphocytes in the lymphoid system (Almieda & Barry 2009). There are two types of lymphoma: Hodgkin's lymphoma (HL) and non-Hodgkin's lymphoma (NHL). HL is mainly characterised by the presence of Reed-Sternberg giant cells. NHL on the other hand is composed of a wide spectrum of disease differing in their pathobiology and therapy. The majority of NHLs originate from B-cells, with diffuse large B- cell lymphoma (DLBCL) making upto 40% (Contran et al. 1989). In 2012, there were approximately 400000 people diagnosed with NHL globally (Ferlay et al. 2012). To add to the burden of cancer population, HIV infection and AIDS increases the risk of developing NHL (Golematis & DeVita Jr 1996). NHL treatment currently involves the use of toxic, high dose chemotherapy, irrespective of HIV status or subtype. The aforementioned therapy has a high relapse rate due to the heterogeneity inherent within the disease. Therefore, there is need for alternative treatment regiments that are less toxic and tailor-made for a specific cancer.

Biomarkers are "*cellular, biochemical or molecular alterations that are measurable in biological media such as tissues, cells or fluids*" (Hulka 1990). Cancer biomarkers can be stratified into three categories: diagnostic biomarkers; prognostic biomarkers and predictive biomarkers (Hamdan 2007). Since proteins are the effectors of all biological processes, they have been the main target for biomarker discovery drives. Therefore, there is substantial interest in applying proteomics technologies towards the identification of disease biomarkers.

Proteomics involves the identification and characterization of a complete set of proteins/peptides at a given time (Chandramouli & Qian 2009). Mass spectrometry (MS) has the potential to yield comprehensive profiles of peptides and proteins in biological tissues without carrying out protein separations by gel electrophoresis (Zhou & Veenstra 2013). This type of approach is highly suited for biomarker identification due to reduced sample concentration requirements and high throughput (Han 2011). The most commonly used platforms for MS analysis are matrix assisted laser desorption/ionisation (MALDI) and electrospray ionisation (ESI). MALDI is frequently combined with time of flight (TOF) mass analysers while ESI is joined to liquid chromatography mass spectrometers (LC-MS) (Khadir & Tiss 2013). There are various sample types that can be used for protein extraction and analysis. One of the untapped tissue treasures include formalin fixed paraffin embedded (FFPE) tissues. FFPE tissue blocks have a history of being excellent substrates for morphological and immunological studies in pathology (Magdeldin & Yamamoto 2012). In addition, emerging studies show that it is possible to use FFPE tissues in mass spectrometry analyses; particularly in biomarker identification investigations (Craven et al. 2013; Sousa et al. 2012; Takadate et al. 2013).

This literature review will cover the scope of protein biomarkers characterised for HIV related and non-HIV related DLBCL. Following the literature review, the thesis will continue onto two chapters of results, each containing a brief discussion. These will then be followed by a final conclusion.

The first chapter of the results will compare the clinical characteristics of HIV and non HIV related DLBCL patients treated at Groote Schuur Hospital (GSH). The second

chapter will compare the differences in protein expression profiles between the two study groups. The expression of one of the differentially expressed protein will be confirmed using IHC and correlated with clinicopathological characteristics of the patients.

1.3 LYMPHOMA

1.3.1 The incidence of lymphomas

Lymphomas and leukaemia's are a part of the broader group of tumours of the hematopoietic and lymphoid tissues (Nayak et al. 2013). Lymphoma is the name applied to a group of blood cell tumours that develop from lymphatic cells (Müller et al. 2005; Hennessy et al. 2004). In 2012, statistics showed that lymphomas developed in 450,000 people worldwide and caused over 200,000 deaths (Ferlay et al. 2012). Lymphomas consist of 3–4% of all cancers, making them the seventh-most common form of cancer. The two main categories of lymphomas are Hodgkin lymphomas and the non-Hodgkin lymphomas, the majority being NHL (Jaffe et al. 2008; Aisenberg 2000). Symptoms of lymphoma may include the so called 'B' symptoms which consist of fever, sweating, and weight loss (Levine 1992).

1.3.2 Differences between lymphoma subtypes

Hodgkin lymphomas differ from non- Hodgkin lymphomas in several ways. The age distribution of HL is bimodal occurring in young adults as well as in the fifth decade while in NHL, the age distribution is unimodal peaking in the sixth decade (Rosolen et al. 2000; Giunti et al. 2002). The 'B' symptoms are more common in HL than in NHL. Initial presentation of HL is usually localised nodal disease in contrast to the generalised

nodal disease with extranodal presentation observed in NHL (Rosenberg & Kaplan 1970). HL are characterised by the presence of Reed-Sternberg cells while NHL mostly originate from B cells but NHL can also arise from T cell and natural killer cells (Giunti et al. 2002; Hennessy et al. 2004). Some NHL subtypes are associated with specific chromosomal translocations of proto-oncogenes into either the immunoglobulin gene or other genes (Rosolen et al. 2000; Vega & Medeiros 2003); while there are no known specific genetic defects in HL (Hudnall 2011)

1.3.3 Non-Hodgkin's lymphoma

Non-Hodgkin lymphomas are a group of neoplasms whose clinical presentation and natural history are more variable than in Hodgkin's lymphoma (Hoffbrand et al. 2006; Nayak et al. 2013). There are more than sixty subgroups of NHL, some having more than one subtype(Campo et al. 2011). Thirty percent of NHLs present in extranodal sites with the gastrointestinal tract being the most frequent site followed by the head and neck region constituting 11-33% of the cases (Nayak et al. 2013). A number of classification systems have been developed to categorize NHL. Early classification systems such as the Rappaport (1966) were based exclusively on lymph node architecture and cell morphology (Hu et al. 2011; Garvin et al. 1980). However, increased understanding of lymphocyte biology led to the Kiel and Lukes-Collins classification system in 1974, the Working Formulation for Clinical Usage (WF) in 1982 and the revised American European lymphoma (REAL) system in 1994 (Hoffbrand et al. 2006). The current World Health Organisation (WHO) discouraged the use of subdividing lymphomas on the basis

of histological appearance only to *'distinct syndromes each with a characteristic morphological, immunophenotypic, genetic and clinical features'*(Jaffe 2009).

1.3.4 Diffuse large B- cell lymphoma

Diffuse large B- cell lymphoma (DLBCL) is the most commonly diagnosed NHL subgroup, comprising of more than a third of NHL cases in South Africa(Lossos & Morgensztern 2006; Tilly & Dreyling 2009; Patel et al. 2015). It is composed of a heterogeneous group with different morphological, histopathological and genetic subtypes (Klapper et al. 2012). According to the WHO classification, diffuse large B-cell lymphoma is subdivided into several subtypes as follows: DLBCL, NOS; T cell/ histiocyte rich large B cell lymphoma; DLBCL associated with chronic inflammation; Epstein Barr positive DLBCL of the elderly (Jaffe 2009).

1.3.4.1 Morphological variants of DLBCL

DLBCL is histologically characterised by a diffuse proliferation of large neoplastic B lymphoid cells with a nuclear size equal to or exceeding that of normal histiocyte nuclei (Campo et al. 2011). In most cases, these neoplastic cells morphologically resemble either centroblasts or immunoblasts, and usually both are present in the same tumour (Dominis et al. 2002). There are three morphological variants of DLBCL, namely: centroblastic, immunoblastic and anaplastic (Figure 1.3) (Xie et al. 2015). The

centroblastic variant contains medium to large lymphoid cells with vesicular nuclei, irregular nuclear membrane and multiple nucleoli(Hunt & Reichard 2008). In the immunoblastic variant, greater than 90% of the cells contain a single, centrally located nucleolus and appreciable amount of cytoplasm (Hu et al. 2011). The anaplastic variant is a rare morphological variant of DLBCL characterized by the presence of large cells with bizarre pleomorphic nuclei that may resemble Reed –Sternberg cells and tumour cells of anaplastic large cell lymphoma (ALCL), thereby rendering them diagnostically challenging (Mona et al. 2014). Centroblastic lymphoma is the most common subtype with better prognosis when compared with immunoblastic and anaplastic variants of DLBCL (Dominis et al. 2002).

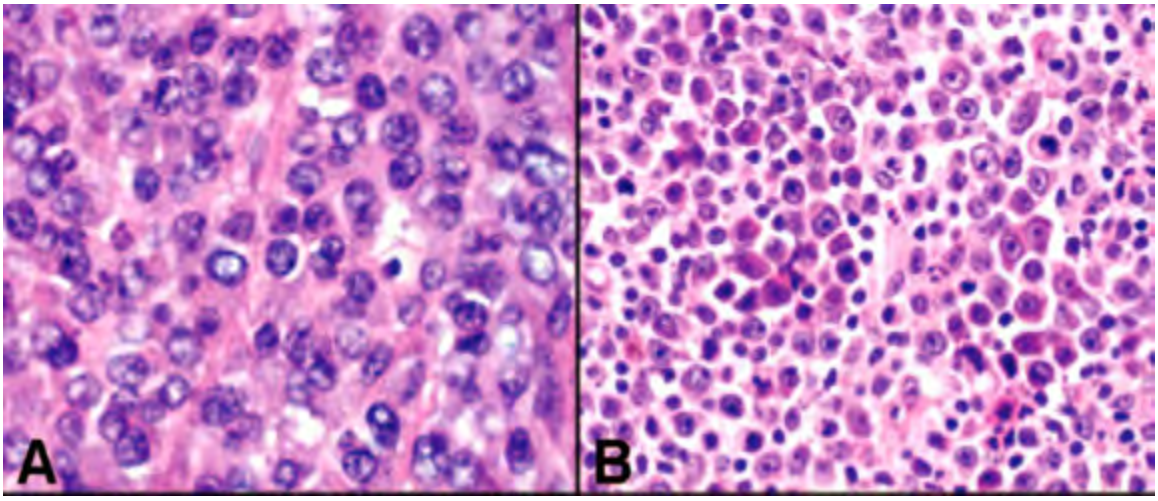


Figure 1.3: H&E sections of the morphological variants of DLBCL(Modified from Hunt & Reichard 2008). (A) The centroblastic variant have large blast cells with vesicular nuclei, irregular nuclear membrane and multiple nucleoli (X600).(B) The immunoblastic variant contains numerous blast cells with prominent single nucleolus, regular nuclear membrane and vesicular chromatin (X400).

1.3.4.2 Genetic variants of DLBCL

DLBCL consists of three genetic variants that point to the origin of the tumour; the so called 'cell of origin' (COO). The study by Alizadeh et al (2000) was the first to suggest that lymphomas retain the gene expression of the COO (Alizadeh et al. 2000; Xie et al. 2015). This suggests that different types of DLBCL arise from different stages of B cell differentiation (Lossos et al. 2000).These genetic DLBCL subtypes are not related to any of the morphological subtypes of DLBCL (Rosenwald et al. 2002). The study used gene

expression profiling (GEP) to classify DLBCL into two groups (Alizadeh et al. 2000).The first group had a high expression of genes expressed in the B- cells of the germinal centre and was named the GCB type. The second group had a genetic signature similar to activated B cells and was named ABC type (Alizadeh et al. 2000). The classification demonstrated prognostic significance with statistically different 5 year survival outcomes (GCB- 76% vs. ABC- 16%) (Alizadeh et al. 2000). A third group named 'type 3' was identified by Rosenwald et al (2002) and his colleagues in 2002 (Rosenwald et al. 2002). This subgroup was heterogeneous, not well defined and had similar prognosis to that of the ABC subtype (Rosenwald et al. 2002). The genetic variants of DLBCL also have differing genetic aberrations and molecular pathways. The GCB type is frequently associated with the t(14;18) (q32;q21) chromosomal aberration as well as gains at 7q, deletion of the tumour suppressor PTEN and losses at 16q (Barrans et al. 2003; Iqbal et al. 2004; Lenz et al. 2008; Guo et al. 2014). However, the non-GCB type is defined by gains at11q24.3 and 3q13.2 (Guo et al. 2014)as well as a constitutively active NF-kB pathway (Davis et al. 2001).

1.3.4.3 Immunohistological algorithms for COO classification

The GEP technology used in classifying DLBCL is expensive and not user friendly thus making it unsuitable for regular diagnostic practise (Hans et al. 2004). This feature gave rise to the search for protein markers that define each genetic DLBCL subtype that can be detected immunohistologically. A number of immunohistochemical algorithms have been

developed to stratify patients according to their COO type, using various markers (Hans et al. 2004; Muris et al. 2006; Natkunam et al. 2001; Choi et al. 2009; Meyer et al. 2010). These algorithms target proteins whose genes have been shown to be highly expressed in either GCB or ABC subtypes (Meyer et al. 2011). Proteins that are highly expressed in GCB include CD10, BCL6 and GCET1. Those highly expressed in ABC include: MUM1, cyclin D, CD44 and FOXP1 (Meyer et al. 2011; Dunleavy et al. 2013). The algorithms developed by Hans et al (2004) and Choi et al (2009) are the only ones that were highly correlated with GEP data. They were able to correctly classify samples according to GEP data by determining the protein expression of the genes that are commonly overexpressed in each subtype. The Hans algorithm uses three markers (CD10, BCL6 and MUM1), while the Choi algorithm uses five markers (GCET1, CD10, MUM1, BCL6, FOXP1) (Hans et al. 2004; Choi et al. 2009). The other algorithms were not correlated with GEP data and relied on prognostic differences between the groups (Muris et al. 2006; Natkunam et al. 2001; Meyer et al. 2010). Both the Hans and Choi algorithms show high (above 80%) concordance with GEP data (Choi et al. 2009). However the Choi algorithm is not user friendly and some of the markers used are not commonly found in clinical practise (Meyer et al. 2011).

1.4. HUMAN IMMUNODEFICIENCY VIRUS INFECTION /ACQUIRED IMMUNE DEFICIENCY SYNDROME (HIV/AIDS)

The human immunodeficiency virus (HIV) belongs to the lentivirus genus of the retroviridae family (Coffin JM, Hughes SH 1997). HIV infects the immune T lymphocytes using receptors that these cells express for entry (Wilen et al. 2012). HIV infection progresses in stages and culminates in acquired immune deficiency syndrome (AIDS). The following criteria are required to diagnose an infected patient as having AIDS: a CD4 count of less than 200 cells/ mm^3 and/or presence of AIDS defining illness (World Health Organisation 2007). The AIDS defining illnesses encompass a spectrum of diseases which can be broadly classified as opportunistic infections and AIDS related malignancies (Deeks et al. 2013). AIDS related malignancies include Kaposi sarcoma (KS), NHL and invasive cervical cancer (Bower et al. 2006; Centers for disease control 1987).

Treatment for HIV infection involves targeting the HIV virus at various points of the life cycle. There are various drugs available that affect the HIV life cycle including: protease inhibitors; chemokine receptor antagonist; integration inhibitors; reverse transcription inhibitors; and fusion inhibitors (Maartens et al. 2014). The current therapeutic strategy (highly active anti-retroviral therapy [HAART]) involves the use of anti-retroviral drugs from at least two distinct classes (World Health Organisation 2013).

1.4.1 Epidemiology of HIV/AIDS in Southern Africa

In 2013, there were 24.7 million people living with HIV in Sub-Saharan Africa. This amounted to 71% of the global population (UNAIDS 2015). South Africa has one of the highest HIV prevalence rates in Sub-Saharan Africa which accounted for 25 % of the global figures in 2013 (UNAIDS 2015). The number of people living with HIV in 2014 was 5.51 million (Shisana et al. 2012). The higher prevalence rate may be partly due to infected people living longer because of access to HAART, thus increasing the disease burden even if new infections are on the decrease (Shisana et al. 2012). In South Africa there are currently 2.6 million people receiving HAART treatment (UNAIDS 2015).

1.4.2 The role of HIV in lymphomagenesis

Currently, there is little understanding of how HIV induces lymphomagenesis in AIDS related lymphomas, which are predominantly of B-cell origin (Monroe & Silberstein 1995). There are two main factors that play a role in HIV lymphomagenesis: cellular immunodeficiency and oncogenic viruses (Besson & Raphaël 2003). The HIV-associated $CD4^+$ T cellular depletion leads to the increase of Epstein Barr virus (EBV) infected B-cells and to the decrease of antitumor immunity. The EBV oncogenic virus is frequently associated with HIV related NHL. Clonal EBV genome is found in lymphoma cells in 30 to 70% of cases of HIV-related NHL. It expresses oncogenic proteins including Latent

membrane protein 1(LMP-1) which behaves like an activated CD40 receptor. LMP-1 induces the expression of intracellular genes which stimulate cell growth and inhibit apoptosis (Besson & Raphaël 2003). In addition the increased production of B-cell stimulating cytokines such as IL-6 andIL-10 further promote lymphomagenesis (Pantanowitz et al. 2015).

1.5. HIV/AIDS Related Lymphomas

AIDS related lymphomas (ARL) are frequently of B-cell phenotype and can be divided into three categories: lymphomas occurring in immunocompetent patients; lymphomas occurring only in HIV infected patients; and lymphomas occurring in other immunodeficient states (Pantanowitz et al. 2015; Vishnu & Aboulafia 2012). DLBCL and Burkitt lymphoma (BL) are the most common lymphoma subtypes seen in HIV infected patients also occurring in immunocompetent patients (Wiggill et al. 2013). ARL also commonly occur as 'grey zone' lymphomas which are intermediates between lymphoma subtypes, such as the lymphoma intermediate between BL and DLBCL subtype (Tirelli et al. 2000). In general, HIV-related lymphomas manifest late during HIV infection. However, these lymphomas can arise in patients with any CD4 count (Akanmu 2006). BL frequently occurs in patients with relatively high CD4counts (>200 cells/mm^3), whereas primary central nervous system lymphoma (PCNSL) and DLBCL

are more likely to be seen in advanced stages of immune deficiency (CD4 <50 cells/mm^3) (Biggar et al. 2007; Pantanowitz et al. 2015).

1.5.1 Risk factors for AIDS related lymphoma

The relative risk of acquiring ARL increases 600 fold in patients diagnosed with AIDS (Levine et al. 2001).Epidemiologic studies have failed to identify major environmental factors associated with the development of lymphoma among HIV-infected individuals (Levine et al. 2001)However, genetic and host factors have been shown to influence the development of AIDS related lymphoma. Whilst a decreasing HIV viral load and increasing CD4 count are protective (Bohlius et al. 2009), older individuals (> 37years), the lowest CD4 cell count and a prior AIDS diagnosis have been identified as risk factors for developing ARL (Bohlius et al. 2009). In addition, longer periods of immunosuppression and uncontrolled viral loads have also been identified as risk factors for the development of ARL (Simard & Engels 2010).

1.6 PROTEOMICS

The human genome is made up of 26000–31000 genes (Baltimore 2001). However, the protein products are in the range of 2-4 million due to splice variants and posttranslational modifications (PTMs) which include glycosylation, phosphorylation, and ubiquitination (Milo 2013; Hirsch et al. 2004). Moreover, the proteomes of cells, tissues, and body fluids are complex and display a wide dynamic range of proteins concentration (Chandramouli & Qian 2009). Proteomics involves the identification and characterization of a complete set of proteins present in a cell, organ, or organism at a given time (Chandramouli & Qian 2009). In general, proteomic techniques have numerous applications including protein profiling, expression analysis, identification of PTMs, and protein–protein interactions (Monti et al. 2005; Shiio & Aebersold 2006).

1.6.1 Mass spectrometry in proteomics

Initial approaches to investigate the proteome of cell lysates and body fluids were performed using 2D polyacrylamide gel electrophoresis (2D-PAGE) (Hirsch et al. 2004). MS is at the heart of virtually all proteomics experiments as it provides the key tools for the analysis of proteins (Guerrera & Kleiner). Mass spectrometers consist of three basic components: an ion source, a mass analyser, and an ion detector. MS measurements are carried out on ionised analytes in the gaseous phase, requiring a method to transfer molecules from solution or solid phase into this state. The two most commonly used

techniques for ionisation are matrix assisted laser desorption/ionisation (MALDI) and electrospray ionisation (ESI). MALDI and ESI are both soft ionisation techniques in which ions are created with low internal energies and thus undergo little fragmentation (Karas et al. 1987; Nguyen & Fenn 2007). After ionization, the sample ions reach the mass analyser which separates ions by their mass-to-charge (*m/z*) ratios. In proteomics research, four basic types of mass analysers are currently being used: time-of-flight (TOF), ion trap, quadrupole, and Fourier transform ion cyclotron resonance (FTICR) analysers (Han et al. 2008). MALDI is usually coupled to TOF analysers, which separate ions according to their flight time down a field-free tube (Calderaro et al. 2014). The ions' time-of-flight is directly related to their m/z values and thus a mass spectrum can be acquired. ESI is most frequently coupled to ion traps where the ions are first captured in the centre of the device for a certain time interval, followed by scanning from the trap to the detector. This type of trap allows the detection of mass and sequence data (Banerjee & Mazumdar 2012).

1.6.2 MALDI mass spectrometry

In MALDI, small volumes of samples are co-crystallised with an equal volume of organic matrix on a metal target. A pulsed laser is used to excite the matrix, which causes rapid thermal heating of the molecules and eventually desorption of ions into the gas phase (Karas et al. 1987). MALDI produces packets of ions rather than a continuous beam because of the usage of a pulsed laser, therefore, it is most often coupled to a mass

analyser that can measure either a complete mass spectrum without scanning a mass range or trap all the ions for subsequent mass analysis (Calderaro et al. 2014). It is the most commonly used technique because of the generation of singly charged ions, low concentration of sample required, simple preparation method, and the ability to analyse complex mixtures (Hirsch et al. 2004; Chung et al. 2007). Limitations of the technology include poor mass resolution of signals above *m/z* 50,000, the ionization of the most abundant proteins, lack of reproducibility and the fact that approximately 90% of the observed signals are below *m/z* 30,000 (Hirsch et al. 2004; Byrd & McEwen 2000).

1.6.3 MALDI imaging mass spectrometry

An adaptation of the MALDI technique-MALDI imaging mass spectrometry (IMS) is used in profiling and imaging of proteins directly from tissue sections (Caprioli et al. 1997). The choice between profiling and imaging proteins depends on the overall experimental objectives. Using tissue profiling methodology, the matrix is deposited in discrete spots on regions of interest on a tissue section. Mass signals amounting to 100 laser shots per spot are detected in *m/z* 500-25000 range (Balluff et al. 2011). The mass profiles between defined areas of the tissue can then be compared. During an imaging experiment, a tissue section is uniformly coated with matrix solution so that protein profiles can be acquired over the entire area (Walch et al. 2008). Mass spectrometry data are acquired using a pattern of laser spots. The distance between spots are fixed and depends on the resolution, typically ranging from 25 to 200 μm. The number of laser shots per spot is predetermined and is usually 20–50 shots are recorded (Cornett et al. 2007). A two dimensional ion density image can be reconstructed from the intensity of a given m/z value in each spectrum.

1.6.4 FFPE tissues for MALDI IMS

Formalin fixation paraffin embedding is a preservation method with a long history in pathology with millions of clinical samples available in biobanks around the world (Eriksson et al. 2013). The formalin fixation caused protein-protein crosslinking as well as nucleic acid-protein crosslinking, thus preserving the integrity of the tissue (Sompuram et al. 2004). However, the formalin also causes the degradation and modification of nucleic acid and proteins (Nam et al. 2014). Despite this, these types of tissues have become invaluable in the process of biomarker identification due to the clinical data accompanying the tissue sample. Therefore, various sample preparation methods have been developed for reversing the crosslinking process and accessing the proteins stored in these tissues. In MALDI IMS, these include the use of *in situ* trypsin digestion (Lemaire et al. 2007) step or an antigen retrieval step with a selected buffer (Ronci et al. 2008; Casadonte & Caprioli 2011; Gustafsson et al. 2010) prior to digestion. Although the latter protocol produces enhanced signal (Ronci et al. 2008), there is an introduction of bias towards the types of proteins 'unlocked' since the buffer used is not constant and has to be optimized for each tissue type (Sompuram et al. 2004).

1.7 BIOMARKERS

Biomarkers are molecular indicators of a biological status, and as biochemical species can be assayed to evaluate the presence of cancer and therapeutic interventions (Tainsky 2009). Cancer biomarkers are discovered and utilized with a specific purpose including (a) early detection of cancer, (b) diagnosis, (c) prognosis, (d) response to anticancer therapies or (e) cancer recurrence (Tainsky 2009). Numerous biomarkers have been identified thus far (Geisler et al. 2015). However, these have failed to be implemented in clinical practice because they failed to meet the high specificity and sensitivity criteria for clinical use. Therefore, panel biomarkers are expected to be more specific and sensitive for a particular cancer.

Biomarker identification workflow involves three stages: discovery phase, verification phase, and validation phase. The discovery phase can be done using two approaches. The first approach is knowledge based, where a biomarker is chosen using known molecular mechanisms of the disease. The second approach is untargeted and a biomarker is chosen from the differentially expressed proteins between two groups (Frantzi et al. 2014). The verifications phase seeks to confirm the presence and levels of the candidate biomarker. The validation phase evaluates the clinical utility and performance of the candidate biomarker in a clinical setting.

1.7.1 Biomarkers for DLBCL

Biomarkers may originate from three main sources within cells: nucleic acids, proteins and metabolites (Tainsky 2009).

1.7.2 International prognostic index (IPI)

The International Prognostic Index (IPI), although not a biomarker, is the gold standard used for predicting the outcome for patients with NHL and is therefore a clinical surrogate for biomarkers (Banham et al. 2005). The IPI is based on a combined score of five clinical parameters (age, performance status [PS], Ann Arbor stage, extranodal involvement, and lactate dehydrogenase [LDH] level (Zhou et al. 2014). It only reflects tumour and host status biology but this alone cannot accurately predict treatment outcome for individual patients (Xie et al. 2014). Therefore, there is a need for additional methods or markers to identify high-risk patients that can benefit from targeted therapy (Banham et al. 2005).

1.7.3 Biomarkers for DLBCL independent of the IPI

There are many individual prognostic biomarkers that have been shown to be significance in DLBCL. These belong to a variety of cell processes such as cell cycle regulators, apoptotic proteins, and proteins involved in B cell differentiation (Lossos & Morgensztern 2006). Below, we will focus on some of the frequently appearing in literature and form part of the clinical diagnostic process.

1.7.3.1 Ki-67

Ki-67 is a nuclear protein that is synthesized at the beginning of cell proliferation and is expressed throughout the active part of the cell cycle (He et al. 2014). Its expression is detected immunohistochemically by dividing the number of Ki-67 positive cells with the total number of cells (Broyde et al. 2009). Ki-67 expression is correlated with the proliferation rate and is therefore used to evaluate the proliferative activity of lymphomas (Miller et al. 1994; He et al. 2014). The prognostic effect of Ki-67 was noticed in retrospective studies when the proportion of Ki-67 positive cells was correlated with grade and outcome in NHL cases (Miller et al. 1994; Broyde et al. 2009; Yoon et al. 2010; Li et al. 2012). However, there is no consensus on the cut-off value for Ki-67 expression that can predict outcome. Generally, a value above 70% seems to have prognostic value. Broyde et al. (2009) determined that a cut off value of 70% was prognostic in DLBCL. Low expressing DLBCL patients had good prognosis compared to the high expressing patients treated with both CHOP and R-CHOP (Broyde et al. 2009). Similar results were also observed when Ki-67 was combined with other prognostic factors such as low IPI and bulky disease (Broyde et al. 2009). A study by Yoon et al (2010) used a prognostic cut off value of 85% in patients treated with R-CHOP and CHOP. Patients with higher Ki-67 expression levels had higher rates of relapse after treatment with R-CHOP (Yoon et al. 2010). In addition, the high expression group had significantly lower 2 year event free survival (EFS) and overall survival (OS) rates (Yoon

et al. 2010). In another Chinese study, a lower cut off value of 75% can distinguish patients with a good or bad prognosis (Li et al. 2012). As with the previous study, high expression of Ki-67 was correlated with poor OS and PFS in DLBCL patients treated with R-CHOP (Li et al. 2012). Therefore, although there is no consensus on the cut-off values for Ki-67, expression levels above 70% seem to be prognostic. In addition, high expression levels indicate a poor prognosis for patients treated with both CHOP and R-CHOP.

1.7.3.2 BCL- 2

B-cell leukemia/lymphoma-2 gene (BCL-2) forms part of a family of apoptotic regulatory proteins. BCL-2 functions as an anti-apoptotic factor that is important in normal B cell development and is frequently overexpressed in DLBCL (Dunleavy & Wilson 2011; Iqbal et al. 2004). One well-characterized mechanism of BCL-2 over-expression is the t(14; 18)(q32; q31) translocation with the Ig gene. This rearrangement is largely restricted to the GCB subtype (Iqbal et al. 2004), whereas in the ABC subtype, the mechanism of BCL2 overexpression is associated with constitutive NF-kB activation, with or without 18q21 amplification (Iqbal et al. 2006). While the BCL-2 gene rearrangement has no prognostic significance (Iqbal et al. 2004; Kramer et al. 1998), the prognostic significance of BCL-2 protein expression is controversial and subtype dependent. In the CHOP era, BCL-2 protein expression was indicator of poor outcome

indicator, especially in the ABC subtype (Sohn et al. 2003; Iqbal et al. 2004; Dunleavy & Wilson 2012). However, in the post rituximab era, BCL-2 is prognostic in the GCB subtype (Perry et al. 2014), raising questions about the differential activity of rituximab within the GCB and ABC subtypes of DLBCL (Dunleavy & Wilson 2012).

1.7.3.3 *BCL 6* Gene Rearrangements

BCL6 is a zinc finger transcription factor that functions as a proto-oncogene, specifically acting as a transcriptional repressor needed for germinal centre formation (Westin & Fayad 2009). The BCL6 gene is found to be frequently translocated and hypermutated in DLBCL. Rearrangement of the *BCL6* proto-oncogene at chromosome 3q27 is the most frequent cytogenetic abnormality in DLBCL, occurring in up to 35% of cases (Shustik et al. 2010; Akyurek et al. 2012). *BCL6* rearrangement is associated with elevated lactate dehydrogenase (LDH) and advanced age (Jerkeman et al. 2002). No association between *BCL6* rearrangement status and expression of BCL2, Ki-67 or TP53 was found. The evidence for the prognostic significance of *BCL6* rearrangement in DLBCL is contradictory. Some studies report no significance (Jerkeman et al. 2002), Kramer et al 1998) while others do (Akyurek et al. 2012). Differences in the rearrangement detection method, cut-off value as well as the classification of DLBCL may account for the differences. What is consistent is that *BCL6* rearrangement is associated with the non-GC phenotype, therefore pointing to the involvement of *BCL6* in the process of B cell differentiation (Akyurek et al. 2012; Shustik et al. 2010). In addition, the detection of

BCL6 protein using IHC does not confer any prognostic significance (Akyurek et al. 2012; Shustik et al. 2010).

1.7.3.4 P53 Genetic mutations and protein expression

The P53 gene is a tumour suppressor gene located on chromosome 17 and encodes a 53kDa protein (Harris 1996). P53 is known to play a critical role in cellular responses since it responds to damaged DNA by induction of G1 arrest, and if DNA is not repaired, p53 may induce apoptosis (Carbone & Gloghini 2014). The dysfunction and mutation of p53 can result in abnormal cell growth, lack of apoptosis and eventually malignant transformation (Zainuddin et al. 2009). P53 mutations (point mutations and deletions) were observed to predict inferior survival in low risk DLBCL cases as well as GC phenotype cases (Leroy et al. 2002; Zainuddin et al. 2009) in the CHOP era. However, the addition of rituximab, resulted in p53 mutations additionally having a prognostic significance for the non-GC phenotype (Stefancikova et al. 2011; Xu-monette et al. 2012). The p53 protein expression was also determined to be prognostically significant with strong expression (>30% staining) having an adverse effect on overall survival (Leroy et al. 2002; Xie et al. 2014). There are limited studies investigating the association between a singular mutation/deletion with survival of DLBCL patients. A recent study attempted to fill this gap by determining the effect of the P53 rs1625895 polymorphism in DLBCL patients treated with R-CHOP (Voropaeva et al. 2015). The study observed an

increased relapse rate and lower overall survival in patients having the G/G phenotype compared to the G/A or the A/A phenotype of the polymorphism (Voropaeva et al. 2015). A study by Young et al (2008) showed that mutations in different domains of the protein have an impact on prognosis. Patients with mutations in the DNA- binding domain had the poorest overall survival rate, while mutations in the loop-L2 region showed no effect on patient survival (Young et al. 2008).

1.7.3.5 FOXP1

Forkhead box protein P1 (FOXP1) is a member of the FOX family of transcription factors characterized by a common DNA-binding, winged-helix domain with N-terminal zinc finger and leucine zipper domains (Sagaert et al. 2006; Jack & Barrans 2004; Barrans et al. 2004). GEP studies have reported upregulated mRNA expression of the FOXP1 forkhead transcription factor in response to B-cell activation(Banham et al. 2005). In addition, FOXP1 mRNA is upregulated in the non-GC subtype of DLBCL (Barrans et al. 2004; Hoeller et al. 2010). FOXP1 protein is therefore overexpressed in a subset of DLBCL cases and is reported to have prognostic significance. FOXP1 protein overexpression is associated with lower disease-specific and overall survival in nodal and non-GCB DLBCL cases (Barrans et al. 2004; Banham et al. 2005; Hoeller et al. 2010). The prognostic significance of FOXP1 expression was independent of the IPI (Barrans et al. 2004; Banham et al. 2005; Tzankov et al. 2015). However, a recent study has reported that the extranodal DLBCL group expressing FOXP1 protein also has worse outcome(Yu et al. 2011). The discrepancies in the data may be due to differences in the IHC antibody used (which may detect related FOX proteins), IHC conditions, as well as the cut-off value for positive expression of FOXP1.

1.7.3.6 *MYC* Gene Rearrangement

The *MYC* proto-oncogene family comprises three genes encoding for the transcription factors MYC, NMYC, and MYCL that regulate multiple functions in normal cells and have a strong oncogenic potential (Karube & Campo 2015). *MYC* gene rearrangement is characteristics of BL and occurs in 10% of DLBCL cases and these alterations are frequently associated with *BCL-2* or *BCL-6* rearrangements (Lin et al. 2012; Karube & Campo 2015).*MYC* gene alterations have prognostic significance in lymphoma cases intermediate between BL and DLBCL (Lin et al. 2012). *MYC* rearrangements were observed to be frequently found in extranodal cases of DLBCL and these cases had a higher DFS rate (Kramer et al. 1998). While in another study *MYC* rearrangements were associated with lower overall survival, and this was apparent in the GCB subtype (Akyurek et al. 2012; Perry et al. 2014).

Lymphomas with BCL-2 and MYC genetic rearrangements are referred to as "Double Hit" lymphomas (Aukema et al. 2011).These genetic aberrations lead to an overexpression of the protein which can have prognostic significance (Johnson et al. 2012). In DLBCL, the overexpression of both MYC and BCL-2 protein has been shown to signify poor prognosis (Johnson et al. 2012). The prognostic significance of MYC and BCL-2 co-expression is similar in DLBCL subtypes, although the aberrations are common in the non-GC subtype (Hu et al. 2013).

1.7.4 Biomarkers for HIV-related DLBCL

Prognosis in HIV related DLBCL is determined by factors relating to the tumour as well the underlying immunodeficiency (Kaplan 2012). Previously in the pre-HAART era, the state of immunodeficiency played a major role in determining prognosis for HIV related NHL cases (Lim et al. 2005). However, in the post-HAART era features relating to the tumour, such as the IPI, have surfaced as prognostic indicators (Bower et al. 2006; Kaplan 2012; Tedeschi et al. 2012). There are limited studies investigating the impact of molecular indicators on survival, specifically in HIV-related DLBCL. Preliminary data indicate that markers including *MYC*, *BCL-6* and *MUM1* are overexpressed in HIV related DLBCL (Chao et al. 2015). In addition, markers that may be significant predictors of survival in HIV related DLBCL are EBV and MYC expression in patients with low CD4 counts (Silverberg et al. 2012). EBV positivity is known to be associated with the NF-kB pathway as well as increased mortality in patients with HIV related DLBCL (Chao et al. 2013). Chadburn et al (2009) concluded that the known markers for prognosis in non HIV related DLBCL do not predict any prognostic significance in HIV related DLBCL. However this study combined patients receiving four different types of treatment in their analyses (Chadburn et al. 2009). Therefore, more studies are required in determining biomarkers for HIV positive DLBCL so as to better understand the pathology of the disease.

1.7.5 Biomarkers for DLBCL identified by proteomic methods

There are a number of proteomic studies performed to determine biomarkers for DLBCL. Differences between early relapse patients and progression free DLBCL patients have been quantified using a SILAC based method in pre-treatment frozen tissues (Rüetschi et al. 2015). The authors found an enrichment of proteins involved in the regulation and organization of the actin cytoskeleton that could discriminate the two patient groups (Rüetschi et al. 2015). Another study that attempted to elucidate how the tumour responds to treatment was conducted using 2D coupled to mass spectrometry. The authors identified 19 differentially expressed proteins between chemoresistance and chemosensitive causing chemotherapy resistance. The study highlighted proteins involved in cell proliferation and apoptosis such as the FGF signaling pathway as well as the Ca2+- calmodulin-related signaling pathways as the main driving force behind chemoresistance (Liu et al. 2013). An additional SILAC based study attempted to segregate the GCB and ABC subtypes of DLBCL. The main proteins that were differentially expressed between the subtypes included cell surface markers CD44 and CD27, transcription factors IRF4 and SPI1 (Deeb et al. 2012). Several studies have used serum to identify biomarkers which have high accuracy and specificity (Zhang et al. 2007). However, the identities of these potential biomarkers are unknown because the classification models can be used as is without identifying the peaks. Researchers need to take a further step and identify the proteins that make up the classification models.

Knowing the identity may assist in better understanding the disease processes and thereby assist in developing alternate therapeutic options.

There is currently no literature on potential biomarkers for HIV related DLBCL discovered using proteomic methods.

1.7.6 Disadvantages of proteomic biomarkers

The search for new biomarkers for cancer is a great step towards finding alternative therapeutic agents that can be used to effectively treat all variants of DLBCL. However, before any of the biomarkers identified can be ready for clinical trials a number of challenges need to be overcome. One of the major challenge is to conquer errors of study design and/or experimental execution (Drucker & Krapfenbauer 2013). Studies need to have well-characterized and adequately stored clinical samples collected using recommended procedures together with detailed information on clinical, pathological parameters and follow-up (Matta et al. 2010). Another challenge is limited sample number. The majority of biomarker studies use limited clinical samples as large numbers of samples is beyond the scope of a single centre (Mayeux 2004). In addition, due to the heterogeneity of cancers, single biomarkers may not offer enough sensitivity and specificity (García-Bilbao et al. 2012). Therefore, the solution would be to validate a panel of potential biomarkers in multicentre studies so as to determine sensitivity and specificity.

1.8 AIMS AND OBJECTIVES

As mentioned in section 1.7.3 there has been considerable focus on determining biomarkers for DLBCL in the HIV negative context as this will aid in advancing knowledge on the development and progression of this disease. In contrast, there are limited studies on molecular markers for DLBCL in the HIV positive context, thus leading to a poor understanding of its development and progression.

1.8.1 Aim

- The aim of the project is to identify biomarkers for HIV and non- HIV related diffuse large B- cell lymphoma

1.8.2 Objectives:

- Determine the clinicopathological differences between HIV and non-HIV related DLBCL.
- Identifying protein biomarkers present in both the HIV and non-HIV related DLBCL using MALDI Imaging MS.
- Characterize the identified proteins using LC-MS/MS.
- Correlate the resulting molecular data with clinicopathological information.
- *Determine the expression levels of microRNA-21 in HIV and non-HV related DLBCL* (*This part of the project was completed and published as seen in Appendix P).*

CHAPTER 2: THE CLINICOPATHOLOGICAL FEATURES OF DLBCL IN HIV NEGATIVE AND POSITIVE PATIENTS

2.1 INTRODUCTION

Diffuse large B cell lymphoma makes up 30% of all NHL cases in developed countries with an even higher frequency in developing countries (Zhou et al. 2013; Mey et al. 2012). DLBCL can occur at all ages but has a median age at diagnosis of 64 years (Hunt & Reichard 2008). The disease is aggressive in nature and occurs at nodal or extranodal (EN) sites (Hunt & Reichard 2008). Up to 40% of DLBCL cases are extranodal with the gastrointestinal tract being the most common site (López-Guillermo et al. 2005; Vishnu & Aboulafia 2012). In 1993, the International prognostic index (IPI) was established as a prognostic predictor for NHL and is now the gold standard for determining prognosis (Sehn et al. 2007; Shipp & The international non-Hodgkins lymphoma prognostic factors 1993). Treatment involves the administration of six to eight cycles of CHOP (cyclophosphamide, doxorubicin, vincristine, prednisone) (Mey et al. 2012) with complete response (CR) rate between 48-60% (López-Guillermo et al. 2005; Montes-Moreno et al. 2010). The addition of the CD20 antibody, rituximab (R), has improved the complete response rate by 10% (Salles et al. 2011) .

HIV associated DLBCL is also the most common HIV associated lymphoma, with a frequency of up to 50% (Wang & Castillo 2011; Patel et al. 2015). Prior to the HAART era, HIV positive patients had a 60-200 times increased risk of developing NHL compared to the HIV negative population (Beral et al. 1991). However, since the introduction of HAART the risk of acquiring NHL has decreased due to increased CD4 counts and decreased viral load (Grogg et al. 2007). The introduction of HAART has also improved the outcome of HIV positive patients with DLBCL (Patel et al. 2015; de Witt et al. 2014).

The morphology of HIV-DLBCL tumours are similar to that seen in HIV uninfected counterparts (Gloghini et al. 2013b). However, there may be significant clinicopathological differences observed between the two groups (Barreto et al. 2012). The aim of the study was to determine the differences in the clinicopathological features of HIV positive and negative patients diagnosed with DLBCL at a single centre tertiary hospital.

2.2 MATERIALS AND METHODS

2.2.1 Ethics approval

The study was approved by the University of Cape Town Human Research Ethics Committee (reference HREC REF: 261/2010). The research proposal was approved by the Department of Clinical Laboratory Sciences Research Committee and the Faculty of Health Sciences Postgraduate Committee at the University of Cape Town.

2.2.2 Case selection

The study was a retrospective analysis of all DLBCL patients diagnosed and treated at Groote Schuur Hospital over 11 years (1st January 2003 until 31th December 2013).An electronic search was performed for cases diagnosed as “malignant lymphoma large cell diffuse, NOS” and “diffuse large cell lymphoma” from the laboratory information system database of the Division of Anatomical Pathology, National Health Laboratory Service,

using the Systemized Nomenclature of Medicine (SNOMED) code: M-96803. The generated list was used to search for patient folders at the Department of Radiation Oncology, GSH. The SNOMED codes, M-95953, M-95913, M-95903 were also used to search for additional cases.

2.2.3 Folder selection

The information from the list generated from the large cell lymphoma SNOMED mentioned in section 2.2.2 was used to look for the Department of Radiation Oncology folder number (RT number) which was different from the main hospital folder number. Baseline demographic and clinicopathological data were extracted, including treatment, response and survival. The following cases were excluded:

- Cases with no RT number
- Cases with missing folders
- Cases not diagnosed as Diffuse Large B cell lymphoma
- Cases with no HIV test result
- Patients with lymphoma intermediate between BL and DLBCL

The following DLBCL subtypes were included:

- Sclerosing DLBCL (4 cases)
- T-cell/histiocyte rich large B cell lymphoma (10 cases)
- DLBCL, NOS

2.2.4 Treatment of DLBCL patients

The majority of DLBCL patients were treated with CHOP therapy without rituximab. Twelve of the HIV negative patients from 2012-2013 were treated with R-CHOP. Intrathecal chemotherapy was used in patients with central nervous system (CNS) disease. Some patients in this study also received dose-reduced chemotherapy, palliative radiotherapy and involved field radiotherapy. HIV negative patients who had poor response to CHOP or relapsed were considered for high-dose salvage chemotherapy and stem cell transplantation. HIV positive patients who responded poorly to CHOP received palliative radiotherapy and/or chemotherapy. All HIV positive patients received HAART either prior or after diagnosis (Pather et al. 2013).

2.2.5 Classification into nodal and extranodal

We decided to classify nodal and extranodal tumours using criteria stated in Moller et al (2004). Lymphomas with clinically dominant lymph node involvement, as well as those presenting at the spleen or Waldeyer's ring, were considered as nodal lymphoma. Patients were defined as having extranodal DLBCL when the disease after the staging procedures was confined to one or more extranodal sites and with no (or only minor) nodal involvement. Primary tumour site was categorized into gastrointestinal (GI) tract, breast, hematologic system (blood, and bone marrow), genitourinary tract, head and neck, liver/pancreas, respiratory system, skeletal tissue, skin/soft tissue and thyroid.

2.2.6 Statistical analysis

The statistical analysis of the data was performed using Stata 12 unless otherwise stated. Comparisons between numerical variables were conducted with an unpaired t test. Categorical variables were compared using chi-square or Fischer's exact test, depending

on frequency levels. The linear regression analyses were completed using STATISTICA 12 with default settings. Categorical variables were coded and correlated with patient outcome (Appendix A). Overall survival was calculated from the date of first treatment to the date of death or last contact. Kaplan- Meier survival curves were drawn for overall survival times using GraphPad Prism 6.

2.3 RESULTS

2.3.1 Baseline characteristics of HIV negative and HIV positive DLBCL

There were 556 cases diagnosed as DLBCL from 2003-2013 using the NHLS database, however only 263 folders were located (Table 2.1). Patient age at diagnosis ranged from 13 years to 92 years, with a median of 52 years (Table 2.1). The majority of the located cases were HIV negative (205/263) but a few of the cases had no record of HIV status (4/263) (Table 2.1). The HIV positive cases had a statistically significant lower age at diagnosis with a median of 40 years, compared to 55 years of the HIV negative cases ($p< 0.0001$) (Table 2.1). There were only 219 cases with recorded gender, with females being the predominant gender (117 females vs. 102 males). This was observed in both HIV negative and HIV positive groups (Table 2.1). There were 192 and 200 cases with B symptoms and bone marrow involvement recorded, respectively. Equal numbers of the HIV negative and HIV positive DLBCL cases presented with B symptoms (35% vs. 34.7%) and bone marrow involvement (BMi) (42% vs. 46%).

Table 2.1: The clinical biological parameters of DLBCL from HIV infected and HIV uninfected patients

	ALL DLBCL		HIV- DLBCL		HIV+ DLBCL		P Values
	N	%	N	%	N	%	
Number of patients	263[1]	100	205	77.9	54	20.5	
Median age (range)	52 (13-92)		55 (13-92)		40(14-72)		<0.0001
Male Gender	102 (219)[2]	46.6	79(166)	47.6	23 (53)	43.2	0.594
N (Total)							
Ann Arbor Stage (Total)	244		195		49		0.057
I	45	18.44	36	18.46	9	18.37	
II	66	27.05	59	30.26	7	14.28	
III	33	13.52	22	11.28	11	22.45	
IV	100	40.92	78	40	22	44.90	
IPI Score (Total)	84		82		2		0.121
0	9	10.71	9	10.98			
1	19	22.62	19	23.2			

[1]4 cases had no recorded HIV status
[2] 44 cases did not have recorded gender

2	14	16.67	14	17.1			
3	20	23.81	20	24.4			
4	16	19.04	14	17.07	2	100	
5	6	7.14	6	7.32			
Elevated LDH (> 480 U/L)	56 (81)	69.14	42 (66)	63.64	14 (15)	93.33	0.025
ECOG Performance Status	238		187		51		0.980
0	18	7.56	14	7.49	4	7.84	
1	107	44.96	85	45.45	22	43.14	
2	40	16.81	30	16.04	10	19.61	
3	48	20.17	38	20.32	10	19.61	
4	25	10.50	20	10.7	5	9.8	
B symptoms	67 (192)	34.9	50 (143)	35	17 (49)	34.7	0.973
Bone Marrow Involvement	57 (200)	28.36	42 (154)	27.27	15 (46)	32.61	0.482
Tumour site							0.353
Nodal	127 (249)	51.00	101 (198)	51.01	26 (51)	50.98	
Extranodal	122 (249)	49.00	97 (198)	48.99	25 (51)	49.02	

There were 244 and 238 cases with Ann Arbor stage and performance status (PS) recorded, respectively (Table 2.1). The majority of the HIV negative cases were classified as stage II and stage IV (30.26% and 40%) (Table 2.1). In contrast, the majority of cases in the HIV positive group were classified as stage III and stage IV (44.90% and 22.45%). The predominant cases (44.96%) with recorded data had a performance status of 1. This was observed for both HIV negative (45.45%) and HIV positive (43.1%). There were 84 cases with IPI recorded, of which 82 were HIV negative. Only 2 of the HIV positive cases had IPI status calculated and recorded.

Just 51% of all DLBCL tumours were nodal lymphomas (Table 2.1). Similar percentages for nodal cases were observed for both HIV negative and HIV positive DLBCL cases (Table 2.1). In addition, no differences in median OS between nodal and extranodal lymphoma cases were observed for both HIV negative (9.1 years vs. 6.3 years) and HIV positive cases (5.0 years vs. 2.7 years) (Appendix B). The most common extranodal sites were the GI tract (40/122), bone (23/122) and the respiratory system (15/122) (Figure 2.1). There were no differences in OS between the three most common exrtanodal sites in the HIV negative group (Figure 2.4). The 10 year OS rates were 65.45% for GI tract, 50.22% for respiratory system and 23.98% for skeletal site

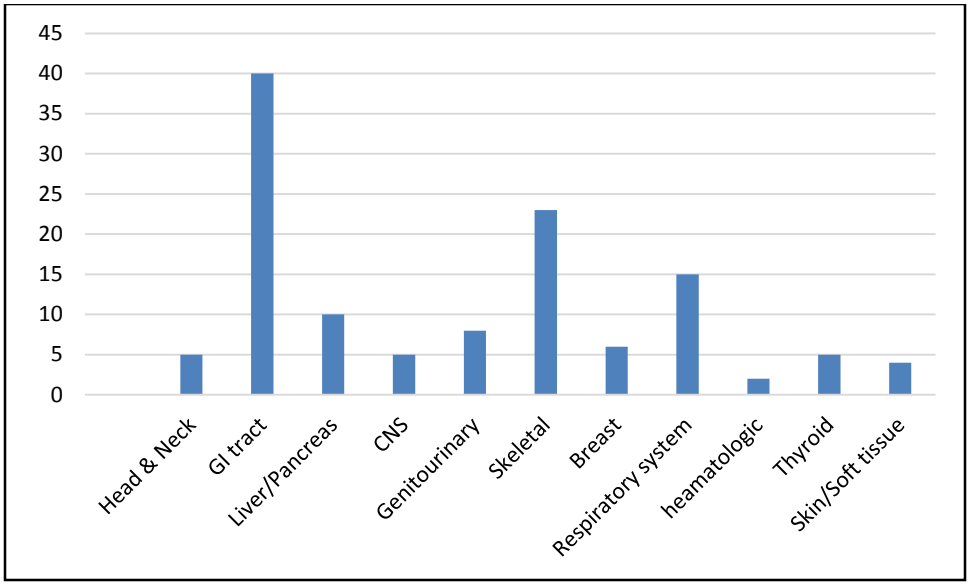


Figure 2.1: The frequency of the various DLBCL extranodal sites. The GI tract and skeletal sites were the most frequently observed.

2.3.2 Statistical analysis of Oncology data

Differences between the clinicopathological parameters of HIV negative and HIV positive DLBCL were compared using chi square and t-test. No statistically significant differences in gender, Ann Arbor stage, performance status, B symptoms, primary tumour site and bone marrow involvement were observed (Table 2.1). The only statistically significant differences observed were in median age and elevated LDH levels. The HIV positive group had a higher percentage of cases with elevated LDH

levels (higher than 480 U/L) compared to the HIV negative group (93% vs. 63%, $p < 0.05$).

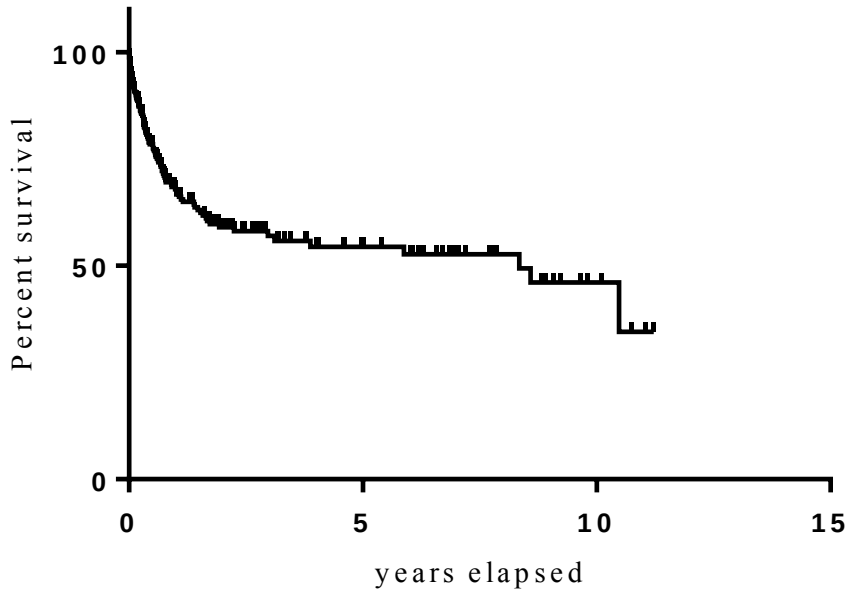


Figure 2.2: Kaplan Meier survival curve of DLBCL cases treated at GSH

2.3.3Treatment outcome of DLBCL patients

The median survival of all DLBCL patients was 8.6 years (Figure 2.2) with the patients receiving CHOP treatment having a median of 2.24 yrs vs. 8.34 yrs for HIV positive and HIV negative patients, respectively (Figure 2.3). The five year OS was 56% vs. 46% for

HIV negative vs. HIV positive DLBCL patients (p = 0.048) (Figure 2.3). The HIV negative patients had higher complete response (CR) rates compared to HIV positive patients (47% vs. 40%) (Table 2.2). The relapse rate in the HIV negative group was higher than the rate of the HIV positive group (10% vs. 8%) (Table 2.2).

Table 2.2: Complete response and relapse rates of the DLBCL patients treated at GSH.

DLBCL		**CR**	**Relapse (after CR)**	**No CR**	**Total**
HIV neg	N	95	20	86	201
	%	47.26	9.95	42.79	100
HIV pos	N	21	4	27	52
	%	40.38	7.7	51.92	100
	Total	116	24	113	253
	%	45.85	9.49	44.66	100

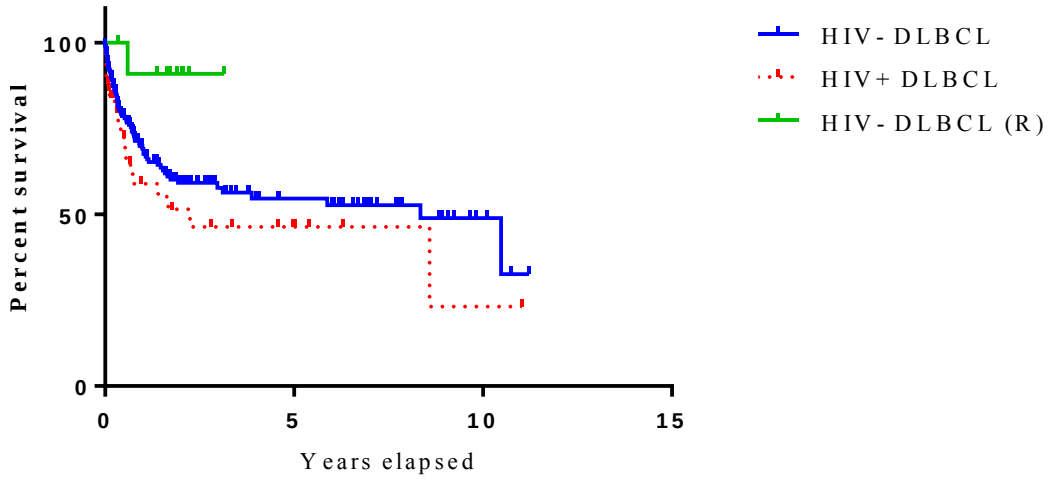


Figure 2.3: Kaplan Meier survival curve of HIV infected (HIV +) and HIV uninfected (HIV-) DLBCL cases

Table 2.3: The surviving proportions of DLBCL patients at 3, 5 and 10 year intervals.

Years elapsed	HIV-DLBCL	HIV+DLBCL	HIV- DLBCL (R)
3	58.40%	46.30%	90.91%
5	56.60%	46.30%	90.91%
10	50.70%	23.10%	90.91%

P value = 0.0482

2.3.4 Prognostic factors affecting overall survival

Linear regression was performed to determine the clinicopathological prognostic indicators affecting patient outcome. The clinicopathological parameters were given codes for the analysis (Appendix A). The analysis was performed on all the DLBCL

patients irrespective of HIV status prior to analysing the two cohorts of the study. Advanced age and stage were associated with poor prognosis for all DLBCL patients (Table 2.4). In addition, bone marrow involvement and female patients were associated with poor prognosis for all DLBCL patients (Table 2.4). The same factors were significant in HIV negative patients with the inclusion of primary extranodal tumour site, which was a poor prognostic indicator (Table 2.4).

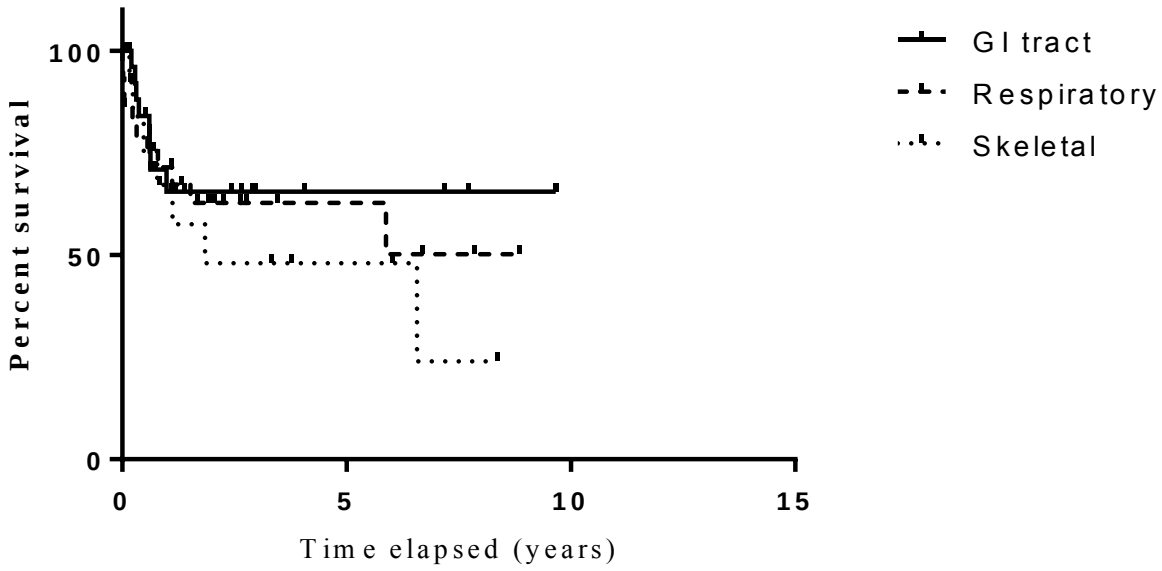


Figure 2.4: Kaplan Meier graph comparing the OS of extranodal lymphoma patients presenting in GI tract, respiratory and skeletal sites in the HIV negative cohort.

Table 2.4: The linear regression analysis of prognostic factors affecting overall survival. The HIV infected cohort did not have significant prognostic indicators.

Factor	ALL DLBCL patients			HIV – DLBCL patients			HIV + DLBCL patients		
	B	S.E of B	p value	B	S.E of B	p value	B	S.E of B	p-value
Age	0.15	0.0615	0.02	0.17	0.0691	0.0014	0.11	0.14	0.44
Gender	-0.15	0.0667	0.03	-0.20	0.0765	0.0098	0.01	0.14	0.95
Stage	0.17	0.0632	p<0.001	0.16	0.0711	0.03	0.2	0.14	0.18
BMi	-0.24	0.0690	p<0.001	-0.23	0.0789	0.004	-0.26	0.14	0.07
EN lymphoma	0.08	0.06	0.23	-0.14	0.0707	0.04	0.13	0.14	0.36
IPI	-0.07	0.06	0.49	-0.07	0.11	0.54	N.D	N.D	N.D
LDH	-0.04	0.07	0.54	-0.04	0.07	0.54	0.003	0.15	0.98
Presence of B-symptoms	0.02	0.07	0.79	0.01	0.1	0.94	-0.02	0.14	0.86
HIV status	0.04	0.06	0.49	N.D	N.D	N.D	N.D	N.D	N.D
CD4 count	-0.09	0.14	0.53	N.D	N.D	N.D	0.06	0.14	0.70

B, regression coefficient; SE, standard error; N.D, Not determined

2.4 DISCUSSION

The study determined the clinicopathological differences between HIV negative and HIV positive DLBCL cases. The DLBCL cases diagnosed and treated at GSH were located by using a NHLS list. The number of patients treated at GSH was lower than expected due to a number of reasons. The cases diagnosed by the NHLS Anatomical Pathology laboratory at Groote Schuur hospital are not all treated at the GSH Department of Radiation Oncology. In addition, the SNOMED code used to create the list is not specific for DLBCL and includes large T cell lymphomas. The SNOMED system used is not aligned to the current WHO classification of lymphomas. We observed a median age for DLBCL cases occurring at the early sixth decade (Table 2.1), thus confirming previous studies (Bürgesser et al. 2013; Flowers et al. 2013; Hunt & Reichard 2008; Castillo et al. 2014). The lower median age at diagnosis for HIV positive DLBCL patients has been previously reported and ranged in the fourth decade in South Africa (Wiggill et al. 2011; de Witt et al. 2014) and fifth decade abroad (M. J. Baptista et al. 2015).

Our study observed a female predominance, in both HIV negative and positive groups which is contrary to the male predominance previously reported (Sarkozy et al. 2014). This may be due to the 44 cases with no recorded gender which could have skewed the data towards a female predominance. However, possible reasons for gender differences in cancer have been elucidated for other cancers, yet this has not been done for lymphoma (Mester et al. 2006). In addition, the increased number of females in the HIV positive group may be due to relatively larger numbers of females that are infected with HIV in South Africa (UNAIDS 2015; Shisana et al. 2012). This is in keeping with a

similar study conducted previously in South Africa showing an overrepresentation of females (Wiggill et al. 2011).

We did not observe any statistical difference in IPI, stage, PS and frequency of B symptoms between the HIV negative and positive DLBCL groups (Table 2.1). This is in contrast with a study done in a Spanish population (M. J. Baptista et al. 2015) which showed that HIV infected patients were more likely to have worse prognostic indicators (higher PS scores, frequent B symptoms and a higher stage). The differences observed may be due to poor socio-economic status impacting on patient survival (Tao et al. 2014). Our study observed that the HIV positive group frequently had higher LDH levels which confirm other studies reporting the difference in LDH levels (Table 2.1) (J. M. Baptista et al. 2015; Williams & Churchill 2007; Lu et al. 2013). The LDH enzyme catalyses the conversion of pyruvate to lactate and is considered to be a key enzyme of glycolysis. It is elevated in many types of cancers and has been linked to tumour growth, maintenance, and invasion (Miao et al. 2013). A disturbed cellular metabolism due to cancer and infectious conditions may trigger release of high levels of LDH into the blood stream (Ramana et al. 2013). In addition, HAART has also been shown to result in elevated LDH in HIV positive individuals due to their role in the inflammatory response (Blanco et al.). Therefore this parameter may be frequently elevated in the HIV positive group due to HIV infection as well as HAART intake.

The HIV positive DLBCL group had significantly lower overall survival, thus affirming published data (Figure 2.3) (J. M. Baptista et al. 2015; de Witt et al. 2014). However, the 3 year OS rate for HIV positive DLBCL in our study was higher than the one published by de Witt (2013) despite study population emanating from the same region. The reason

for the differences observed may be the varying treatment regimens received in our patients, whereas his cohort received CHOP only. The CR rates were higher in the HIV negative cohort confirming previously published data (J. M. Baptista et al. 2015). Although a higher relapse rate in the HIV positive group was expected, it was interestingly higher in the HIV negative group (Table 2.2). This could have resulted from a higher death rate in the HIV positive group, leading to a low percentage of the surviving population which may have low relapse rate (Figure 2.3, Table 2.3).

We did not observe any statistically significant differences in the distribution of nodal and extranodal cases between the two DLBCL groups (Table 2.1). This is different from the study by Pather et al (2013) who found HIV positive cases to be more likely to have extranodal tumours. The contradiction may be due to differences in sample size as well as the inclusion of plasmablastic lymphoma cases in their analysis. In addition, differences in the classification systems used to classify nodal and extranodal lymphomas may also result in contradictions. For example, some studies include the Waldeyer's ring in nodal, while others exclude this (Castillo et al. 2014; López-Guillermo et al. 2005). This may lead to discrepancies in the frequency of nodal and extranodal cases. Our study confirms previous data suggesting that the GI tract is the most common extranodal site (López-Guillermo et al. 2005). In addition our study indicated that extranodal tumour site was associated with worse outcome in the HIV negative group, and this was also seen in previous studies (Castillo et al. 2014). Our results indicated that skeletal/bone extranodal sites have a lower median survival in the HIV negative group although this was not significant (Figure 2.4). We could not however, determine the effect that each extranodal site has on outcome due to low sample size on most groups, especially in the HIV positive group. However, previous studies indicate that the GI tract and liver/pancreas

demonstrated the worst outcome (Castillo et al. 2014). A larger cohort made up of multiple centres may assist in determining the prognostic effect of each extranodal site in the Southern African population. The prognostic effect of extranodal tumours may lie in the genetic and molecular differences inherent between nodal and extranodal tumours (Al-humood et al. 2011; López-Guillermo et al. 2005). Extranodal lymphomas have frequent gains of chromosomal arms 1p, 7p, 12q24.21-12q24.31 and loss of chromosome 4, 6q, and 18q22.3-23 while nodal lymphomas frequent 18q amplification and BCL2 protein overexpression (Al-humood et al. 2011).

Our finding regarding age, stage, bone and marrow involvement indicated that they were prognostic indicators for all DLBCL patients, including the HIV negative cases (Table 2.2). The study by Klapper et al (2012) showed that increased age was associated with poor prognosis. In addition, age, stage and extranodal sites are part of the International Prognostic Index (Shipp et al. 1993) highlighting their importance as prognostic markers. Bone marrow involvement has been previously identified to negatively affect overall survival (Shim et al. 2013). Our study reports that the female gender is associated with worse prognosis. This result is contrary to the studies performed in other parts of the world which indicate that males have worse prognosis (Yıldırım et al. 2015; Sarkozy et al. 2014). The poor prognosis in females in our cohort may be due to the increased risk of anthracycline therapy induced cardiotoxic side effects in females (Volkova & Russell 2011).

The study was limited by the lack of proper record keeping in the patient folders. This resulted in inability to determine the differences in IPI score distribution between the two groups. The lack of alternative treatment options for DLBCL and workload of oncologists may affect the ability to accurately record prognostic factors for each patient.

In addition, the study population was not uniform since mediastinal DLBCL and T cell rich large B cell lymphoma subtypes were included. These are considered as distinct entities by WHO and may affect the results (Jaffe 2009). The patients were also not given uniform treatment as mentioned in section 2.2.4 which may mask the effect of CHOP therapy on DLBCL subtypes. Moreover, some of the recent cases (2012-2013) belonging to the HIV negative group were given R-CHOP and may have had a better prognosis compared to their HIV positive counterparts (Rovira et al. 2015; Cultrera & Dalia 2012). Interestingly, R-CHOP has been reported to not offer survival improvement in Southern Africa, so it is possible that the addition of rituximab did not influence patient survival (Sissolak et al. 2013). Finally, given adequate financial resources, the study would have benefitted from comparisons of patients who belonged to the same molecular subtype as well as given similar treatment type.

In summary, our study observed that HIV negative and HIV positive DLBCL patients have similar clinicopathological features but, they differ in median age at diagnosis and frequency of elevated LDH levels. The HIV positive DLBCL group was also shown to have poorer overall survival. Finally, age, gender, stage and bone marrow involvement were calculated to be prognostic indicators for all DLBCL patients irrespective of HIV status. In addition, the presence of lymphoma in extranodal sites was also prognostic for the HIV negative group.

CHAPTER 3: BIOMARKER IDENTIFICATION USING MALDI IMAGING MASS SPECTROMETRY (IMS) AND LC-MS

3.1 INTRODUCTION

MALDI IMS is a powerful technology which investigates the content and distribution of molecules in tissues while preserving their morphological structure (Balluff et al. 2011). Therefore, it has great potential in the field of biomarker discovery as proteins, lipids and metabolites can be analysed without prior knowledge of their presence or the need for labelling (Caprioli et al. 1997; Balluff et al. 2011). MALDI IMS has various applications including: molecular classification of tissue, analysis of intra-tumour diversity and the determination of drug metabolism kinetics (Castellino et al. 2011; Schöne et al. 2013). Molecular classification of tissues can lead to the identification of markers used for patient diagnosis (Rauser et al. 2010), prognosis (Balluff et al. 2015), and response to therapy (Reyzer & Caprioli 2005).

The previous chapters highlighted the need for biomarkers for DLBCL, especially in the HIV context. Below, we describe the identification of biomarkers for HIV related and non-HIV related DLBCL, using MALDI IMS.

3.2 MATERIALS AND METHODS

3.2.1 Study samples

The study consisted of 12 FFPE tissue samples including four reactive nodes and eight DLBCL samples. There were four DLBCL samples from HIV negative samples and four from HIV positive samples. Similarly, two reactive nodes were taken from HIV negative and HIV positive samples, respectively. The baseline demographics of the study cohorts can be seen in Appendix G. In addition, 52 samples were used to validate the expression of Hsp70 using immunohistochemistry (Appendix J).

3.2.2 Tissue preparation

FFPE tissue blocks were sectioned on a rotary microtome at 10 µm. Tissue sections were picked up poly-L-lysine coated (Appendix E) indium tin oxide (ITO) glass slides and baked on a hot-plate at 60 °C for 10 min. The slides were then brought down through three xylene washes of 5 min each to enable dewaxing. This was followed by washing in three successive 100% ethanol baths to facilitate clearing of tissues. Sections were then washed well in distilled water for 5 min and left to air dry at room temperature (RT) prior to processing.

3.2.3 Hematoxylin & Eosin (H&E) Stain

Tissue sections were cut at 3µm and prepared as mentioned above before incubating in Mayers Hematoxylin solution (Appendix F) for 5 min at RT. This was followed by

washing and bluing of nuclei with ammoniated water (Appendix F). The cytoplasm of the tissues was stained with phloxine eosine solution (Appendix F) for 2 min. The sections were then dehydrated with alcohol baths, cleared with xylene baths prior to mounting on cover slips. The H&E stained slides were then viewed by a pathologist and the tumour region demarcated.

3.2.4 Methodology for MALDI IMS

3.2.4.1 Tissue digestion for MALDI IMS

The tissue sections were prepared as described in section 3.2.2andwere left to dry at RT. The trypsin solution (5 ng/µl in 50% acetonitrile; 50 mM ammonium bicarbonate, pH 8) was automatically spotted on the tissue using the trypsin deposition settings on ImagePrep (Bruker). The slide was then incubated at 100% humidity at 37 °C for 15 hr to facilitate digestion.

3.2.4.2 MALDI Imaging mass spectrometry of digested FFPE tissues

Following on tissue digestion, the α-cyano-4-hydroxycinnamic acid (HCCA) matrix (7 g/l in 60% acetonitrile, 0.2% trifluoroacetic acid) was deposited onto sections using the HCCA settings on Imageprep. External calibration was performed using Bruker peptide calibration standard II. The Imaging analysis of the tissue was carried out by rastering at 100 µm by 100 µm pixel sizes. The spectra were acquired by firing a 337 nm laser (20

μm laser spot size) at a rate of 50 shots per raster spot. A mass spectrum was acquired for each set of spatial coordinates based on 2500 shots per pixel area accumulated in a random walk mode. The spectra were collected in positive reflectron mode in the *m/z*700 -3500 range. Image acquisition was carried out using the FlexControl 3.0 software package (Bruker). MALDI parameters were set as follows: 25 kV-22.45 kV acceleration voltage, 8.00 kV lens voltage, 26.71 kV reflector voltage and 13.35 kV final acceleration, with up to 400 Da suppression.

3.2.4.3 IMS data analysis

3.2.4.3.1 Data exportation

Regions of Interests (ROIs) in the tumour region of tissue sections were drawn on the scanned tissue section image on FlexImaging 3.0 software (Bruker). The follicular regions in the control tissue sections were chosen as ROI (Appendix C). For each sample, the spectra within the ROIs were exported as xml files. Statistical analyses were performed on pooled spectral data from each sample group. Fifty randomly chosen spectra from each ROIs were selected using Spectral Importer software (Bruker) and comparison groups were created and given a colour. The spectra within the groups were then recalibrated and normalized prior to statistical analyses on Clinprotools 3.0 (Bruker).

3.2.4.3.2 Spectral recalibration and normalisation

Recalibration of the spectra on Clinprotools 3.0 (Bruker) was performed at a resolution of a 1000, with a mass range between *m/z*100 -5000 and a data reduction of two. The background from the spectra was subtracted using the TopHat baseline correction and normalised using the total ion count (TIC) method. The selection of statistically relevant peaks was done on the total average spectra selecting peaks with a signal to noise ratio of at least 20 (Appendix Q).

3.2.4.3.3 Principal Component Analysis (PCA) and Hierarchical Clustering

PCA was performed on recalibrated spectra with a 'level' scaling method keeping 95% variance of the data. Hierarchical clustering was performed on PCA data with Euclidean distances and the ward linkage method used to measure the distances between two points. The data was reduced to 95% of the explained variances and the tree drawn with a maximum of 5 path lengths.

3.2.5 Protein extraction for LC-MS/MS

The same tissue sections that were used for MALDI IMS were washed in 70% ethanol to remove the matrix. They were then subjected to *in situ* trypsin digestion using the settings stated in section 3.2.4.1. The peptides were extracted with, 20 µl of 10% acetonitrile (ACN) was pipette up and down the tissue section for 1 min. The extract was then concentrated using a speedvac for 1 hr. A 2 µl aliquote was then mixed with 10 µl 0.1%

trifluoroacetic acid (TFA) and taken through the zip tip method (Appendix E). This was then subjected to nLC-MS/MS analysis.

3.2.6 Protein Identification

3.2.6.1 nLC-MS separation of peptide extract

The peptide extract was first separated usingliquidchromatographyon a Thermo Scientific EASY-nLC II connected to a Proteineer fc II protein spotter controlled through HyStar software. The separation was performed on an EASY column (2cm, 75µm ID, 5µm, C18) pre-column followed by an analytical column (10cm, 75µm ID, 3µm, C18) with a flow rate of 100µl/hr using a 48 minute gradient run (Table 3.1). The eluent were spotted on a MALDI plate (MTP 384 target plate) with a saturated HCCA matrix every 15s. The plate was further processed with MALDI-TOF MS.

3.2.6.2 MALDI-TOF of spotted peptides

MALDI-TOF MS and LIFT MS/MS was performed using an UltrafleXtreme MALDI ToF/ToF system (Bruker Daltonics, Bremen, Germany) with instrument control through Flex control 3.4. Peptides were ionized with a 337 nm laser and spectra acquired in reflector positive mode at 28kV using 100 laser shots per spectrum with a scan range of *m/z* =700 -4000. Spectra were internally calibrated using peptide calibration standard II(Bruker). Peptide spectra of accumulated 4,000 shots were automatically processed using WARP LC 3.2 software (Bruker)

Table 3.1: The gradient run for the nLC-MS analysis

Time (min)	**Function**	**Value**
0	Flow rate	300nl/min
0	Solvent Mix	98% A , 2% B
44	Solvent Mix	65% A , 35% B
48	Solvent Mix	60% A , 40% B
48.10	Solvent Mix	0% A , 100% B
60	Solvent Mix	0% A , 100% B
60.10	Solvent Mix	98% A , 2% B
70	Solvent Mix	98% A , 2% B

A: 0.05%TFA/H2O, B: 0.05%TFA/ACN

3.2.6.3 Data analysis

The results were exported with the Mascot algorithm using the SwissProt database on a ProteinScape 3.0 workstation. The search parameters were as follows: Taxonomy- Homo Sapiens and Viruses; Enzyme-trypsin; Missed cleavages-1; Fixed modification- carbamidomethyl (C); Variable modification- oxidation (M); Precursor tolerance-50 ppm; Fragment tolerance-0.7 Da. The identities of the m/z ions detected in MALDI IMS were matched with the peptides from nLC-MS/MS using peptide mass (*m/z*).

3.2.7 Panther and network analysis

The PANTHER (Protein Analysis THrough Evolutionary Relationships) (http://pantherdb.org/) gene ontology classification system was used to annotate and cluster the proteins according to molecular function and pathways. The retrieval of interacting genes/proteins (STRING) database (http://string-db.org/) was used to determine the physical and functional interactions among the identified proteins. STRING defines a metric called "confidence score" to define interaction confidence. Interactions with a highest confidence score of≥ 0.9 were viewed.

3.2.8 Immunohistochemistry of p24 and Heat Shock 70 Proteins

The detection of p24 was performed on the MALDI IMS samples while heat shock (Hsp) 70 protein was also detected in a separate cohort of DLBCL samples (Appendix J).

3.2.8.1 Antigen retrieval step

FFPE tissue sections placed on coated slides were dewaxed and cleared prior to boiling in appropriate buffer (Table 3.2) in a pressure cooker for 90 sec. They were then left to stand in the buffer for a further 20 min to allow for cooling before washing with distilled water.

3.2.8.2 Immunohistochemistry

Following antigen retrieval, the slides were then washed with distilled water prior to incubation with 3% hydrogen peroxide for 10 min. This was followed by washing in

PBS-Tween (0.02 % v/v) for 10 min before blocking with goat serum (5% in PBS) for 15 min. The tissue slides were then incubated with primary antibody (Table 3.2). The primary antibody was then washed in PBS-T and the tissue incubated with the Envision anti- mouse HRP labelled polymer for 30 min. This was then followed by incubation for 10 min with DAB to stain the antibody- antigen complex. The tissue slides were then counterstained with Mayers Haematoxylin, and mounted with DPX.

Table 3.2: The primary antibody information for p24 and Hsp70

Antigen retrieval buffer	**Antigen**	**Clone**	**Supplier**	**Antibody dilution**	**Control tissue**
1mM Tris and 1mM EDTA, pH 9	p24	Kal1	Dako	1:20, overnight	HIV positive reactive lymph node
0.01M Citric acid, pH6	Hsp 70	4872	Cell Signalling	1:60, overnight	Normal colon and colon cancer tissue

3.2.8.3 Assessment of immunohistochemical stains

Immunopositivity of p24 in lymphoid tissue was considered when mononuclear cells, macrophages or the meshwork of follicular dendritic cells were staining (Moonim et al. 2010). Immunopositivity of Hsp70in tumour cells was evaluated by determining the proportion of cells staining for Hsp70 as well as the intensity of the stain. Staining intensity was graded as 0 (negative), 1 (weak), 2 (moderate), and 3 (strong); percentage of positive cells examined was scored as 0 (negative), 1 (< 10%), 2 (10-25%), 3 (25-50%), and 4 (>50%).

3.3 RESULTS

3.3.1 MALDI IMS analysis results

The representative protein profile of control lymph node and DLBCL were characterised using the histology-directed MALDI IMS approach. Mass spectra from the ROIs in follicular regions of reactive nodes as well as tumour sections of DLBCL (Appendix C) were extracted and statistically compared using Clinprotools software. There were thirty five different mass ions that were differentially expressed between reactive node spectra and DLBCL as well as between DLBCL subtypes. However, only twenty five peptides were successfully identified using nLC-MS/MS (Appendix K).

3.3.1.1 Proteins differentiating between reactive nodes and DLBCL in the HIV negative cohort

The peak picking function identified twenty peptides in the range of *m/z* 700- 2000 that were differentially expressed between the reactive nodes and DLBCL cases (Table 3.3). Fifteen (15/20) of these peptides were successfully identified using nLC/MS (Table 3.3). Nine peptides were upregulated in DLBCL cases including ribosomal protein L40 (*m/z* 1039), annexin A5 (*m/z* 1105) and histone H2A (*m/z* 944). The majority of the identified proteins were downregulated in DLBCL samples including enolase (*m/z* 704), glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (*m/z* 805) and histone H3 (*m/z* 1032).

GAPDH *(m/z* 805) was found to be highly expressed throughout the reactive node but had moderate expression in the tumour region of HIV negative DLBCL, with some

regions having little or no expression of GAPDH (Figure 3.1). In contrast, in HIV positive samples, GAPDH was highly expressed throughout the reactive node and tumour samples (Appendix O). Histone H3 (*m/z* 1032) was highly expressed only in the follicular region and had very little expression throughout the tumour region of HIV negative samples (Figure 3.1). Some expression of histone H3 was also observed in the venules of the fatty layer (Figure 3.1). Similar results were observed for the HIV positive samples (Appendix O).

Table 3.3: Proteins differentially expressed between reactive nodes (RN) and DLBCL tissues in the HIV negative cohort.

Ion mass	Regulation pattern	Identity	Common name
703.5155897	Down		
704.5124537	Down	ENOA_HUMAN	Enolase
713.5527754	Up	H4_HUMAN	Histone H4
788.6839193	Down	H31T_HUMAN	Histone H3
789.6806242	Down	RS16_HUMAN	Ribosomal protein S16
801.6711308	Up		
805.6951772	Down	G3P_HUMAN	GAPDH
849.6560995	Up		
861.2877784	Up		
927.7955744	Down	ALBU_HUMAN	Albumin
944.7990104	Up	H2A3_HUMAN	Histone H2A
945.7859354	Up	H2A3_HUMAN	Histone H2A
966.7704072	Up		
1032.899619	Down	H31T_HUMAN	Histone H3
1039.815191	Up	RL40_HUMAN	Ribosomal protein L40
1105.856544	Up	ANXA5_HUMAN	Annexin 5
1199.120245	Down	HSP7C_HUMAN	Heat shock protein 70
1287.158619	Down	IGHG1_HUMAN	Ig gamma-1 chain C region
1506.326079	Down	KPYM_HUMAN	Pyruvate kinase
1792.315581	Down	ACTG_HUMAN	Actin

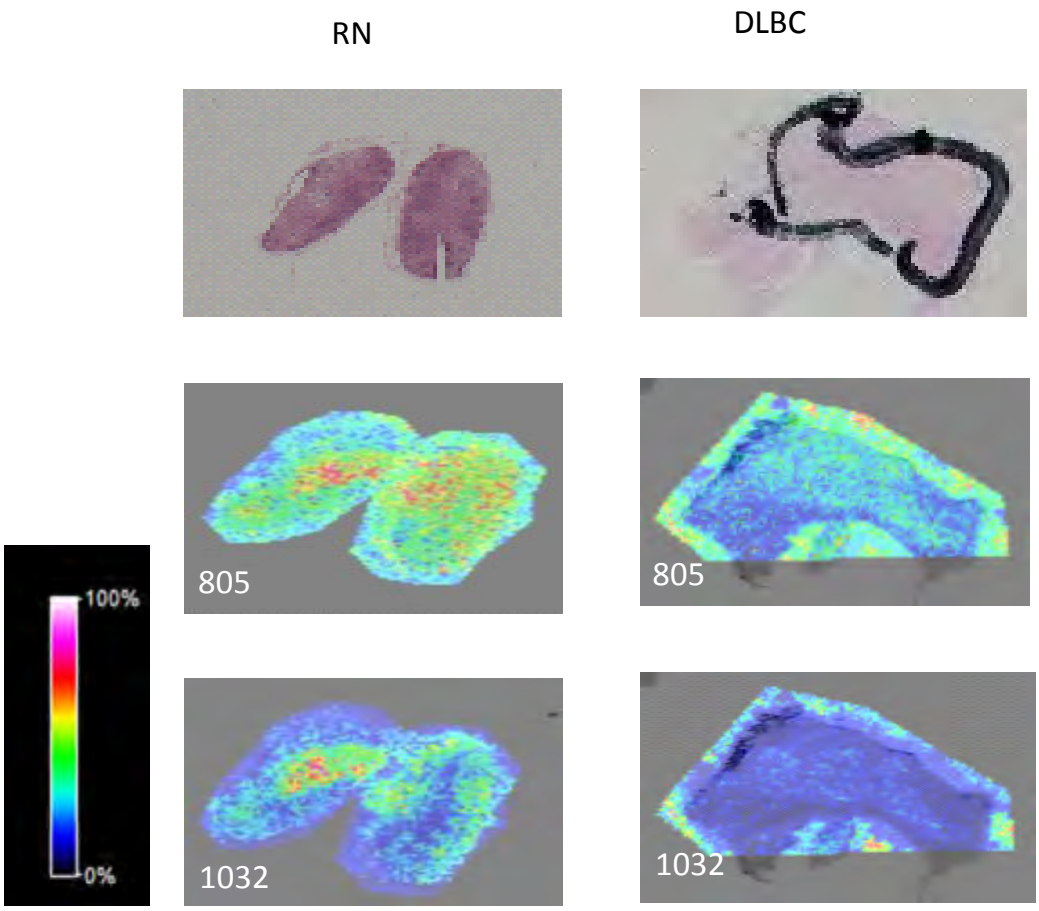


Figure 3.1: The distribution of selected peptides from representative samples of the HIV negative cohort. GAPDH (*m/z* 805) was highly expressed throughout the reactive node and moderately expressed in the tumour of DLBCL. Histone H3 (*m/z* 1032) was highly expressed in the germinal centres of the reactive nodes with low expression in DLBCL tumours.

3.3.1.2 Proteins differentiating between reactive nodes and DLBCL tissues in the HIV positive cohort

The peak picking function identified twenty five statistically significant peptides that were differentially expressed between reactive nodes and DLBCL from HIV positive patients (Table 3.4). The identity of twenty peptides was successfully determined using nLC-MS and matched to various proteins. Fourteen (14/25) of these peptides were upregulated and eleven were downregulated. Most of these proteins were also expressed in the HIV negative cohort (Table 3.3). There were new proteins which were not identified in the HIV negative context including collagen (*m/z* 848), albumin (*m/z* 927) and heterogeneous nuclear ribonuclear protein A2 (*m/z* 1378). There were some proteins which were regulated in a similar manner as in the HIV negative context including annexin A5 (*m/z* 1105) and histone H4 (*m/z* 713). However, other proteins were regulated differently to the HIV positive context. For example, constitutive heat shock protein 70 (*m/z* 1199) was downregulated in the HIV negative cohort but upregulated in the HIV positive cohort. Similarly, histone H3 (*m/z* 1032) was downregulated in the HIV negative context but upregulated in the HIV positive context (Table 3.4).

The distribution of ribosomal protein S16 (RPS16) (*m/z* 789) and heat shock protein (Hsp) 70 (*m/z* 1199) in HIV positive samples are shown in figure 3.2. RPS16 (*m/z* 789) is expressed throughout the reactive nodes, while it is expressed in a subset of tumour cells (Appendix O, Figure 3.2). In contrast in HIV negative samples, RPS16 was only expressed in follicular regions of reactive nodes. In addition, it was expressed in some tumour cells as well as the venules in the fatty layer (Appendix O). In contrast the ion *m/z*

1199 is expressed in the follicular regions only and diffusely expressed in the tumour region (Figure 3.2, Appendix O). Similar results were observed in the HIV negative group.

Table 3.4: Proteins differentially expressed between RN and DLBCL cases of the HIV positive cohort.

Ion Mass	Regulation pattern	Identity	Common name
788.727194	Up	H31T_HUMAN	Histone H3
789.721836	Up	RS16_HUMAN	Ribosomal protein S16
801.739062	Down		
805.739438	Down	G3P_HUMAN	GAPDH
806.729231	Down	ENOA_HUMAN	Enolase
816.723438	Down	H2B1K_HUMAN	Histone H2B
848.716095	Down	COA62_HUMAN	Collagen
849.719735	Down		
850.760654	Down	H2A3_HUMAN	Histone H2A
898.807199	Up	RL15_HUMAN	Ribosomal protein L15
927.841058	Down	ALBU_HUMAN	Albumin
928.805379	Up		
929.809243	Up		
944.860026	Up	H2A3_HUMAN	Histone H2A
945.84584	Up	H2A3_HUMAN	Histone H2A
957.859394	Up	RL13_HUMAN	Ribosomal protein L13
966.844867	Up		
976.809754	Down	ACTG_HUMAN	Actin
1032.93542	Up	H31T_HUMAN	Histone H3
1039.88407	Up	RL40_HUMAN	Ribosomal protein L40
1105.93637	Up	ANXA5_HUMAN	Annexin A5
1199.16965	Up	HSP7C_HUMAN	Heat shock protein 70
1350.14995	Down	RL15_HUMAN	Ribosomal protein L15
1378.17787	Up	ROA2_HUMAN	heterogeneous nuclear ribonucleoproteins
1792.39323	Down	ACTG_HUMAN	Actin

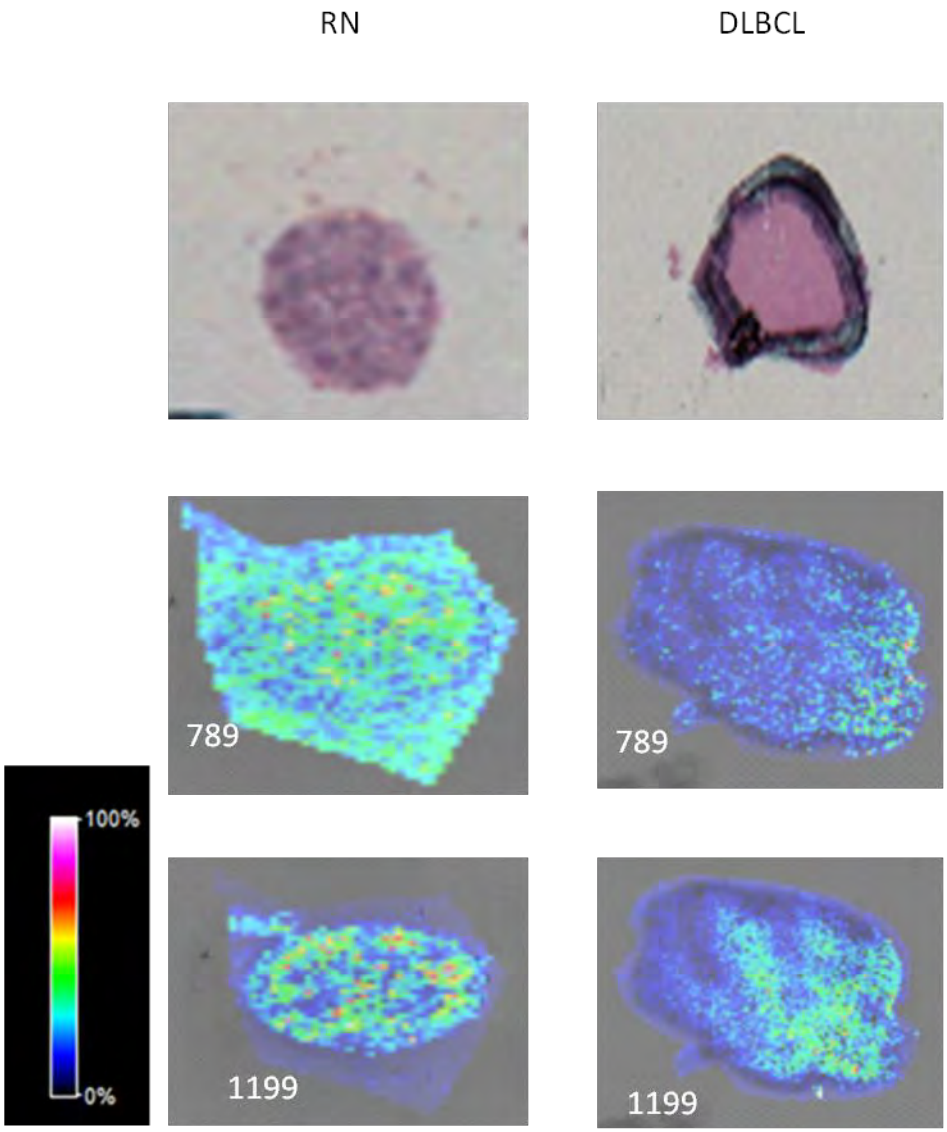


Figure 3.2: The distribution of selected ions in representative samples of the HIV positive cohort's tissues. Ribosomal protein S16 (*m/z* 789) and Hsp 70 (*m/z*1199) were highly expressed in germinal centres and a subset of tumour cells.

Differences between DLBCL from HIV negative and HIV positive cases were confirmed comparing spectra from the two groups. The peak picking algorithm generated seventeen peptides that were similar to the peptides identified in the previous comparison sets (Table 3.5).

Some peptides were still upregulated in the HIV positive DLBCL such as histone H3 (*m/z* 1032), collagen (*m/z* 848) and annexin A5 (*m/z* 1105), while there were other ions which had a change in the pattern of regulation when compared to HIV negative DLBCL samples. GAPDH (*m/z* 805) was downregulated in the HIV positive group vs. reactive nodes, but was upregulated in the HIV positive group vs. HIV negative group (Table 3.5). Similarly, the histone proteins H2A (*m/z* 945), H4 (*m/z* 713) were upregulated in both DLBCL groups vs. RN but were downregulated in the HIV positive DLBCL group vs. HIV negative group.

Table 3.5: Differentially expressed proteins between DLBCL tissues from HIV negative and HIV positive patients.

Ion Mass	Regulation pattern	Identity	Common name
713.5709	Down	H4_HUMAN	Histone H4
788.6688	Down	H31T_HUMAN	Histone H3
789.6591	Up	RSL16_HUMAN	Ribosomal protein L16
801.6639	Down		
805.6682	Up	G3P_HUMAN	GAPDH
816.6507	Down	H2B1K_HUMAN	Histone H2B
848.63	Down	CO6A2_HUMAN	Collagen
849.6382	Down		
850.675	Down	H2A3_HUMAN	Histone H2A
944.7813	Down	H2A3_HUMAN	Histone H2A
945.7534	Down	H2A3_HUMAN	Histone H2A
957.7765	Up	RL13_HUMAN	Ribosomal protein L13
966.7574	Down		
1032.847	Up	H31T_HUMAN	Histone H3
1039.79	Down	RL40_HUMAN	Ribosomal protein L40
1105.809	Up	ANXA5_HUMAN	Annexin A5
1199.082	Up	HSP7C_HUMAN	Heat shock protein 70

3.3.1.3 Proteins differentiating between DLBCL subtypes

The presence of the two clustering groups in the PCA analyses of the DLBCL cohorts (Appendix H) prompted the comparison of the immunohistochemical subtypes of DLBCL. The cases were classified into GC and non-GC using the Hans classification algorithm (Appendix D). There were twenty seven differentially expressed peaks between GC and non-GC for the HIV negative cohort. Eighteen of these peaks were successfully identified which correspond to fifteen different proteins (Table 3.6). The majority of the proteins were upregulated in the non-GC subtype including histone proteins H4 (*m/z* 713), H3 (*m/z* 788), ribosomal proteins S16 (*m/z* 789) and L40 (*m/z* 1039). Downregulated proteins includedcollagen α1(I) (*m/z* 836), collagen α2 (VI) (*m/z* 848), annexin A5 (*m/z* 1105) and actin (*m/z* 1792). In addition, unique proteins were identified that were not differentially expressed in the other comparison sets including the ion *m/z* 848 belonging to the collagen alpha chain. There were also ions that were not differentially expressed in this group including albumin (*m/z* 927), heterogeneous ribonucleoprotein (*m/z* 1378) and enolase (*m/z* 704).

The GC and non-GC comparison set from the HIV positive cohort produced ten differentially expressed peaks. We identified nine of these peaks as belonging to seven different proteins (Table 3.7). The majority of these peptides were also upregulated in the non-germinal centre group, including GAPDH (*m/z* 805) and heat shock protein (*m/z*1199). Only annexin A5 (*m/z* 1105) was downregulated in the non-germinal centre group vs. the germinal centre in the HIV positive cohort.

Table 3.6: Proteins differentially expressed between germinal centre (GC) and non-germinal centre (non-GC) DLBCL tissues of the HIV negative cohort.

Ion Mass	Regulation pattern	Identity	Common name
713.5531	Up	H4_HUMAN	Histone H4
757.6075	Up	RL13_HUMAN	Ribosomal protein L13
788.6508	Up	H31T_HUMAN	Histone H3
789.641	Up	RS16_HUMAN	Ribosomal protein L13
801.6496	Up		
805.6532	Up	G3P_HUMAN	GAPDH
816.6384	Down	H2B1K_HUMAN	Histone H2B
836.6418	Down	CO1A1_HUMAN	Collagenα1(I)
845.3651	Up		
848.605	Down	CO6A2_HUMAN	Collagenα2 (VI)
849.6152	Down		
850.6619	Down	H2A3_HUMAN	Histone H2A
861.2948	Up		
898.6963	Down	RL15_HUMAN	Ribosomal protein L13
940.7182	Up		
944.7667	Up	H2A3_HUMAN	Histone H2A
945.7494	Up	H2A3_HUMAN	Histone H2A
957.7619	Up	RL13_HUMAN	Ribosomal protein L13
966.7403	Up		
971.7643	Up		
978.8001	Up		
988.7306	Up		
1032.823	Up	H31T_HUMAN	Histone H3
1039.766	Up	RL40_HUMAN	Ribosomal protein L13
1105.822	Down	ANXA5_HUMAN	Annexin A5
1199.041	Up	HSP7C_HUMAN	Heat shock protein 70
1792.199	Down	ACTG_HUMAN	actin

Table 3.7: Proteins differentially expressed between GC and non-GC tissues of the HIV positive cohort.

Ion Mass	Regulation pattern	Identity	Common name
788.71	Up	H31T_HUMAN	Histone H3
789.71	Up	RS16_HUMAN	Ribosomal protein S16
805.72	Up	G3P_HUMAN	GAPDH
928.78	Up		
944.797	Up	H2A3_HUMAN	Histone H2A
945.78	Up	H2A3_HUMAN	Histone H2A
1032.93	Up	H31T_HUMAN	Histone H3
1039.35	Up	RL40_HUMAN	Ribosomal protein L40
1105.56	Down	ANXA5_HUMAN	Annexin A5
1199.16	Up	HSP7C_HUMAN	Heat shock protein 70

3.3.2 Pathway analysis of differentially expressed peaks

STRING predicts protein-protein interactions using information gathered from experimental and computational sources. PANTHER is a large curated database of gene/protein families and their functionally related subfamilies that can be used to classify gene/protein lists. The differentially expressed peaks found in the study were subjected to further analysis using STRING and PANTHER software in order to further understand the identified proteins. The majority of the proteins belonged to pathways such as glycolysis, Huntington's disease, inflammatory pathways mediated by the chemokine and cytokine signalling pathways (Figure 3.3). In addition, the proteins were found to have a binding function and were involved in the structural integrity of the cell

(Figure 3.4). STRING analysis of the differentially expressed proteins identified fourteen interactions that formed four clusters (p value = 1.285 e-5). The first cluster included the glycolysis pathway proteins consisting of enolase, GAPDH and pyruvate kinase. Enolase and GAPDH were also reported to bind to each other (Figure 3.5). The second cluster was made up of the ribosomal proteins which are connected to ubiquitin protein by a binding, reaction and catalysis interaction. The ribosomal proteins bind and react with each other and are collectively involved in catalytic reactions. The third cluster consisted of the four histone proteins which bind and react with each other (Figure 3.5). The last cluster included the two collagen chains which react together to form collagen. Proteins such as annexin A5, albumin and actin did not show any relationship with the other protein clusters that were identified in the study (Figure 3.5).

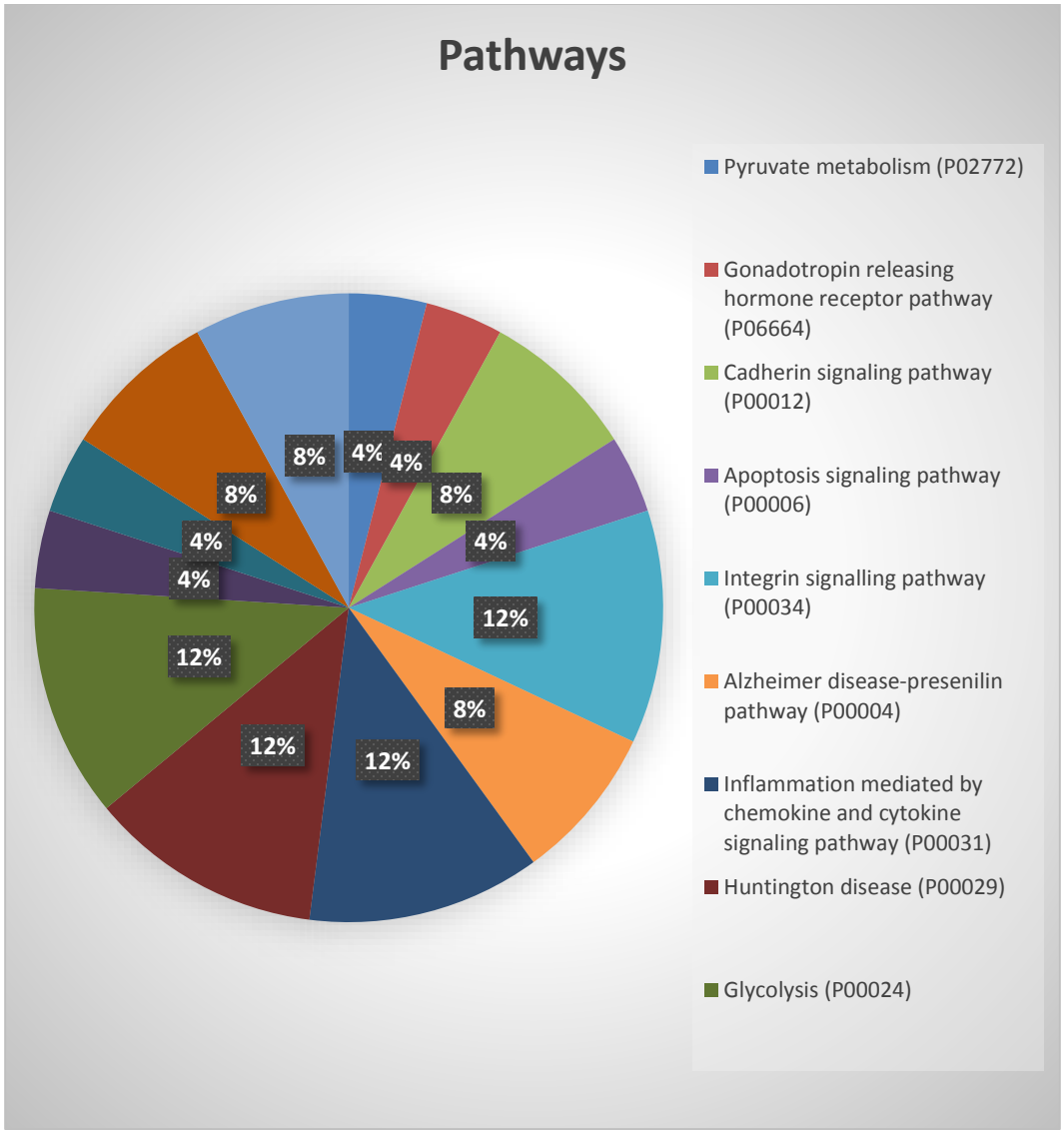


Figure 3.3: PANTHER analysis of the pathways associated with the differentially expressed proteins. A number of proteins participate in the glycolysis pathway.

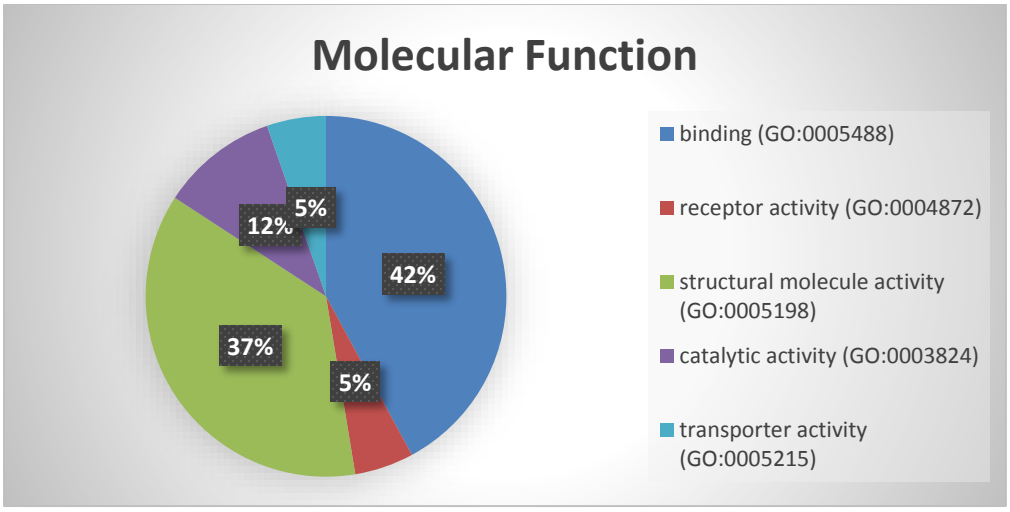


Figure 3.4: PANTHER analysis of the molecular function of the differentially expressed peptides/proteins. The majority of the proteins had a binding function.

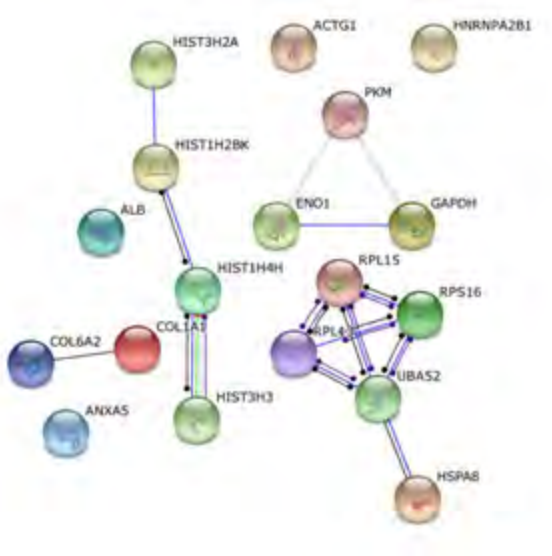


Figure 3.5: STRING analysis of all the differentially expressed proteins. The relationships between the proteins form four main clusters of glycolytic enzymes, ribosomal proteins, histones and collagen. Key: blue-binding, purple- catalysis, black-reaction, green- activation.

3.3.3 The detection of HIV in DLBCL FFPE tissues

HIV infection plays a role in the pathogenesis and progression of DLBCL tumours (Monroe and Silberstein 1995).The presence of HIV in DLBCL samples was detected using two approaches. The first approach involved searching viral databases for hits using the nLC-MS data from the HIV positive control and tumour samples. The viral database searches did not harvest any hits for HIV peptides/ proteins (Appendix L). The second approach involved the detection of the HIV p24 antigen using IHC (Appendix J). In addition, we observed negative p24 staining on the majority of control and tumour samples (Appendix J). Only one reactive node sample stained positive for p24, which was visible in the follicular dendritic network (Figure 3.6), however, no DLBCL sample showed any positivity for p24.

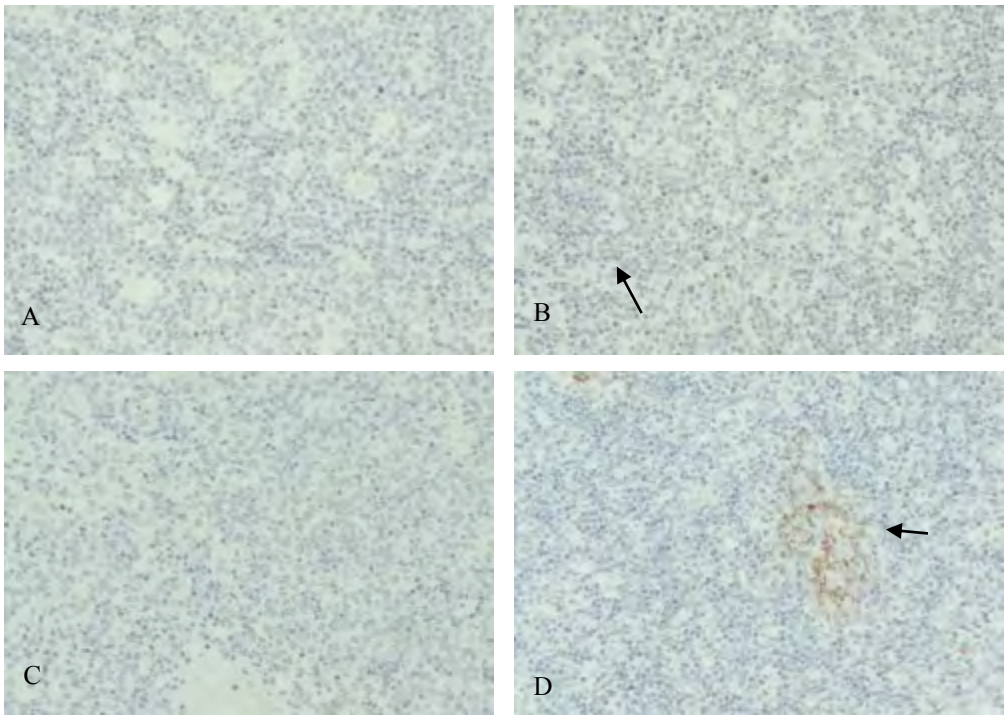


Figure 3.6: The immunohistochemical staining of p24 in the negative control (A) and monocuclear and macrophages in the positive control (B). p24 staining in the follicular dendritic network of a HIV positive lymph node (D). All DLBCL samples were negative for p24 (C). Magnification 5X (A, C) and 10X (B, D).

3.3.4 Confirmation of the expression pattern of heat shock protein 70

3.3.4.1 Expression of Hsp70 in DLBCL tissues

In order to validate the expression of proteins detected from the MALDI IMS data, IHC was performed to detect one of the thirty five proteins identified in the study (Hsp70). Hsp70 was chosen because it was one of the proteins which had a differential expression pattern across the comparison groups (Table 3.3 vs. Table 3.4). Hsp70 expression was detected in 52 samples including the ten samples used in the MALDI IMS analysis (Appendix J). Only eleven DLBCL samples were positive for Hsp70 expression which is evident by the cytoplasmic staining (Table 3.8). The majority (41/52) of the samples were negative for Hsp70 expression; however, some of them (23/41) expressed Hsp70 in plasma cells (Table 3.8, Figure 3.7). All the B-cells in the reactive nodes included in the confirmation cohort were also negative for Hsp70 expression, however, Hsp70 was expressed in the plasma cell (Figure 3.7, Table 3.8, Appendix J). The HIV positive DLBCL group had a higher percentage of cases expressing Hsp70 than the HIV negative DLBCL group (31.25% vs. 20.7%). In the HIV positive cohort, there were equal numbers of Hsp70 expressing cases for both GC and non-GC subtype (40% vs. 40%). In contrast, the non-GC subtype in the HIV negative cohort had higher percentage of cases expressing Hsp70 (66.6% vs. 16.6%).

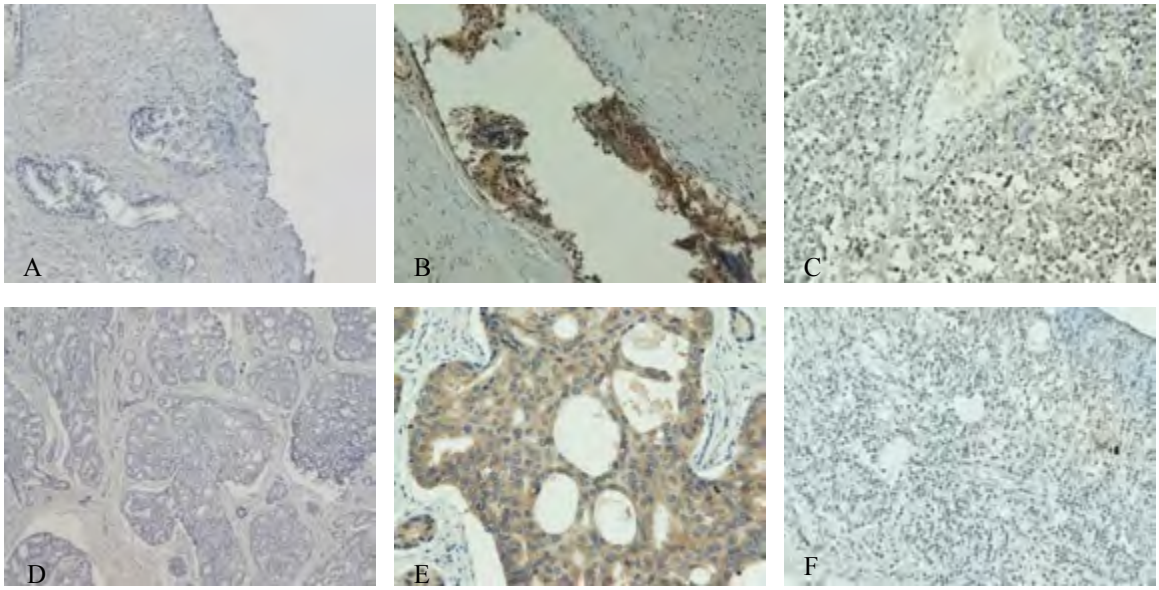


Figure 3.7: The control tissue used in the detection ofHsp70 expression was both normal gastric (A, B) and cancer tissue (D, E). DLBCL samples show Hsp70 staining in tumour cells (C) and/or plasma cells (F). Original magnifications 5X (A), 10X (B, C, D, F), 20X (E).

3.3.4.2 Correlation of Hsp70 expression with patient outcome

The effect of Hsp70 expression in DLBCL tissue samples was evaluated using a Fischer's exact test. Hsp70 expression significantly associated with outcome in HIV negative DLBCL cases ($p = 0.033$) (Table 3.8) with Hsp70 positive cases having a worse outcome/prognosis (Appendix N). There was no sign association between hsp70 and outcome in both HIV positive DLBCL cases and with the combined DLBCL groups ($P > 0.05$). In addition, there was no association betweenHsp70 expression in plasma cell and outcome (Appendix N).

Table 3.8: A summary of the Hsp70 IHC results. Hsp70 expression was correlated with outcome (alive or dead following treatment) in HIV negative DLBCL (DLBCL -) and HIV positive DLBCL cases (DLBCL +)

	Total (N)	Hsp pos	% Hsp pos	Hsp neg (plasma cell)	Correlation with outcome (P value)
DLBCL-	**29**	**6**	**20,68966**	22 (7)	0.033
GC	6	1	16,66667	5 (1)	
non-GC	13	4	66,66667	8 (4)	
Unknown subtype	10	1	16,66667	9 (2)	
RN	5	0	0	5 (3)	
DLBCL+	**16**	**5**	**31,25**	11 (2)	1
GC	5	2	40	3 (1)	
non-GC	7	2	40	5(1)	
Unknown subtype	4	1	20	3 (0)	
RN	2	0	0	2 (2)	
Total	52	11		41(23)	0.09

3.4 DISCUSSION

3.4.1 MALDI IMS analyses of differentially expressed proteins

The differentially expressed peptides in the MALDI IMS analyses were not all identified using nLC-MS due to several reasons. The identification of ions generated by MALDI IMS has been recently initiated and 100% identification has not yet been achieved because of technical differences between the MALDI and nLC- MS techniques (Djidja et al. 2010; Schober et al. 2011; Schey et al. 2013). The nLC-MS generates and identifies a higher number of ions compared to MALDI-MS due to chromatographic separation and a more efficient fragmentation (Gustafsson et al. 2012).

Further, the matches established from MALDI and nLC may not be accurate, since multiple peptides can be linked to one particular ion/ mass. However, some of the peptides identified such as those belonging to histone H2A (*m/z* 944), histone H3 (*m/z* 1032) and actin (*m/z* 1790) have exact matches have been previously observed in other studies, thus increasing our confidence in the matches (Djidja et al. 2009; Cole et al. 2014).

In our study, the majority of proteins that were identified belonged to pathways such as glycolysis and the integrin signalling pathway (Figure 3.4) and had a binding and structural activity function (Figure 3.3), which again confirms previous findings (Álvarez-chaver et al. 2014; Hwang et al. 2007). The glycolytic pathway enzymes included GAPDH, enolase and pyruvate kinase (Berg et al. 2002). GAPDH converts

glyceraldehyde-3-phosphate to1,3-diphosphoglycerate while enolase-α catalyses the dehydratation of 2-phosphoglycerate to phosphoenolpyruvate. The third glycolytic enzyme, pruvate kinase transfers a phosphate group from phosphoenolpyruvate to ADP, thus producing pyruvate (Diaz-Ramos et al. 2012). The glycolytic pathway proteins are frequently overexpressed in cancer due to the high energy requirements of the proliferating cancer cells (Tang 2012; Marín-Hernández et al. 2011). However, these proteins were shown to be downregulated in our tumour cohort compared to reactive nodes. This is probably due to the fact that the follicular region of reactive nodes have high rate of proliferating B cells which require more energy. In addition, DLBCL arises from the later stages of B cell development, where the proliferation rate of B cells may have decreased (Evans & Hancock 2003; Rickert 2013; Basso & Dalla-favera 2015). It is also possible that these proteins may not be playing a role in tumour-genesis.

The glycolytic enzymes have other functions not related to their roles in energy generation. Enolase functions as a plasminogen receptor, concentrating proteolytic plasmin activity on the cell surface of hematopoietic and endothelial cells, suggesting that it may play an important role in regeneration, immune response, and metastasis (Pancholi 2001; Diaz-Ramos et al. 2012; Kumari & Malla 2015). Downregulation of enolase in DLBCL compared to normal tissue (Table 3.3- Table 3.4) may also be a result of the cancer cells evading immune surveillance as high concentrations of enolase elicit an immune response (Kumari & Malla 2015).Furthermore, the downregulated in enolase we observed contradicts other cancer studies which show an upregulation of enolase which

result in invasion and also provides protection against therapy (Capello et al. 2011; Song et al. 2014; Hsiao et al. 2013).

GAPDH, another glycolytic enzyme is involved in several cellular processes such as, apoptosis and proliferation. Depending on whether it is expressed in the nucleus or cytoplasm, it may promote apoptosis or proliferation (Barbini et al. 2007). Overexpression in the nucleus promotes apoptosis whilst expression in both nuclear and cytoplasmic compartments promotes proliferation (Barbini et al. 2007). The upregulated GAPDH in HIV positive DLBCL may promote tumour-genesis, however this assertion requires further investigations with regard to the specific location of GAPDH expression (Table 3.3, Table 3.4). Our findings from the distribution of GAPDH (*m/z* 805) show an intense expression in germinal centres and moderate expression in DLBCL tumours (Figure 3.1). The heat map intensity profile generated from our study confirms previously published data generated by the protein atlas project. The protein atlas project showed intense staining of GAPDH in lymph node and low to moderate expression in NHL tumours which confirms our findings (Pontén et al. 2008; Uhlén et al. 2005).

A study by Chiche et al (2014) overexpression GAPDH is associated with aggressiveness and vascularisation of non-Hodgkin's lymphoma via a HIF-1 induction by NF-kB. This may be a possible reason for the overexpression of GAPDH in HIV positive DLBCL tumours compared to HIV negative tumours (Table 3.5), which are highly vascularised (Chiche et al. 2014; Liapis et al. 2013). GAPDH was also overexpressed in the non-

germinal centre subtype (Table 3.6, Table 3.7), which has a constitutively active NF-KB pathway, thus suggesting a mechanism for the aggressiveness of this subtype (Davis et al. 2001).

Moreover, GAPDH has a protective role against HIV infection and this is perhaps why there is increased expression in HIV positive tumour cells. Kishimoto et al (2012) observed that viruses produced from cells overexpressing GAPDH showed decreased infectivity and reverse transcription efficiency (Kishimoto et al. 2012).

STRING analyses from our study suggest that enolase and GAPDH function by binding to each other (Figure 3.5). A few studies observed the binding of enolase and GAPDH regulates host-pathogen interaction in bacterial walls (Antikainen et al. 2007). In addition, this binding activity is also observed in renal epithelial cells where it regulates salt reabsorption (Renigunta et al. 2011).

Pyruvate kinase is the other glycolytic enzyme identified in our study. It is expressed in mammalian tissues as four different isozymes: L, R, M1 and M2 (Gupta & Bamezai 2010). Our study identified the molecular ion *m/z* 1506 as belonging to either pyruvate kinase M1 or pyruvate kinase M2, which share high homology (Table 3.3). The downregulation of pyruvate kinase observed in HIV negative DLBCL tumours compared to normal tissues, confirms the findings from other studies which reported downregulation of pyruvate kinase in gastric and colorectal cancer (Yoo et al. 2004; Martinez-balibrea et al. 2009; Israelsen & Vander 2015) These results suggest that the

tumours may be channelling glucose to other biosynthetic pathways to synthesize nucleotides and other molecules (Israelsen & Vander 2015) .

Another protein identified in this study was heat shock protein 70 which acts as a molecular chaperone. Chaperones can either assist in protein folding or direct misfolded proteins for the ER-associated proteosomal degradation system (Muller et al. 2013; Young et al. 2003). The Hsp70 proteins bind to misfolded proteins during translation or after stress-mediated protein damage (Patterson & Hohfeld 2006) by recognizing short segments of hydrophobic amino acids flanked by basic residues in substrate peptides (Patterson & Hohfeld 2006). The human Hsp70 family consists of at least eight highly homologous members that differ by intracellular localization and expression pattern (Rohde et al). The two ubiquitously expressed Hsp70 proteins include, the constitutively active Hsp70 (Hsc) and the inducible form (Hspi) (Rohde et al. 2005). In our study the constitutively active form of Hsp70 was identified by nLC to correspond to *m/z* 1199 (Table 3.3, Appendix K). Most IHC antibodies detect both HSP isoforms since they have high sequence similarity (Wehner et al. 2003; Leopardi et al. 2001; Qi et al. 2005).

Hsp70 expression affects tumour cell survival through the inhibition of the apoptotic process (Rohde et al. 2005; Tanaka et al. 2014). Hsp70 has also been shown to affect transcription factors involved in the expression of the Bcl-2 family such as p53 thereby interferes its function (Goloudina et al. 2012). The upregulation of Hsp70 in normal tissue of HIV negative patients may be due to the high proliferation of B cells as well as the constant protein synthesis required B cell development (Table 3.3). The

downregulation of Hsp70 in DLBCL samples compared to control in the HIV negative cohort may suggest that this particular isoform may not be important for the development and progression of cancer. However, other forms of Hsp70 may be overexpressed in this context (Murphy 2013; Rohde et al. 2005). In addition, the upregulation of this isoform of Hsp70 in HIV positive DLBCL samples may suggest that it has an important role in the development and progression of cancer in these patients.

Ribosomal proteins were also identified and in conjunction with ribosomal RNA, make up the ribosome which is involved in protein synthesis (Bhavsar et al. 2010; Shenoy et al. 2012). There are eighty different types of ribosomal proteins that make up the subunits of the ribosome (Zhou et al. 2015). The biogenesis of the ribosome as well as protein translation are finely coordinated and essential for cell growth, proliferation and differentiation. Interruptions in any of these two cellular processes can severely retard cell growth and perturb development (Zhou et al. 2015). Cellular stress resulting from nutrient deficiency and toxic chemicals can result in an accumulation of ribosomal proteins in the cytoplasm. These 'ribosome- free' proteins then stimulate immune signalling or tumour-genesis (Zhou et al. 2015).

Ribosomal proteins have been observed to be upregulated in various tumours and cell lines (Henry et al. 1993; Kreunin et al. 2007; Mao-de & Jing 2007) and are said to be involved in the progression of colorectal ,oesophageal and gastric cancer (Wang et al. 2006). Our study showed that the specific ribosomal proteins were differentially regulated depending on the subtype (Table 3.3-Table3.6). While most were

downregulated in DLBCL tumours, there were specific subtypes which were upregulated in tumours. Similarly, the majority of ribosomal proteins were upregulated in the non-GC subtype of DLBCL yet some subtypes were downregulated. This data supports the study by Kasai et al (2003) which observed a differential regulation of ribosomal protein subtypes in normal and colorectal tumours.

Abnormal levels of RPL13a in humans interfere with the cell cycle thus inducing apoptosis (Chen & Ioannou 1999). In our study, RPL13 was overexpressed in both HIV positive DLBCL cases and the non-GC subtype of the HIV negative group suggesting a role in tumour aggressiveness (Table 3.4-Table 3.6).RPS16 was downregulated in HIV negative DLBCL but upregulated in HIV positive tumours (Table 3.3-Table 3.5). The distribution of RPS16 (*m/z* 789) in DLBCL (Figure 3.2), which was also upregulated and concentrated in non-germinal centre subtypes, suggests that not all tumour cells express this protein. The human protein atlas also showed a low to moderate cytoplasmic staining of RPS16 in NHL cases indicating inter- tumour and intra-tumour heterogeneity (Pontén et al. 2008; Uhlén et al. 2005).

Ubiquitin (Ub), which was identified to be fused to a ribosomal protein (RPL40) in our study, is a highly conserved protein comprised of 76 amino-acid residues with diverse intracellular activities (Patterson & Hohfeld 2006). In humans the Ub genes, *Uba80* (HUBCEP80) and *Uba52* (HUBCEP52) encode Ubprotein fused to the ribosomal proteins S27a (RPS27a) and L40 (RPL40), respectively (Han et al. 2012). The STRING analyses suggest that Ubiquitin binds to ribosomal proteins therefore, indicating the

ubiquitin-fusion proteins (Figure 3.3). Our study identified RPL40 to be upregulated in HIV negative tumours but downregulated in HIV positive tumours (Table 3.3, Table 3.5). It was also overexpressed in the non-GC subtype of DLBCL in both HIV negative and positive contexts (Table 3.6, Table 3.7).These results are similar to other reports which show an overexpression of RPL40 and RPS27a in cancer inferring a role in tumour-genesis and tumour aggressiveness (Barnard et al. 1995; Mafune et al. 1991).

STRING analysis show that Ubiquitin also binds to Hsp70 (Figure 3.5), and this is signified via the ubiquitin-like (UbL) protein binding to the ATPase domain of the Hsp70-like Stch protein in yeast (Kaye et al. 2000)-.

Collagen is an abundant structural protein comprises one third of the total protein in humans. Twenty-eight different types of collagen composed of at least 46 distinct polypeptide chains have been identified in vertebrates (Shoulders & Rainers 2010). Its fundamental structural unit is a long (300-nm), thin (1.5-nm-diameter) protein that consists of three coiled subunits of two α1 (I) chains and one α2 (I) chain (Lodish et al. 2000).

Our findings showed that collagen chains, α1 (I) and α2 (VI) were downregulated in DLBCL tumours compared to normal tissues of the HIV positive group as well as in the non- GC subtype of the HIV negative group (Table 3.3-Table 3.5). Both of these groups have a poor prognosis and this may be related to a less responsive host immune system (Staudt & Dave 2005; Rosenwald et al. 2002). Collagen has been previously detected as a

potential biomarker for DLBCL in the molecular profiling study by Rosenawald et al (2002). It was shown to be associated with a 'lymph node signature', which indicated a host immune response and was indicative of a good prognosis (Staudt & Dave 2005). In addition, the collagen genes *COL1A2* and *COL4A1* were downregulated in advanced stage DLBCL (Nishiu et al. 2002) thus confirming upregulated collagen expression being a marker for good prognosis. In addition, some DLBCL tumours are heavily infiltrated with macrophages, thus leading to increased expression of collagen, since macrophages secrets collagen VII, (Staudt & Dave 2005; Schnoor et al. 2008).

Histone proteins are involved in the packaging of DNA and were observed to be differentially expressed in DLBCL. The DNA is hierarchically packed in the nucleus with the aid of proteins to form a complex called chromatin. The nucleosome core particle is composed of two copies of each of histones H2A, H2B, H3 and H4, assembled in an octamer with 146bp of DNA wrapped around it (Marino-Ramirez et al. 2005; Cohen et al. 2011). While the core histones are highly conserved, there are variants which may play role in gene expression and epigenetic silencing (Marino-Ramirez et al. 2005). It is the possible that the histones proteins identified in this project may be variants of core histones which may have functions in tumour development and progression.

The core histones have N-terminal tails of variable length that are subject to extensive PTMs, which have been implicated in transcriptional activation, silencing, chromatin assembly and DNA replication (Marino-Ramirez et al. 2005). The specific PTMs can

either repress or activate gene expression. Ubiquitination of histone H2A on lysine K119 and methylation of lysine 27 on histone H3repress genes while methylation of histone H3 at lysine 4 (H3K4me3) activates gene expression (Ezponda & Licht 2014).

The various PTMs are regulated by many proteins including histone acetyltransferases (HATs), histone deacetylases (HDAC), kinases and histone lysine methyltransferases (Bannister & Kouzarides 2011).

The deregulated activity of the enzymes may play a role in the development of cancer by either altering gene expression programmes of oncogenes/tumour suppressor genes, or by histone modifications affecting the integrity of the chromosome (Bannister & Kouzarides 2011). The deregulated activity of chromatin modifiers is particularly linked to cancer formation and progression (Ellis et al. 2009; Esteller 2011; Geutjes et al. 2012; Gajer et al. 2015). Inhibitors of HAT and HDAC result in growth arrest, differentiation, and apoptosis (Gajer et al. 2015; Lane & Chabner 2009) and are used in clinical trials (Manzo et al. 2009).

The various histone protein subtypes are commonly isolated in MALDI IMS experiments(Seeley & Caprioli 2008; Djidja et al. 2010; Hardesty et al. 2011)due to a technical limitation of MALDI IMS which favours high abundant and soluble proteins (Seeley & Caprioli 2011). Proteins such as Hsp70, albumin and GAPDH are amongst the most abundant and supports the observation made by Seeley and Caprioli (2008). Histone H3 (*m/z* 1032) had intense, widespread distribution in germinal centres and tumour cells (Figure 3.2) which supports the data generated by the atlas protein project in which

intense nuclear expression of histone H3 protein is also expressed in germinal centre cells and NHL tumours (Uhlén et al. 2005; Pontén et al. 2008).

It is encouraging to note that, there are similar proteins differentially expressed between normal and tumour of both HIV negative DLBCL and HIV positive DLBCL suggesting a similar pathology. However, other proteins such as annexin A5 and albumin, were also differentially expressed in our cohort however, the STRING analysis did not find a relationship between these proteins and other proteins identified in the study (Figure3.5). The differences in albumin levels observed may be due to different levels if serum present in the tissue sections (Londoño et al. 1992). However, albumin is decreased in both HIV negative and HIV positive DLBCL samples compared to controls (Table 3.3 and Table 3.4) and not statistically different in the other comparison sets thus, implying that this observation is specific to tumour processes and not due to background.

Annexin A5 is a calcium binding protein that is known to play a role in the development and progression of various cancers (Deng et al. 2013; Prognosis et al. 2009). An increase in annexin A5 has been correlated with advanced stage, metastasis and poor prognosis in colorectal cancer, breast and pancreatic cancer (Prognosis et al. 2009; Deng et al. 2013). In our study, annexin A5 was increased in DLBCL tumours compared to control tissue suggesting a role in the development of the DLBCL. In addition, annexin A5 was

decreased in the non-germinal centre subtype, in both HIV contexts, suggesting it may have a role in good prognosis.

3.4.2 The detection of HIV proteins

It has been postulated that HIV initiates the development of cancer by the chronic stimulation of B cells (He et al. 2006; Gloghini et al. 2013a).In order to determine whether HIV does this in close proximity to the tumour, the HIV capsid protein, p24, was immunohistochemically detected in the MALDI IMS samples. P24 protein was detected in the network of follicular dendritic cells of a single reactive node sample (Figure 3.10), thus confirming previously published data (Moonim et al. 2010). In addition, a lack of expression was to be expected as HIV infects $CD4^{+}$T helper cells which may be limited by the diffuse spread of tumorous B cells (Campo & Rule 2015). Furthermore, there is a depletion of CD4 T cells in the region due to HIV infection and lysis (Liapis et al. 2013). The viral database search for HIV proteins using the peptide mass list also did not produce any match against HIV proteins which is probably due to 'silencing' by other abundant proteins. The peptides from all the HIV positive samples were mixed and ran on the nLC instrument. If only one sample has the HIV proteins, it is possible that the protein quantity may have been insufficient for successful detection.

3.4.3 Confirmation of the expression pattern of heat shock protein 70

There are several aspects of HSP70 function that are important for cancer development and progression. At the initial stages of tumour-genesis, Hsp70 can protect cells undergoing transformation from oncogenic stress induced by overexpression of oncogenes (E.A. Afanasyeva, E.Y. Komarova, L.G. Larsson, F. Bahram, B.A. Margulis 2007). In addition, Hsp70 has been shown to suppress cellular senescence which is an important anti-tumour mechanism at the early stages of tumour-genesis and also important in how cancer cells respond to therapy (Sherman et al. 2007). Hsp70 is an essential factor for tumour cell survival and tumour growth. Elevated levels of Hsp70 in tumours correlate with poor prognosis and higher resistance to cancer treatment for cancer patients (Ciocca & Calderwood 2005).

It is for these reasons that Hsp70 was chosen as the protein whose expression was going to be confirmed in a separate cohort of DLBCL samples. Most of the DLBCL cases did not express Hsp70 confirming previously observed data in non-small lung cancer (Qi et al. 2005). The lack of Hsp70 expression in the GC cells in RN samples contradicts published data where the blast cells were intensely positive (Leopardi et al. 2001). Differences in antibodies and IHC conditions may account for the discrepancies. Our study utilised a polyclonal antibody while Leopardi used a monoclonal antibody which recognise all forms of Hsp70. Interestingly, this is a first report showing Hsp70 expression in GC plasma cells (Figure 3.7, Appendix J). It is therefore possible that the

expression in plasma cells may be the responsible for the expression of Hsp70 in the samples used in MALDI IMS (Figure 3.2). In addition, HIV infected tumours express high levels of plasma cells (Liapis et al. 2013), and this may explain the higher Hsp70 expression levels in the HIV positive cohort (Table 3.8). Interestingly equal proportion of GC and non-GC subtypes expressed Hsp70 in the HIV positive cohort, contrasting with the MALDI IMS analysis results (Table 3.3). However, the sample size used in the study may not give representative results. The Hsp70 expression levels in the non-GC subtype suggest that Hsp70 may be involved in treatment resistance (Fang et al. 2013; Carbone & Gloghini 2014), since high levels of Hsp70 inhibit apoptosis and promote apoptosis (Qi et al. 2005; Liu et al. 2011; Mjahed et al. 2012). The GC subtype expresses high levels of BCL-2 to inhibit apoptosis (Shaffer et al. 2002). However, Hsp70 has been shown to stabilise BCL-2 under oxidative stress (Jiang et al. 2011). It is possible that Hsp70 and BCL-2 work concertedly to inhibit apoptosis and promote proliferations of DLBCL GC tumours.

In summary, Hsp70 was overexpressed in HIV positive samples of control and tumour samples as well as in the non-GC subtype of DLBCL in the HIV negative cohort, thus confirming the MALDI IMS data (Table 3.8).

The expression of Hsp70 was correlated with survival where Hsp70 expression signified a poor prognosis in the HIV negative group but not in the HIV positive group (Table 3.8, Appendix N). Patients with Hsp70 expression in tumour cells either relapsed after treatment or died during treatment (Appendix F), thus confirming its role in

chemotherapy resistance (Liu et al. 2011; Fang et al. 2013). The study had a lower number of HIV positive sample and the prognostic significance of Hsp70 in this group may increase with a larger sample size. In addition, no correlations with clinic-biological features were possible due to only eleven samples being positive (Table 3.8).

The study was met with a few limitations which were unavoidable. The MALDI IMS analysis had a low sample size due to the high expense of this technique; which was increased during the Hsp70 IHC analysis. However, the sample size of this cohort also decreased due to lack of Hsp70 expression in a substantial amount of the samples (Table 3.7). In addition, the study samples were varied in terms of tissue origin as well as in basic clinical characteristics thus increasing the variation. This was unavoidable as the study population also differed in some clinical characteristics and contained mostly extranodal DLBCL cases (Chapter 2) which would diversify the tissue origin. We also encountered technical limitations that are inherent to the MALDI IMS machinery such as low resolution and detection of abundant proteins (Seeley & Caprioli 2011).Finally, the study was limited by using only one protein to confirm the MALDI IMS data, which lowered the sensitivity and specificity as previously published (García-Bilbao et al. 2012).

In summary, the study identified proteins such as the histones, GAPDH, Hsp70 and ribosomal proteins, to discriminate between normal and DLBCL tumour, as well as between DLBCL subtypes. Furthermore, these proteins were identified in all DLBCL cases irrespective of HIV status, suggesting that DLBCL tumour-genesis is similar

between the two cohorts. The expression pattern of one of the proteins, HSP70, was confirmed by immunohistochemical staining using a separate cohort of cases. Expression of Hsp70 was correlated with worse prognosis in HIV negative DLBCL cases but not HIV positive cases.

CHAPTER 4: CONCLUSIONS AND FUTURE RECOMMENDATIONS

DLBCL is divided into prognostically significant groups using the IPI. However, this is not able to sufficiently classify all patients, and does not highlight potential therapeutic targets. There are various molecular markers for DLBCL, which are not used clinically due to sensitivity and specificity issues. However, these have assisted in advancing knowledge on the development and progression of the disease. These have not been studied in the HIV context, therefore leading to a lack of understanding on the molecular pathology of the disease. The aim of the study was to determine biomarkers for HIV and non- HIV related diffuse large B cell lymphoma. The specific objectives of the study included:

1. Determining differences in the clinicipathologic features of HIV negative and HIV positive DLBCL.
2. Identifying protein biomarkers for HIV negative and HIV positive DLBCL using MALDI Imaging MS.
3. Characterizing the identified proteins using LC-MS/MS.
4. Correlating the resulting molecular data with clinicopathological information.

The study showed that DLBCL from HIV negative and HIV positive patients presents in a clinically similar manner except for differences in median age and frequency of elevated LDH levels, thus suggesting possible similarities in lymphomagenesis pathways

between the two cohorts. In addition, the study identified potential clinical biomarkers that may be used for prognosis in addition to current prognostic indicators. Two of the clinical parameters were indicative of good prognosis while the other two were indicative of worse prognosis. The study may have been limited by record keeping and confounding, yet the prognostic markers highlighted herein have been previously identified by other studies. In addition, some of them form part of the IPI thus confirming their legitimacy. These clinical biomarkers were found to be prognostic for all DLBCL cases, suggesting that these may be used without the knowledge of the patient's HIV status. Identical parameters were observed to be prognostic for the HIV negative group, with the addition of extranodal tumour site, suggesting that the extranodal site might be prognostic. We did not observe statistically significant difference in OS rates between the three major extranodal sites of the HIV negative cohort. However, comparisons with the other less frequently detected extranodal sites might shed more light. We were not able to determine prognostic factors specific for the HIV positive group due to low sample size. This group was smaller than the HIV negative group and with the added complication of missing information, resulted in non-significant prognostic indicators.

This study also identified protein biomarkers for both HIV negative DLBCL and HIV positive DLBCL including clusters of inter-connected proteins differentially expressed between the two cohorts. These included glycolytic proteins, histones, ribosomal proteins and collagen. These proteins have been previously identified in other cancers by MALDI IMS technology, signifying a restriction to identify certain proteins. However, these

proteins have been previously shown to have roles in cancer development and progression, thus supporting their potency. MALDI IMS is a powerful technology yet it is often limited by high costs and numerous pre-analytical steps, however, the data presented in this thesis highlights the value it may bring to cancer research.

The identification of similar proteins in HIV negative and HIV positive DLBCL highlights the similarities between the two cohorts. However, the proteins were sometimes differentially regulated pointing to different roles within each context. Therefore, the protein markers identified herein require further studies use as for potential therapeutic targets.

The expression of one of the proteins identified, Hsp70, was confirmed using a separate cohort of DLBCL samples. Our results confirm the expression pattern observed in MALDI IMS analyses. This substantiates the use of MALDI IMS technology to correctly determine the distribution of different molecules or ions within tissues. We showed that Hsp70, despite low sample size, has prognostic significance for DLBCL samples. The expression pattern of this protein in DLBCL samples suggests that more studies are required to determine its role in DLBCL development and progression.

In conclusion, all the objectives of this project were successfully accomplished. Biomarkers for DLBCL were identified including clinical and molecular markers which may be used for diagnostic and prognostic purposes. However, more studies are required prior to this. Therefore, future studies would involve:

1. The confirmation of these results in a multicentre cohort consisting of DLBCL cases from HIV negative and HIV positive patients.
2. The determination of the distribution and prognosis of the other proteins identified herein, both individually and in a synchronous manner.
3. Studying the pathways of the identified proteins for additional therapeutic targets.

CHAPTER 5: REFERENCES

Aisenberg, A., 2000. Historical Review of lymphomas. *British Journal of Haematology*, 109, pp.466–476.

Akanmu, A.S., 2006. AIDS-associated malignancies. *African journal of medicine and medical sciences*, 35 Suppl, pp.57–70.

Akyurek, N. et al., 2012. Prognostic Significance of MYC , BCL2 , and BCL6 Rearrangements in Patients With Diffuse Large B-Cell Lymphoma Treated With Cyclophosphamide , Doxorubicin , Vincristine , and Prednisone Plus Rituximab. *Cancer*, 118, pp.4173–4183.

Al-humood, S.A. et al., 2011. Genotypic and Phenotypic Differences between Nodal and Extranodal Diffuse Large B-Cell Lymphomas. *Journal of Histochemistry & Cytochemistry*, 59(10), pp.918–31.

Alizadeh, A.A. et al., 2000. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*, 403, pp.503–511.

Almieda, C.A. & Barry, S.A., 2009. *Cancer: Basic Science and Clinical Aspects*, Wiley-Blackwell.

Álvarez-chaver, P. et al., 2014. Proteomics for discovery of candidate colorectal cancer biomarkers. *World journal of gastroenterology : WJG*, 20(14), pp.3804–3824.

Antikainen, J. et al., 2007. pH-Dependent Association of Enolase and Glyceraldehyde-3-Phosphate Dehydrogenase of Lactobacillus crispatus with the Cell Wall and Lipoteichoic Acids. *Journal of bacteriology*, 189(12), pp.4539–4543.

Aukema, S.M. et al., 2011. Double-hit B-cell lymphomas. *blood*, 117(8), pp.2319–2331.

Balluff, B. et al., 2015. *De novo* discovery of phenotypic intratumour heterogeneity using imaging mass spectrometry. *The Journal of Pathology*, 235(1), pp.3–13.

Balluff, B. et al., 2011. MALDI imaging mass spectrometry for direct tissue analysis: technological advancements and recent applications. *Histochemistry and cell biology*, 136(3), pp.227–44.

Baltimore, D., 2001. Our genome unveiled. *Nature*, 409, pp.814–816.

Banerjee, S. & Mazumdar, S., 2012. Electrospray Ionization Mass Spectrometry : A Technique to Access the Information beyond the Molecular Weight of the Analyte. *Internation journal of analytical chemistry*, 2012, p.ID 282574.

Banham, A.H. et al., 2005. Expression of the FOXP1 Transcription Factor Is Strongly Associated with Inferior Survival in Patients with Diffuse Large B-Cell Lymphoma. *Clinical Cancer Research*, 11, pp.1065–1072.

Bankovich, A.J. et al., 2007. Structural Insight into Pre-B Cell Receptor Function. *Science*, 316(11), pp.291–294.

Bannister, A.J. & Kouzarides, T., 2011. Regulation of chromatin by histone modifications. *Nature Publishing Group*, 21(3), pp.381–395.

Baptista, J.M. et al., 2015. HIV-infection impact on clinical – biological features and outcome of diffuse large B-cell lymphoma treated with R-CHOP in the combination antiretroviral therapy era. *AIDS*, 29, pp.811–818.

Barbini, L. et al., 2007. Glyceraldehyde-3-phosphate dehydrogenase exerts different biologic activities in apoptotic and proliferating hepatocytes according to its subcellular localization. *Molecular and Cellular Biochemistry*, 300, pp.19–28.
Barnard, G. et al., 1995. Ubiquitin fusion proteins are overexpressed in colon cancer but not in gastric cancer. *Biochim Biophys Acta.*, 1272(3), pp.147–53.
Barrans, S.L. et al., 2004. Brief report Strong expression of FOXP1 identifies a distinct subset of diffuse large B-cell lymphoma (DLBCL) patients with poor outcome. *Blood*, 104(9), pp.2933–2936.
Barrans, S.L. et al., 2003. The t (14 ; 18) Is Associated with Germinal Center-derived Diffuse Large B-Cell Lymphoma and Is a Strong Predictor of Outcome. *Clinical Cancer Research*, 9, pp.2133–2139.
Barreto, L., Azambuja, D. & Morais, J.C. De, 2012. Expression of immunohistochemical markers in patients with AIDS-related lymphoma. *The Brazilian journal of infectious diseases : an official publication of the Brazilian Society of Infectious Diseases*, 16(1), pp.74–7.
Basso, K. & Dalla-favera, R., 2015. Germinal centres and B cell lymphomagenesis E l. *Nature Publishing Group*, 15(3), pp.172–184.
Beral, P. et al., 1991. AIDS-associated non-Hodgkin lymphoma. *The Lancet*, 337(8745), pp.805–809.
Berg, J., Tymoczko, J. & Stryer, L., 2002. The Glycolytic Pathway Is Tightly Controlled. In *Biochemistry. 5th edition*. New York: W H Freeman.
Besson, C. & Raphaël, M., 2003. Lymphoma genesis in the context of HIV infection. *Annales de médecine interne*, 154(8), pp.523–8.
Bhavsar, R.B., Makley, L.N. & Tsonis, P.A., 2010. The other lives of ribosomal proteins. *Human genomics*, 4(5), pp.327–344.
Biggar, R.J. et al., 2007. AIDS-Related Cancer and Severity of Immunosuppression in Persons With AIDS. *J Nat Cancer Inst*, 99(12), pp.962–72.
Blanco, F. et al., First-line therapy and mitochondrial damage: different nucleosides, different findings. *HIV clinical trials*, 4(1), pp.11–9.
Bohlius, J. et al., 2009. Prognosis of HIV-associated non-Hodgkin lymphoma in patients starting combination antiretroviral therapy. *AIDS*, 23(15), pp.2029–37.
Bower, M., Palmieri, C. & Stebbing, J., 2006. AIDS associated malignancies. *Update on Cancer Therapeutics*, 1(2), pp.221–234.
Broyde, A. et al., 2009. Role and prognostic significance of the Ki-67 index in non-Hodgkin ' s lymphoma. *Am J Heamtol*, 84, pp.338–343.
Bürgesser, M.V. et al., 2013. Clinicopathological features of aggressive B-cell lymphomas including B-cell lymphoma, unclassifiable, with features intermediate between diffuse large B-cell and Burkitt lymphomas: a study of 44 patients from Argentina. *Annals of diagnostic pathology*, 17(3), pp.250–5.
Byrd, H.C.M. & McEwen, C.N., 2000. The Limitations of MALDI-TOF Mass Spectrometry in the Analysis of Wide Polydisperse Polymers. *Analytical Chemistry*, 72(19), pp.4568–4576.
Calderaro, A. et al., 2014. MALDI-TOF mass spectrometry applied to virus

identification. *Scientific Reports*, 4, p.6803.
Campo, E. et al., 2011. The 2008 WHO classification of lymphoid neoplasms and beyond : evolving concepts and practical applications. *Blood*, 117(19), pp.5019–5033.
Campo, E. & Rule, S., 2015. Mantle cell lymphoma: evolving management strategies. *Blood*, 125(1), pp.48–55.
Capello, M. et al., 2011. a -enolase : a promising therapeutic and diagnostic tumor target. *FEBS Journal*, 278, pp.1064–1074.
Caprioli, R.M., Farmer, T.B. & Gile, J., 1997. Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Analytical chemistry*, 69(23), pp.4751–60.
Carbone, A. & Gloghini, A., 2014. Diffuse large B cell lymphoma : using pathologic and molecular biomarkers to define subgroups for novel therapy. *Ann Hematol*, 93, pp.1263–1277.
Casadonte, R. & Caprioli, R.M., 2011. Proteomic analysis of formalin-fixed paraffin-embedded tissue by MALDI imaging mass spectrometry. *Nature protocols*, 6(11), pp.1695–709.
Castellino, S., Groseclose, M.R. & Wagner, D., 2011. MALDI imaging mass spectrometry: bridging biology and chemistry in drug development. *Bioanalysis*, 3, pp.2427–2441.
Castillo, J.J., Winer, E.S. & Olszewski, A.J., 2014. Sites of extranodal involvement are prognostic in patients with diffuse large B-cell lymphoma in the rituximab era : An analysis of the Surveillance , Epidemiology and End Results database. *Am J Heamtol*, 89(3), pp.310–314.
Centers for disease control, 1987. Revision of the CDC Surveillance case definition for Acquired Immunodeficienct syndrome. *MMWR*, 36(1S), pp.1–15.
Chadburn, A. et al., 2009. Immunophenotypic analysis of AIDS-related diffuse large B-cell lymphoma and clinical implications in patients from AIDS Malignancies Consortium clinical trials 010 and 034. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 27(30), pp.5039–48.
Chandramouli, K. & Qian, P., 2009. Proteomics : Challenges , Techniques and Possibilities to Overcome Biological Sample Complexity. *Human genomics and proteomics*, 2009.
Chao, C. et al., 2015. A comparative study of molecular characteristics of diffuse large B-cell lymphoma from patients with and without human immunodeficiency virus infection. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 21(6), pp.1429–37.
Chao, C. et al., 2013. Epstein-Barr virus infection and expression of B-cell oncogenic markers in HIV-related diffuse large B-cell lymphoma. *Clin Cancer Res*, 18(17), pp.1–23.
Chen, F.W. & Ioannou, Y.A., 1999. Ribosomal proteins in cell proliferation and apoptosis. *International reviews of immunology*, 18(5-6), pp.429–48.
Chiche, J. et al., 2014. GAPDH enhances the aggressiveness and the vascularization of

non-Hodgkin's B lymphomas via NF-κB-dependent induction of HIF-1α. *Leukemia*.
Choi, W.W.L. et al., 2009. A New Immunostain Algorithm Classifies Diffuse Large B-Cell Lymphoma into Molecular Subtypes with High Accuracy A New Immunostain Algorithm Classifies Diffuse Large B-Cell Lymphoma into Molecular Subtypes with High Accuracy. *Clinical Cancer Research*, 15(17), pp.5494–5503.
Chung, C.H. et al., 2007. Genomics and proteomics: Emerging technologies in clinical cancer research. *Critical Reviews in Oncology/Hematology*, 61(1), pp.1–25.
Ciocca, D.R. & Calderwood, S.K., 2005. Heat shock proteins in cancer: diagnostic, prognostic, predictive, and treatment implications. *Cell Stress Chaperones*, 10, pp.86–103.
Coffin JM, Hughes SH, V.H., 1997. *Retroviruses*, Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.
Cohen, I. et al., 2011. Histone Modifiers in Cancer: Friends or Foes ? *Genes & Cancer*, 2(6), pp.631–647.
Cole, L.M. et al., 2014. MALDI-MSI and label-free LC-ESI-MS / MS shotgun proteomics to investigate protein induction in a murine fibrosarcoma model following treatment with a vascular disrupting agent. *Proteomics*, 14, pp.890–903.
Contran, R., Kumar, V. & Robbins, S., 1989. *Robbins Pathologic Basis of Disease* 4th editio., Philadelphia: WB Saunders.
Cornett, D. et al., 2007. MALDI imaging mass spectrometry: molecular snapshots of biochemical systems. *Nat Methods.*, 10, pp.828–33.
Craven, R. a et al., 2013. Proteomic analysis of formalin-fixed paraffin-embedded renal tissue samples by label-free MS: assessment of overall technical variability and the impact of block age. *Proteomics. Clinical applications*, 7(3-4), pp.273–82.
Cultrera, J.L. & Dalia, S.M., 2012. Diffuse Large B-Cell Lymphoma : Current Strategies and Future Directions. *Cancer control*, 19(3), pp.204–213.
Dave, S.S. et al., 2006. Molecular Diagnosis of Burkitt's Lymphoma. *The New England journal of medicine*, 354, pp.2431–2442.
Davis, R.E. et al., 2001. Constitutive Nuclear Factor KB Activity Is Required for Survival of Activated B Cell – like Diffuse Large B Cell Lymphoma Cells. *The journal of experimental medicine*, 194(12), pp.1861–1874.
Deeb, S.J. et al., 2012. Super-SILAC allows classification of diffuse large B-cell lymphoma subtypes by their protein expression profiles. *Molecular & cellular proteomics*, 11, pp.77–89.
Deeks, S.G., Lewin, S.R. & Havlir, D. V, 2013. The end of AIDS : HIV infection as a chronic disease. *The Lancet*, 382(9903), pp.1525–1533.
Deng, S. et al., 2013. Annexin A1 , A2 , A4 and A5 play important roles in breast cancer , pancreatic cancer and laryngeal carcinoma , alone and / or synergistically. *Oncology letters*, 5, pp.107–112.
Diaz-Ramos, A. et al., 2012. α -Enolase , a Multifunctional Protein : Its Role on Pathophysiological Situations. *Journal of Biomedicine and Biotechnology*, 2012.
Djidja, M. et al., 2010. Novel molecular tumour classification using MALDI – mass spectrometry imaging of tissue micro-array. *Anal Bioanal Chem*, 397, pp.587–601.

Djidja, M.-C. et al., 2009. MALDI-ion mobility separation-mass spectrometry imaging of glucose-regulated protein 78 kDa (Grp78) in human formalin-fixed, paraffin-embedded pancreatic adenocarcinoma tissue sections. *Journal of proteome research*, 8(10), pp.4876–84. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19673544.

Dominis, M. et al., 2002. Diffuse Large B-Cell Lymphoma and Its Variants. *Croat Med J*, 43(5), pp.535–540.

Drucker, E. & Krapfenbauer, K., 2013. Pitfalls and limitations in translation from biomarker discovery to clinical utility in predictive and personalised medicine. *The EPMA journal*, 4(1), p.7.

Dunleavy, K., Grant, C. & Wilson, W.H., 2013. Using biologic predictive factors to direct therapy of diffuse large B-cell lymphoma. *theraeutics advances in hematology*, 4(1), pp.43–57.

Dunleavy, K. & Wilson, W.H., 2012. How I treat HIV-associated lymphoma. *Blood*, 119(14), pp.3245–55.

Dunleavy, K. & Wilson, W.H., 2011. The differential role of BCL-2 within molecular subtypes of DLBCL. *Clin Cancer Res*, 17(24), pp.7505–7507.

E.A. Afanasyeva, E.Y. Komarova, L.G. Larsson, F. Bahram, B.A. Margulis, I.V.G., 2007. Drug-induced Myc-mediated apoptosis of cancer cells is inhibited by stress protein Hsp70. *Int. J. Cancer*, 121, pp.2615–2621.

Ellis, L., Atadja, P. & Johnstone, R., 2009. Epigenetics in cancer: targeting chromatin modifications. *Mol Cancer Ther*, 8, pp.1409–1420.

Eriksson, S. et al., 2013. Comparative analysis of measures of viral reservoirs in HIV-1 eradication studies. *PLoS pathogens*, 9(2), p.e1003174.

Esteller, M., 2011. Epigenetic changes in cancer. *Biol Rep*, 3, p.9.

Evans, L.S. & Hancock, B.W., 2003. Non-Hodgkin lymphoma. *Lancet*, 362(9378), pp.139–46.

Ezponda, T. & Licht, J.D., 2014. Molecular Pathways : Deregulation of Histone H3 Lysine 27 Methylation in Cancer — Different Paths , Same Destination. *Clin Cancer Res*, 20(19), pp.5001–5009.

Fang, X. et al., 2013. Blockade of PI3K / AKT pathway enhances sensitivity of Raji cells to chemotherapy through down-regulation of HSP70. *Cancer Cell International*, 13(1), p.1. Available at: Cancer Cell International.

Ferlay, J. et al., 2012. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. *Lyon, France: International Agency for Research on Cancer; 2013. Available from: http://globocan.iarc.fr, accessed on day/month/year*.

Flowers, C.R. et al., 2013. Examining racial differences in diffuse large B-cell lymphoma presentation and survival. *Leukemia & lymphoma*, 54(2), pp.268–76.

Frantzi, M., Bhat, A. & Latosinska, A., 2014. Clinical proteomic biomarkers: relevant issues on study design & technical considerations in biomarker development. *Clinical and translational medicine*, 3(1), p.7.

Gajer, J.M. et al., 2015. Histone acetyltransferase inhibitors block neuroblastoma cell growth in vivo. *Oncogenesis*, 4, p.e137.

García-Bilbao, A. et al., 2012. Identification of a biomarker panel for colorectal cancer diagnosis. *BMC cancer*, 12, p.43.
Garvin, A. et al., 1980. The Rappaport classification of non-Hodgkin's lymphomas: a closer look using other proposed classifications. *Semin Oncol*, 7(3), pp.234–43.
Gatto, D. & Brink, R., 2012. Mechanisms of allergic diseases The germinal center reaction. *Journal of Allergy and Clinical Immunology*, 126(5), pp.898–907.
Geisler, C. et al., 2015. Identification and validation of potential new biomarkers for prostate cancer diagnosis and prognosis using 2D-DIGE and MS. *Biomed Res Int.*, 2015, p.454256.
Geutjes, E., Bajpe, P. & Bernards, R., 2012. Targeting the epigenome for treatment of cancer. *Oncogene*, 31, pp.3827–3844.
Giunti, M. et al., 2002. Hodgkin ' s lymphoma : the pathologist ' s viewpoint. *J Clin Pathol*, 55, pp.162–176.
Gloghini, A., Dolcetti, R. & Carbone, A., 2013a. Lymphomas occurring specifically in HIV-infected patients: From pathogenesis to pathology. *Seminars in Cancer Biology*, 23(6 PA), pp.457–467.
Golematis, B.C. & DeVita Jr, V.T., 1996. *AIDS and malignancies : current concepts and prospectives* K. N. Syrigos & M. M. Konstantoulakis, eds., New York: Wiley.
Goloudina, A.R., Demidov, O.N. & Garrido, C., 2012. Inhibition of HSP70 : A challenging anti-cancer strategy. *Cancer Letters*, 325(2), pp.117–124.
Grogg, K.L., Miller, R.F. & Dogan, A., 2007. HIV infection and lymphoma. *J Clin Pathol*, 60, pp.1365–1372.
Guerrera, I.C. & Kleiner, O., Application of mass spectrometry in proteomics. *Bioscience reports*, 25(1-2), pp.71–93.
Guo, Y. et al., 2014. Array-comparative genomic hybridization profiling of immunohistochemical subgroups of diffuse large B-cell lymphoma shows distinct genomic alterations. *Cancer Science*, 105(4), pp.481–489.
Gupta, V. & Bamezai, R.N.K., 2010. Human pyruvate kinase M2 : A multifunctional protein. *Protein science*, 19, pp.2031–2044.
Gustafsson, J.O.R. et al., 2010. Citric Acid Antigen Retrieval (CAAR) for Tryptic Peptide Imaging Directly on Archived Formalin-Fixed Paraffin-Embedded Tissue research articles. *Journal of proteome research*, 9, pp.4315–4328.
Gustafsson, J.O.R. et al., 2012. Internal calibrants allow high accuracy peptide matching between MALDI imaging MS and LC-MS / MS ☆. *Journal of Proteomics*, 75(16), pp.5093–5105. Available at: http://dx.doi.org/10.1016/j.jprot.2012.04.054.
Hamdan, M.H., 2007. *Cancer Biomarkers: Analytical Techniques for Discovery* N. . Hoboken, ed., Wiley-Interscience.
Hampe, C.S., 2012. B Cells in Autoimmune Diseases. *Scientifica*, 2012.
Han, H., 2011. A high performance profile-biomarker diagnosis for mass spectral profiles. *BMC systems biology*, 5 Suppl 2(Suppl 2), p.S5.
Han, X. et al., 2012. Altered dynamics of ubiquitin hybrid proteins during tumor cell apoptosis. *Cell death and diseases*, 3, p.e255.
Han, X., Aslanian, A. & Yates, J.R., 2008. Mass spectrometry for proteomics. *Current*

opinion in chemical biology, 12(5), pp.483–90.
Hans, C.P. et al., 2004. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood*, 103, pp.275–282.
Hardesty, W.M. et al., 2011. Protein signatures for survival and recurrence in metastatic melanoma. *Journal of proteomics*, 74(7), pp.1002–14.
Harris, C.C., 1996. Structure and Function of the p53 Tumor Suppressor Gene: Clues for Rational Cancer Therapeutic Strategies. *JNCI Journal of the National Cancer Institute*, 88(20), pp.1442–1455.
He, B. et al., 2006. HIV-1 envelope triggers polyclonal Ig class switch recombination through a CD40-independent mechanism involving BAFF and C-type lectin receptors. *Journal of immunology (Baltimore, Md. : 1950)*, 176(7), pp.3931–41.
He, X. et al., 2014. Ki-67 is a valuable prognostic predictor of lymphoma but its utility varies in lymphoma subtypes : evidence from a systematic meta-analysis. *BMC Cancer*, 14(1), pp.1–13. Available at: BMC Cancer.
Hennessy, B.T., Hanrahan, E.O. & Daly, P.A., 2004. Non-Hodgkin lymphoma : an update. *Lancet Oncol*, 5, pp.341–353.
Henry, J.L., Coggin, D.L. & King, C.R., 1993. High-Level Expression of the Ribosomal Protein L19 in Human Breast T \ imors That Overexpress erbB-21. *Cancer research*, 53, pp.1403–1408.
Hirsch, J. et al., 2004. Proteomics : current techniques and potential applications to lung disease. *Am J Physiol Lung cell mol phyio;*, 287, pp.1–23.
Hoeller, S. et al., 2010. FOXP1 protein overexpression is associated with inferior outcome in nodal diffuse large B-cell lymphomas with non-germinal centre phenotype , independent of gains and structural aberrations at 3p14 . 1. *Histopathology*, 57, pp.73–80.
Hoffbrand, A., Moss, P. & Pettit, J., 2006. *Essential Haematology* 5th editio., Blackwell.
Hsiao, K. et al., 2013. Surface a -Enolase Promotes Extracellular Matrix Degradation and Tumor Metastasis and Represents a New Therapeutic Target. *PloS one*, 8(7), p.e69354.
Hu, S. et al., 2013. MYC/BCL2 protein coexpression contributes to the inferior survival of activated B-cell subtype of diffuse large B-cell lymphoma and demonstrates high-risk gene expression signatures: a report from The International DLBCL Rituximab-CHOP Consortium Program. *Blood*, 121(20), pp.4021–4032.
Hu, Y., Yang, K. & Krause, J.R., 2011. Diffuse Large B-cell Lymphoma , Differential Diagnosis and Molecular Stratification. *North American journal of medicine and Science*, 4(2), pp.67–76.
Hudnall, S.D., 2011. *Hematology: A Pathophysiologic Approach*, Elsevier Health Sciences.
Hulka, B., 1990. *Overview of biological markers. In: Biological markers in epidemiology* B. Hulka, J. Griffith, & T. Wilcosky, eds., New York: Oxford University Press.
Hunt, K.E. & Reichard, K.K., 2008. Diffuse large B-cell lymphoma. *Archives of pathology & laboratory medicine*, 132(1), pp.118–24.

Hwang, S.-I. et al., 2007. Direct cancer tissue proteomics: a method to identify candidate cancer biomarkers from formalin-fixed paraffin-embedded archival tissues. *Oncogene*, 26(1), pp.65–76.
Iqbal, J. et al., 2006. BCL2 Expression Is a Prognostic Marker for the Activated B-Cell – Like Type of Diffuse Large B-Cell Lymphoma. *Journal of clinical oncology*, 24, pp.961–8.
Iqbal, J. et al., 2004. BCL2 Translocation Defines a Unique Tumor Subset within the Germinal Center B-Cell-Like Diffuse Large B-Cell Lymphoma. , 165(1), pp.159–166.
Israelsen, W.J. & Vander, M.G., 2015. Seminars in Cell & Developmental Biology Pyruvate kinase : Function , regulation and role in cancer. *Seminars in Cell and Developmental Biology*, 43, pp.43–51.
Jack, a. & Barrans, S., 2004. Recent advances in the understanding of aggressive B-cell lymphomas. *Current Diagnostic Pathology*, 10(5), pp.360–373.
Jaffe, E.S. et al., 2008. Classification of lymphoid neoplasms: the microscope as a tool for disease discovery. *Blood*, 112(12), pp.4384–99.
Jaffe, E.S., 2009. The 2008 WHO classification of lymphomas: implications for clinical practice and translational research. *Hematology / the Education Program of the American Society of Hematology. American Society of Hematology. Education Program*, pp.523–31.
Jerkeman, M. et al., 2002. Prognostic implications of BCL6 rearrangement in uniformly treated patients with diffuse large B-cell lymphoma - a Nordic Lymphoma Group study. *International Journal of Oncology*, 20(1), pp.161–165.
Jiang, B. et al., 2011. Increased stability of Bcl-2 in HSP70-mediated protection against apoptosis induced by oxidative stress. *Cell stress & chaperones*, 16(2), pp.143–52.
Johnson, N.A. et al., 2012. Concurrent expression of MYC and BCL2 in diffuse large B-cell lymphoma treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. *Journal of Clinical Oncology*, 30(28), pp.3452–3459.
Kaplan, L.D., 2012. HIV-associated lymphoma. *Best practice & research. Clinical haematology*, 25(1), pp.101–17.
Karas, M. et al., 1987. Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *International Journal of Mass Spectrometry and Ion Processes*, 78, pp.53–68.
Karube, K. & Campo, E., 2015. MYC Alterations in Diffuse Large B-Cell Lymphomas. *Seminars in Hematology*, 52(2), pp.97–106.
Kaye, F.J. et al., 2000. A family of ubiquitin-like proteins binds the ATPase domain of HSp70 like Stch. *FEBS Letters*, 467, pp.348–352.
Khadir, A. & Tiss, A., 2013. Proteomic aproaches towards early detection and diagnosis of cancer. *J Carcinogene Mutagene*, S14, p.002.
Kishimoto, N. et al., 2012. Glyceraldehyde 3-phosphate dehydrogenase negatively regulates human immunodeficiency virus type 1 infection. *Retrovirology*, 9(1), p.1.
Klapper, W. et al., 2012. Patient age at diagnosis is associated with the molecular characteristics of diffuse large B-cell lymphoma. *Blood*, 119(8), pp.1882–7.

Kramer, M. et al., 1998. Clinical Relevance of BCL2, BCL6, and MYC Rearrangements in Diffuse Large B-Cell Lymphoma. *Blood*, 92(9), pp.3152–3162.
Kreunin, P. et al., 2007. Differential expression of ribosomal proteins in a human metastasis model identified by coupling 2-D liquid chromatography and mass spectrometry. *Cancer genomics & Proteomics*, 4, pp.329–340.
Kumari, S. & Malla, R., 2015. New Insight on the Role of Plasminogen Receptor in Cancer Progression. *Cancer growth and metastasis*, 8, pp.35–42.
Kuppers, R. et al., 1999. Cellular Origin of Human B -cell lymphomas. *The New England journal of medicine*, 341(20), pp.1520–1529.
Lane, A.A. & Chabner, B.A., 2009. Histone Deacetylase Inhibitors in Cancer Therapy. *J Clin Oncol*, 27(32), pp.5459–5468.
Lebien, T.W. & Tedder, T.F., 2008. B lymphocytes : how they develop and function. *Blood*, 112, pp.1570–1580.
Lemaire, R. et al., 2007. Direct Analysis and MALDI Imaging of Formalin-Fixed , Paraffin-Embedded Tissue Sections. *J Proteome Res.*, 6(4), pp.1295–1305.
Lenz, G. et al., 2008. Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. *Proceedings of the National Academy of Sciences of the United States of America*, 105(36), pp.13520–5.
Leopardi, O. et al., 2001. HSP70 is selectively over-expressed in the blast cells of the germinal centres and paracortex in reactive lymph nodes. *Histopathology*, 39, pp.566–571.
Leroy, K. et al., 2002. p53 gene mutations are associated with poor survival in low and low-intermediate risk diffuse large B-cell lymphomas. *Annals of Oncology*, 13, pp.1108–1115.
Levine, A.M., 1992. Acquired immunodefiency syndrome-related lymphoma. *Blood*, 80(1), pp.8–20.
Levine, A.M., Seneviratne, L. & Tulpule, A., 2001. Incidence and management of AIDS-related lymphoma. *Oncology (Williston Park, N.Y.)*, 15(5), pp.629–39;
Li, X., Liu, Z. & Cao, J., 2012. Rituximab in combination with CHOP chemotherapy for the treatment of diffuse large B cell lymphoma in China : a 10-year retrospective follow-up analysis of 437 cases from Shanghai Lymphoma Research Group. *Ann Hematol*, 91, pp.837–845.
Liapis, K. et al., 2013. The microenvironment of AIDS-related diffuse large B-cell lymphoma provides insight into the pathophysiology and indicates possible therapeutic strategies. *Blood*, 122(3), pp.424–33.
Lim, S.-T. et al., 2005. Prognostic factors in HIV-related diffuse large-cell lymphoma: before versus after highly active antiretroviral therapy. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 23(33), pp.8477–82.
Lin, P. et al., 2012. Prognostic Value of MYC Rearrangement in Cases of B-Cell Lymphoma , Unclassifiable , With Features Intermediate Between Diffuse Large B-Cell Lymphoma and Burkitt Lymphoma. *cancer*, 118, pp.1566–1573.
Liu, W. et al., 2011. Down-regulation of HSP70 sensitizes gastric epithelial cells to apoptosis and growth retardation triggered by H . pylori. *BMC Gastroenterology*,

11(1), p.146.
Liu, Y., Zeng, L. & Zhang, S., 2013. Identification of differentially expressed proteins in chemotherapy-sensitive and chemotherapy-resistant diffuse large B cell lymphoma by proteomic methods.
Lodish, H., Berk, A. & Zipursky, S., 2000. Collagen:The Fibrous Proteins of the Matrix. In *Molecular Cell Biology, 4th Edition*. New York: W. H. Freeman, p. Section 22.3.
Londoño, I., Leclerc, Y. & Bendayan, M., 1992. Ultrastructural localization of endogenous albumin in human aortic tissue by protein A-gold immunocytochemistry. *The American journal of pathology*, 140(1), pp.179–91.
Longo, D.L., 2010. *Harrison's Hematology and Oncology*, New York: McGraw-Hill Medical.
López-Guillermo, A. et al., 2005. Diffuse large B-cell lymphoma: clinical and biological characterization and outcome according to the nodal or extranodal primary origin. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 23(12), pp.2797–804.
Lossos, I.S. et al., 2000. Ongoing immunoglobulin somatic mutation in germinal center B cell-like but not in activated B cell-like diffuse large cell lymphomas. *Proceedings of the National Academy of Sciences of the United States of America*, 97(18), pp.10209–13.
Lossos, I.S. & Morgensztern, D., 2006. Prognostic biomarkers in diffuse large B-cell lymphoma. *Journal of clinical oncology*, 24(6), pp.995–1007.
Lu, R. et al., 2013. Lactate Dehydrogenase 5 Expression in Non-Hodgkin Lymphoma Is Associated with the Induced Hypoxia Regulated Protein and Poor Prognosis. *PloS one*, 8(9), p.e74853.
Maartens, G., Celum, C. & Lewin, S.R., 2014. HIV infection: epidemiology, pathogenesis, treatment, and prevention. *The Lancet*, 384(9939), pp.258–271. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0140673614601641.
MacLennan, I.C.M., 1994. Germinal Centers. *Annual Review of Immunology*, 12(1), pp.117–139.
Mafune, K. et al., 1991. Ubiquitin hybrid protein gene expression during human colon cancer progression. *Archives of surgery (Chicago, Ill. : 1960)*, 126(4), pp.462–6.
Magdeldin, S. & Yamamoto, T., 2012. Toward deciphering proteomes of formalin-fixed paraffin-embedded (FFPE) tissues. *Proteomics*, 12(7), pp.1045–58.
Manzo, F. et al., 2009. Histone acetyltransferase inhibitors and preclinical studies. *Expert opinion on therapeutic patents*, 19(6), pp.761–74.
Mao-de, L. & Jing, X., 2007. Ribosomal Proteins and Colorectal Cancer. *Current genomics*, 8, pp.43–49.
Marín-Hernández, A. et al., 2011. Modeling cancer glycolysis. *Biochimica et biophysica acta*, 1807(6), pp.755–67.
Marino-Ramirez, L. et al., 2005. Histone structure and nucleosome stability. *Expert Rev Proteomics.*, 2(5), pp.719–729.
Martinez-balibrea, E. et al., 2009. A proteomic approach links decreased pyruvate kinase M2 expression to oxaliplatin resistance in patients with colorectal cancer and in

human cell lines. *Mol Cancer Ther*, 8(4), pp.771–779.
Matta, A. et al., 2010. Mass spectrometry-based clinical proteomics: head-and-neck cancer biomarkers and drug-targets discovery. *Mass spectrometry reviews*, 29(6), pp.945–61.
Mayeux, R., 2004. Biomarkers: potential uses and limitations. *NeuroRx : the journal of the American Society for Experimental NeuroTherapeutics*, 1(2), pp.182–8.
Mester, B. et al., 2006. Occupation and malignant lymphoma: a population based case control study in Germany. *Occup Environ Med*, 63, pp.17–26.
Mey, U. et al., 2012. Diagnosis and treatment of diffuse large B-cell lymphoma. *Swiss medical weekly*, 142(January), p.w13511.
Meyer, P.N. et al., 2010. Immunohistochemical Methods for Predicting Cell of Origin and Survival in Patients With Diffuse Large B-Cell Lymphoma Treated With Rituximab. *J Clin Oncol*, 29, pp.200–207.
Meyer, P.N. et al., 2011. Immunohistochemical Methods for Predicting Cell of Origin and Survival in Patients With Diffuse Large B-Cell Lymphoma Treated With Rituximab. *Journal of Clinical Oncology*, 29(2), pp.200–207.
Miao, P. et al., 2013. Lactate dehydrogenase A in cancer: a promising target for diagnosis and therapy. *IUBMB life*, 65(11), pp.904–10.
Miller, B.T.P. et al., 1994. P rognostic Significance of the Ki-67 -Associated Proliferative Antigen in Aggressive Non-Hodgkin's Lymphomas: A Prospective Southwest Oncology Group Trial. *Blood*, 83(6), pp.1460–1466.
Milo, R., 2013. Insights & Perspectives What is the total number of protein molecules per cell volume? A call to rethink some published values. , pp.1050–1055.
Mjahed, H. et al., 2012. Heat shock proteins in hematopoietic malignancies. *Experimental cell research*, 318, pp.1946–1958.
Mona, B., Kiran, A. & Shailaja, S., 2014. Anaplastic Variant of Diffuse Large B Cell Lymphoma (DLBCL) – A Rare and Dilemmatic Entity. *World Journal of Surgical Medical and Radiation Oncology*, 3, p.9.
Monroe, J.G. & Silberstein, L.E., 1995. HIV-mediated B-lymphocyte activation and lymphomagenesis. *Journal of clinical immunology*, 15(2), pp.61–8.
Montes-Moreno, S. et al., 2010. Aggressive large B-cell lymphoma with plasma cell differentiation: immunohistochemical characterization of plasmablastic lymphoma and diffuse large B-cell lymphoma with partial plasmablastic phenotype. *Haematologica*, 95(8), pp.1342–9.
Monti, S. et al., 2005. Molecular pro ling of diffuse large B-cell lymphoma identi es robust subtypes including one characterized by host in ammatory response. *Response*, 105(5), pp.1851–1861.
Moonim, M.T. et al., 2010. Identifying HIV infection in diagnostic histopathology tissue samples--the role of HIV-1 p24 immunohistochemistry in identifying clinically unsuspected HIV infection: a 3-year analysis. *Histopathology*, 56(4), pp.530–41.
Müller, A.M.S. et al., 2005. Epidemiology of non-Hodgkin's lymphoma (NHL): trends, geographic distribution, and etiology. *Annals of hematology*, 84(1), pp.1–12.
Muller, P. et al., 2013. C-terminal phosphorylation of Hsp70 and Hsp90 regulates

alternate binding to co-chaperones CHIP and HOP to determine cellular protein folding/degradation balances. *Oncogene*, 32(25), pp.3101–10.
Muris, J.J.F. et al., 2006. Immunohistochemical profiling based on Bcl-2, CD10 and MUM1 expression improves risk stratification in patients with primary nodal diffuse large B cell lymphoma. *The Journal of pathology*, 208(5), pp.714–23.
Murphy, M.E., 2013. The HSP70 family and cancer. *Carcinogenesis*, 34(6), pp.1181–1188.
Nam, S. et al., 2014. Effects of Fixation and Storage of Human Tissue Samples on Nucleic Acid Preservation. *The Korean journal of pathology*, 48, pp.36–42.
Natkunam, Y. et al., 2001. Analysis of MUM1/IRF4 protein expression using tissue microarrays and immunohistochemistry. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc*, 14(7), pp.686–94.
Nayak, P. et al., 2013. Centroblastic variant of diffuse large B-cell lymphoma: Case report and review of literature. *J Oral Maxillofac Pathol*, 17, pp.261–5.
Nguyen, S. & Fenn, J.B., 2007. Gas-phase ions of solute species from charged droplets of solutions. *Proceedings of the National Academy of Sciences of the United States of America*, 104(4), pp.1111–7.
Nishiu, M. et al., 2002. Microarray Analysis of Gene-expression Profiles in Diffuse Large B-cell Lymphoma : Identification of Genes Related to Disease Progression. *Jpn J Cancer Res*, 93, pp.894–901.
Pancholi, V., 2001. Multifunctional alpha-enolase: its role in diseases. *Cell Mol Life Sci*, 58(7), pp.902–20.
Pantanowitz, L., Carbone, A. & Dolcetti, R., 2015. Microenvironment and HIV-related lymphomagenesis. *Seminars in Cancer Biology*, pp.8–13.
Patel, M. et al., 2015. The impact of human immunodeficiency virus infection on cervical preinvasive and invasive neoplasia in South Africa. *Journal of Cancer Therapy*, 6, pp.527–535.
Pather, S. et al., 2013. Large cell lymphoma: correlation of HIV status and prognosis with differentiation profiles assessed by immunophenotyping. *Pathology oncology research : POR*, 19(4), pp.695–705.
Patterson, C. & Hohfeld, J., 2006. Molecular Chaperones and the Ubiquitin – Proteasome System. In R. Mayer, A. Ciechanover, & M. Rechsteiner, eds. *Protein degradation, Vol 2: The Ubiquitin-Proteasome system*. Weinheim: Wiley-VCH, pp. 1–30.
Perry, A.M. et al., 2014. MYC and BCL2 protein expression predicts survival in patients with diffuse large B-cell lymphoma treated with rituximab. *British Journal of Haematology*, 165, pp.382–391.
Pontén, F., Jirström, K. & Uhlen, M., 2008. The Human Protein Atlas--a tool for pathology. *The Journal of pathology*, 216(4), pp.387–93.
Prognosis, P. et al., 2009. Expression of Annexin A5 Is Associated With Higher Tumor. , 43(9), pp.831–837.
Qi, H. et al., 2005. Expression of Heat Shock Protein 70 and 27 in Non-small Cell Lung Cancer and Its Clinical Significance. *Journal of Huazhong University of Science and Technology*, 25(6), pp.693–695.

Ramana, K. V et al., 2013. Elevated activities of serum lactate dehydrogenase in human immunodeficiency virus sero-positive patients in highly active antiretroviral therapy era. *Journal of Dr. NTR University of Health Sciences*, 2(3), pp.162–166.

Rauser, S. et al., 2010. Classification of HER2 receptor status in breast cancer tissues by MALDI imaging mass spectrometry. *Journal of proteome research*, 9, pp.1854–1863.

Reif, K. et al., 2002. Balanced responsiveness to chemoattractants from adjacent zones determines B-cell position. *Nature*, 416.

Renigunta, A. et al., 2011. Cellular Physiology and Biochemistr y Biochemistry The Glycolytic Enzymes Glyceraldehyde 3-Phos- phate Dehydrogenase and Enolase Interact with the Renal Epithelial K + Channel ROMK2 and Regu- late its Function. *Cell Physiol Biochem*, 28, pp.663–672.

Reyzer, M.L. & Caprioli, R.M., 2005. MALDI mass spectrometry for direct tissue analysis: a new tool for biomarker discovery. *Journal of proteome research*, 4(4), pp.1138–42. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16083264.

Rickert, R.C., 2013. New insights into pre-BCR and BCR signalling with relevance to B cell malignancies. *Nature reviews Immunology*, 13, pp.578–591.

Rohde, M. et al., 2005. Members of the heat-shock protein 70 family promote cancer cell growth by distinct mechanisms. *Genes & development*, 19(5), pp.570–82.

Ronci, M. et al., 2008. Protein unlocking procedures of formalin-fixed paraffin-embedded tissues: application to MALDI-TOF imaging MS investigations. *Proteomics*, 8(18), pp.3702–14.

Rosenberg, S.A. & Kaplan, H.S., 1970. Hodgkin's Disease and Other Malignant Lymphomas. *The Western Journal of Medicine*, 130(4), pp.23–38.

Rosenwald, A. et al., 2002. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *The New England journal of medicine*, 346(25), pp.1937–47.

Rosolen, A. et al., 2000. Non Hodgkin lymphoma in adolescents and young adults.

Rovira, J. et al., 2015. Prognosis of patients with diffuse large B cell lymphoma not reaching complete response or relapsing after frontline chemotherapy or immunochemotherapy. *Ann Hematol*, 94, pp.803–812.

Rüetschi, U. et al., 2015. SILAC-Based Quantitative Proteomic Analysis of Diffuse Large B-Cell Lymphoma Patients. *International journal of proteomics*, 2015.

Sagaert, X. et al., 2006. Forkhead Box Protein P1 Expression in Mucosa-Associated Lymphoid Tissue Lymphomas Predicts Poor Prognosis and Transformation to Diffuse Large B-Cell Lymphoma. *Journal of Clinical Oncology*, 24(16), pp.2490–2497.

Salles, G. et al., 2011. Prognostic significance of immunohistochemical biomarkers in diffuse large B-cell lymphoma: a study from the Lunenburg Lymphoma Biomarker Consortium. *Blood*, 117(26), pp.7070–8.

Sarkozy, C. et al., 2014. Impact of BMI and Gender on Outcomes in DLBCL Patients Treated with R-CHOP: A Pooled Study from the LYSA. *Lymphoma*, 2014.

Schey, K.L., Anderson, D.M. & Rose, K.L., 2013. Spatially-Directed Protein Identi fi

cation from Tissue Sections by Top- Down LC-MS/MS with Electron Transfer Dissociation. *Anal. Chem.*, 85, pp.6767–6774.
Schnoor, M. et al., 2008. Production of Type VI Collagen by Human Macrophages : A New Dimension in Macrophage Functional Heterogeneity 1 , 2.
Schober, Y. et al., 2011. Protein identi fi cation by accurate mass matrix-assisted laser desorption / ionization imaging of tryptic peptides. *Rapid communications in mass spectrometry*, 25, pp.2475–2483.
Schöne, C., Höfler, H. & Walch, A., 2013. MALDI imaging mass spectrometry in cancer research: combining proteomic profiling and histological evaluation. *Clinical biochemistry*, 46(6), pp.539–45.
Seeley, E.H. & Caprioli, R.M., 2011. MALDI imaging mass spectrometry of human tissue: method challenges and clinical perspectives. *Trends in biotechnology*, 29(3), pp.136–43.
Seeley, E.H. & Caprioli, R.M., 2008. Molecular imaging of proteins in tissues by mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America*, 105(47), pp.18126–31.
Sehn, L.H. et al., 2007. The revised International Prognostic Index (R-IPI) is a better predictor of outcome than the standard IPI for patients with diffuse large B-cell lymphoma treated with R-CHOP. *Blood*, 109, pp.1857–1861.
Shaffer, A.L., Rosenwald, A. & Staudt, L.M., 2002. Lymphoid malignancies : the dark side of b-cell differentiation. *,Immunology* 2: 1-13.
Shenoy, N. et al., 2012. Alterations in the ribosomal machinery in cancer and hematologic disorders. ???, 5(1), p.1.
Sherman, M.Y. et al., 2007. Molecular chaperones regulate p53 and suppress senescence programs. *FEBS letters*, 581(19), pp.3711–5.
Shiio, Y. & Aebersold, R., 2006. Quantitative proteome analysis using isotope-coded affinity tags and mass spectrometry. , 1(1), pp.139–145.
Shim, H. et al., 2013. Prognostic impact of concordant and discordant cytomorphology of bone marrow involvement in patients with diffuse , large , B-cell lymphoma treated with R-CHOP. , pp.420–425.
Shipp, M. et al., 1993. A predictive model for aggressive non-Hodgkin's lymphoma: The international non-Hodgkin's lymphoma prognostic factors project. *N engl J Med*, 329, pp.987–94.
Shipp, M. & The international non-Hodgkins lymphoma prognostic factors, 1993. A predictive model for aggressive non-Hodgkin's lymphoma. *N Engl J Med*, 329, pp.987–94.
Shisana, O. et al., 2012. *South African National HIV Prevalence, Incidence and Behaviour Survey, 2012*, Cape Town.
Shoulders, M. & Rainers, R., 2010. Collagen structure and stability. *Annu. Rev. Immunol.*, 78, pp.929–958.
Shustik, J. et al., 2010. Correlations between BCL6 rearrangement and outcome in patients with diffuse large B-cell lymphoma treated with CHOP or R-CHOP. *Haematologica*, 95(1), pp.96–101.

Silverberg, M.J. et al., 2012. Effect of immunodeficiency and tumor marker expression on HIV-related diffuse large B-cell lymphoma prognosis. *Infectious Agents and Cancer*, 7(Suppl 1), p.P47.
Simard, E.P. & Engels, E. a, 2010. Cancer as a cause of death among people with AIDS in the United States. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 51(8), pp.957–62.
Sissolak, G. et al., 2013. Transfusion and Apheresis Science Tissue microarray in a subset of South African patients with DLBCL. *Transfusion and Apheresis Science*, 49(2), pp.120–132.
Sohn, S.K. et al., 2003. Prognostic Significance of bcl-2 , bax , and p53 Expression in Diffuse Large B-Cell Lymphoma. , 107(September 2002), pp.101–107.
Sompuram, S.R. et al., 2004. A Molecular Mechanism of Formalin Fixation and Antigen Retrieval. , pp.190–199.
Song, Y. et al., 2014. Alpha-enolase as a potential cancer prognostic marker promotes cell growth , migration , and invasion in glioma. *Molecular Cancer*, 13(1), pp.1–12. Available at: Molecular Cancer.
Sousa, J.F. et al., 2012. Proteomic Profiling of Paraffin-Embedded Samples Identifies Metaplasia-Specific and Early-Stage Gastric Cancer Biomarkers. , 181(5), pp.1560–1572.
Staudt, L.M. & Dave, S., 2005. The biology of human lymphoid malignancies revealed by gene expression profiling. *Adv. Immunol*, 87, pp.163–208.
Stefancikova, L. et al., 2011. Prognostic impact of p53 aberrations for R-CHOP- treated patients with diffuse large B-cell lymphoma. *Internation journal of oncology*, 39, pp.1413–1420.
Stevenson, F. & et al, 2001. The occurrence and significance of V gene mutations in B cell-derived human malignancy. *Adv. Cancer Res.*, 83, pp.81–116.
Tainsky, M. a, 2009. Genomic and proteomic biomarkers for cancer: a multitude of opportunities. *Biochimica et biophysica acta*, 1796(2), pp.176–93.
Takadate, T. et al., 2013. cancer identified by coupled shotgun and targeted proteomics. *Int. J. Cancer*, 132, pp.1368–1382.
Tanaka, M. et al., 2014. Hsc70 contributes to cancer cell survival by preventing Rab1A degradation under stress conditions. *PloS one*, 9(5), p.e96785.
Tang, Z., 2012. Over-expression of GAPDH in human colorectal carcinoma as a prefered target of 3-Bromopyruvate propyl ester. *J Bioenerg Biomembr.*, 44(1), pp.117–125.
Tao, L. et al., 2014. Socioeconomic disparities in mortality after diffuse large B-cell lymphoma in the modern treatment era. *Blood*, 123(23), pp.3553–3563.
Tedeschi, R. et al., 2012. Assessment of immunovirological features in HIV related non-Hodgkin lymphoma patients and their impact on outcome. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology*, 53(4), pp.297–301.
Tilly, H. & Dreyling, M., 2009. Diffuse large B-cell non-Hodgkin's lymphoma: ESMO clinical recommendations for diagnosis, treatment and follow-up. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO*,

20 Suppl 4(Supplement 4), pp.110–2.
Tirelli, U. et al., 2000. Epidemiological, biological and clinical features of HIV-related lymphomas in the era of highly active antiretroviral therapy. *AIDS (London, England)*, 14(12), pp.1675–1688.
Tsubata, T., 2005. B cell abnormality and autoimmune disorders. *Autoimmunity*, 38(5), pp.331–7.
Tzankov, A. et al., 2015. Multiparameter analysis of homogeneously R-CHOP-treated diffuse large B cell lymphomas identifies CD5 and FOXP1 as relevant prognostic biomarkers : report of the prospective SAKK 38 / 07 study. *Journal of Hematology & Oncology*, pp.1–11.
Uhlén, M. et al., 2005. A human protein atlas for normal and cancer tissues based on antibody proteomics. *Molecular & cellular proteomics : MCP*, 4(12), pp.1920–32.
UNAIDS, 2015. *The gap report*,
Vega, F. & Medeiros, L.J., 2003. Chromosomal Translocations Involved in Non-Hodgkin Lymphomas. *Arch Pathol Lab Med*, 127, pp.1148–60.
Vinuesa, C.G., 2013. T-cell subsets in the germinal. *Immunological reviews*, 252, pp.146–155.
Vishnu, P. & Aboulafia, D.M., 2012. AIDS-Related Non-Hodgkin's Lymphoma in the Era of Highly Active Antiretroviral Therapy. *Advances in hematology*, 2012, p.485943.
Volkova, M. & Russell, R., 2011. Anthracycline Cardiotoxicity : Prevalence , Pathogenesis and Treatment. *current Cardiology Reviews*, 7, pp.214–220.
Voropaeva, E.N. et al., 2015. Prognostic impact of the TP53 rs1625895 polymorphism in DLBCL patients. *British journal of haematology*, 169, pp.32–35.
Walch, A. et al., 2008. MALDI imaging mass spectrometry for direct tissue analysis: a new frontier for molecular histology. *Histochemistry and cell biology*, 130(3), pp.421–34.
Wang, C.-C. & Castillo, J.J., 2011. Management of HIV-associated lymphomas. *Medicine and health, Rhode Island*, 94(1), pp.4–6.
Wang, H. et al., 2006. proliferation in gastric cancer. , 8(Dcc), pp.1–8.
Wehner, P., Nielsen, B. & Hokland, M., 2003. Expression levels of hsc70 and hsp60 are developmentally regulated during B-cell maturation and not associated to childhood c-ALL at presentation or relapse. *Eur J Haematol*, 71(14), pp.100–108.
Westin, J.R. & Fayad, L.E., 2009. Beyond R-CHOP and the IPI in Large- Cell Lymphoma : Molecular Markers as an Opportunity for Stratifi cation. *Current Hematologic Malignancy Reports*, 4, pp.218–224.
Wiggill, T.M. et al., 2011. Changing pattern of lymphoma subgroups at a tertiary academic complex in a high-prevalence HIV setting: a South African perspective. *Journal of acquired immune deficiency syndromes (1999)*, 56(5), pp.460–6.
Wiggill, T.M., Mayne, E.S. & Willem, P., 2013. Challenges in lymphoma diagnosis in HIV positive patients in the South African setting. *Transfusion and apheresis science : official journal of the World Apheresis Association : official journal of the European Society for Haemapheresis*, 49(2), pp.157–62. Available at:

http://www.sciencedirect.com/science/article/pii/S1473050213002474.
Wilen, C., Tilton, J. & Doms, R., 2012. HIV: cell binding and entry. *Cold Spring Harb Perspect Med.*, 2(8), p.a006866.
Williams, I. & Churchill, D., 2007. Management of advanced HIV disease with no other complications in women and in Africans. *International journal of clinical practice*, 61(1), pp.119–129.
de Witt, P. et al., 2014. Treatment outcomes in AIDS-related diffuse large B-cell lymphoma in the setting roll-out of combination antiretroviral therapy in South Africa. *J Acquir Immune Defic Syndr*, 64(1), pp.66–73.
World Health Organisation, 2007. *WHO case definitions of hiv for surveillance and revised clinical staging and immunological classification of hiv-related disease in adults and children*, World Health Organisation, 2013. *World health statistics 2013*,
Xie, Y. et al., 2014. p53 Expression Is a Strong Marker of Inferior Survival in De Novo Diffuse Large B-Cell Lymphoma and May Have Enhanced Negative Effect With MYC Coexpression A Single Institutional Clinicopathologic Study. , pp.593–604.
Xie, Y., Pittaluga, S. & Jaffe, E.S., 2015. The Histological Classi fi cation of Diffuse Large B-cell Lymphomas. *Seminars in Hematology*, 52(2), pp.57–66. A
Xu-monette, Z.Y. et al., 2012. Mutational profile and prognostic significance of TP53 in diffuse large B-cell lymphoma patients treated with R-CHOP : report from an International DLBCL Rituximab-CHOP Consortium Program Study. *Blood*, 120(19), pp.3986–3997.
Yıldırım, M. et al., 2015. Systematic review/Meta-analysis The role of gender in patients with diffuse large B cell lymphoma treated with rituximab-containing regimens: a meta-analysis. *Arch Med Sci*, 11(4), pp.708–714.
Yoo, B.C. et al., 2004. Decreased pyruvate kinase m2 activity linked to cisplatin. *Int. J. Cancer*, 108, pp.532–539.
Yoon, D.H. et al., 2010. Ki-67 expression as a prognostic factor in diffuse large B-cell lymphoma patients treated with rituximab plus CHOP. *European journal of haematology*, 85, pp.149–157.
Young, J.C., Barral, M. & Hartl, F.U., 2003. More than folding : localized functions of cytosolic chaperones. *Trends in Biochemical Sciences*, 28(10), pp.541–547.
Young, K.H. et al., 2008. Structural profiles of TP53 gene mutations predict clinical outcome in diffuse large B-cell lymphoma: an international collaborative study. *Blood*, 112(8), pp.3088–3098.
Yu, B., Zhou, X. & Li, B., 2011. FOXP1 expression and its clinicopathologic significance in nodal and extranodal diffuse large B-cell lymphoma. *Ann Hematol*, 90, pp.701–708.
Zainuddin, N. et al., 2009. TP53 mutations predict for poor survival in de novo diffuse large B-cell lymphoma of germinal center subtype. *Leukemia Research*, 33, pp.60–66.
Zhang, X.X. et al., 2007. Serum diagnosis of diffuse large B-cell lymphomas and further identification of response to therapy using SELDI-TOF-MS and tree analysis patterning. *BMC cancer*, 7, p.235.

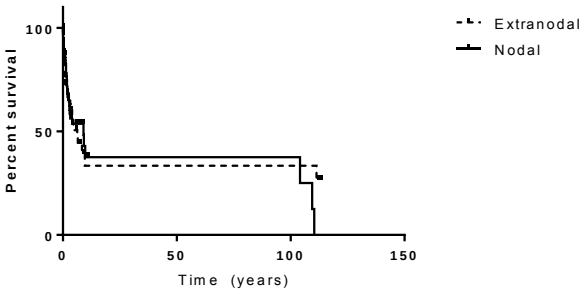
Zhou, D. et al., 2013. Prognostic Values of Various Clinical Factors and Genetic Subtypes for Diffuse Large B-cell lymphoma Patients : A Retrospective Analysis of 227 Cases. , 1, pp.929–934.
Zhou, M. & Veenstra, T., 2013. *Proteomics for Biomarker Discovery: Methods and protocols* J. M. Walker, ed., Humana Press.
Zhou, X. et al., 2015. Ribosomal proteins : functions beyond the ribosome. *Journal of Molecular Cell Biology*, 7(2), pp.92–104.
Zhou, Z. et al., 2014. An enhanced International Prognostic Index (NCCN-IPI) for patients with diffuse large B-cell lymphoma treated in the rituximab era. *Blood*, 123(6), pp.837–843.

APPENDICES

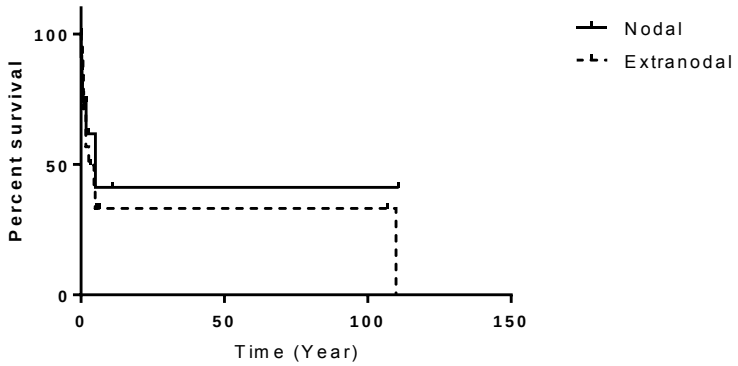
APPENDIXA: LINEAR REGRESSION CODES

Codes for Linear Regression	0	1	2
Survival status	Alive	Death	
Stage	I/II	III/IV	
LDH	High	Normal	
Gender	Female	Male	
IPI scores	0/1	2/3	4/5
CD4 count	<200	>/=200	
Nodal/extranodal	Extranodal	Nodal	
B symptoms	No B symptoms	B symptoms	
HIV status	negative	positive	
Bone marrow involvement	involved	clear	

APPENDIX B: CHAPTER 2 SUPPLEMENTARY RESULTS

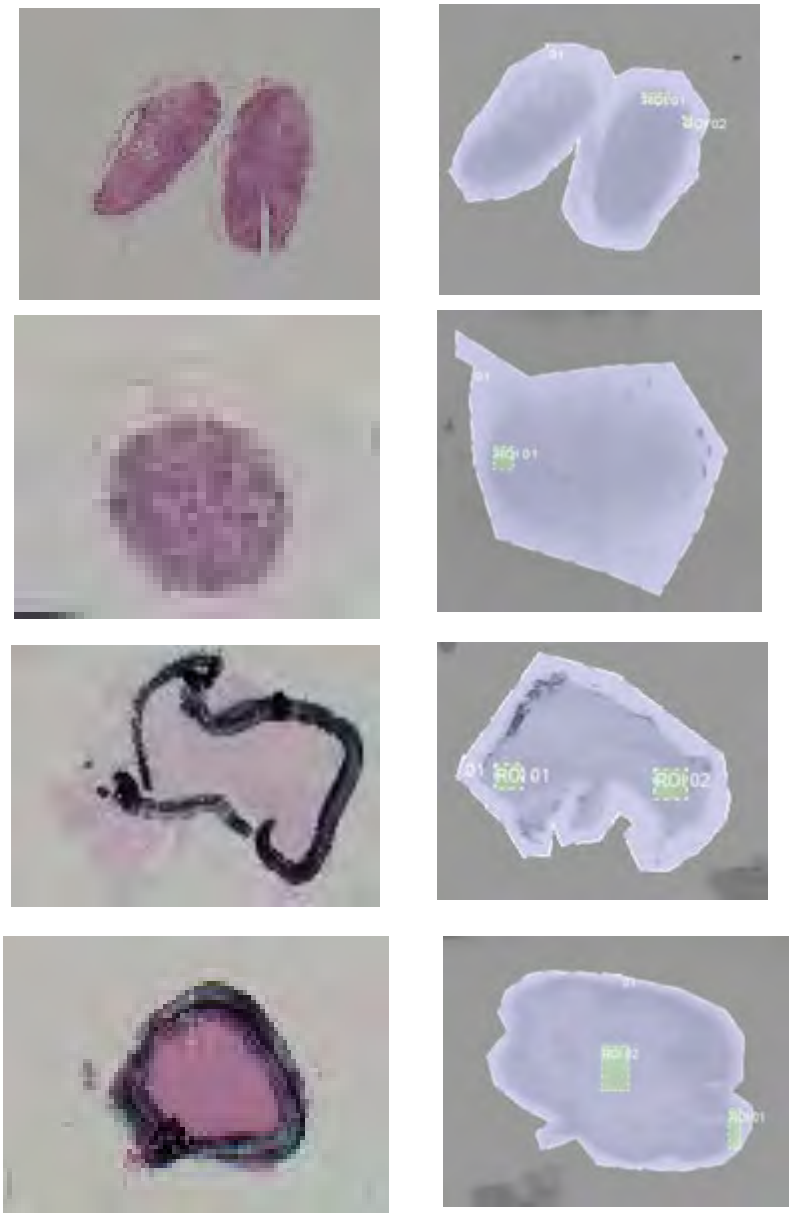


Kaplan Meier survival plot of HIV negative DLBCL patients presenting in nodal and extranodal sites



Kaplan Meier survival plot of HIV positive DLBCL patients presenting in nodal and extranodal sites

APPENDIX C: ROIS DRAWN IN MALDI IMS ANALYSIS



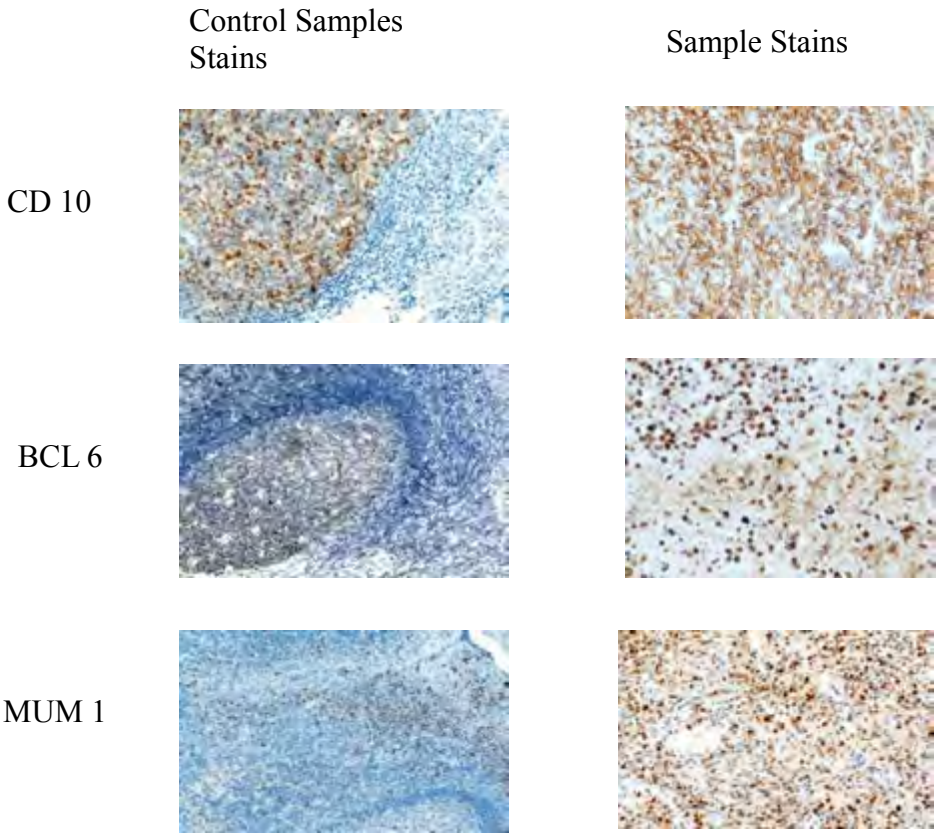
APPENDIX D: HAN'S CLASSIFICATION ALGORITHM ANTIBODIES

The experimental conditions for the antibodies used in the Hans classification algorithm.

Antigen retrieval buffer	Antigen	Clone	Clonality	Supplier	Antibody dilution
1mM Tris 1mM EDTA, pH 9	CD10	56C6	Monoclonal	Dako	1:40,1 hr
1mM Tris 1mM EDTA, pH 9	BCL6	PG-B6p	Monoclonal	Dako	1:20,1 hr
1mM Tris 1mM EDTA, pH 9	MUM1	MUM1p	Monoclonal	Dako	1:50,1 hr

Assessment of immunohistochemical stains

A qualitative assessment was used to classify cases as positive or negative. Cases were considered positive, if 30%tumour cells were stained with an antibody (Hans et al 2004). According to the College of American Pathologists guidelines, weak and patchy nuclear staining in the tumour was also designated positive.



Representative of positive stains of CD10, BCL6 and MUM 1 of the Hans classification.

APPENDIX E: OTHER PROTOCOLS

Slide coating protocol

- Mix 1:1 poly-L-lysine with water in an Eppendorf tube
- Add 1 µl of IGEPAL per 1.5 ml
- Streak out 2 drops of a Pasteur pipette onto an ITO coated slide using a tongue depressor
- Dry slide for 15 min on 60°C

Zip Tip Method

- Wet Zip tip in 20 µl ACN (aspirating and dispensing 10X)
- Equilibrate zip tip in 20 µl of solvent A (0.1% TFA)
- Bind Sample (2 µl of sample and 8 µl solvent A)
- Wash zip tip in solvent A
- Elute in 10 µl Solvent B (50 % ACN; 0.1 % TFA)

APPENDIX F: COMPOSITION OF SOLUTIONS

Mayers Hematoxylin solution

Haematoxylin	1 g
Potassium/ammonium alumminium	50 g
Sodium iodate	0,2 g
Citric acid	1 g
Chloral hydrate	50 g
Distlled water	1000 ml

Dissolve the haematoxylin, potassium alumminium and sodium iodate in water. Add chloral hydrate and citric acid. Boil for 5 min. Cool and filter.

Ammoniated water

Ammonium hydroxide (concentrated)	2 ml
Distilled water	1000 ml
Mix well.	

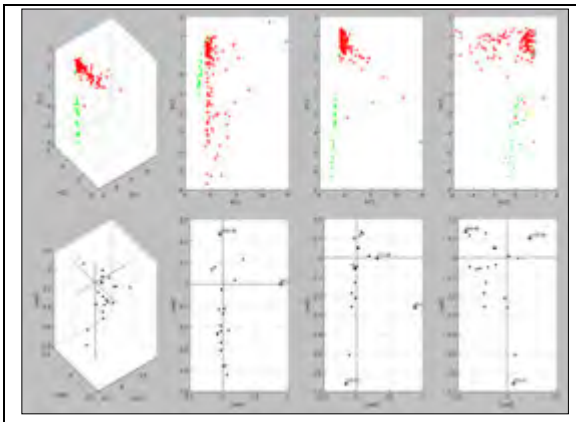
Phloxine-eosine solution

1 % Phloxine B	10 ml
1 % Eosin Y	100 ml
95 % Alcohol	780 ml
Glacial acetic acid	4 ml

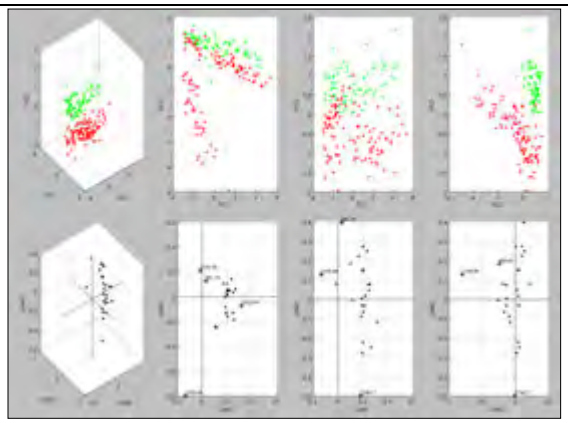
APPENDIX G: THE DEMOGRAPHICAL CHARACTERISTICS OF THE PATIENTS USED IN MALDI MSI ANALYSIS

	Reactive Nodes		**DLBCL**	
	HIV Negative Group	**HIV Positive Group**	**HIV Negative Group**	**HIV Positive Group**
Mean Age	47.5	40	56.25	39.25
N	2	2	8	8
Gender (F/M)	0/2	0/2	5/3	5/3
Subtype (GC/non-GC)	N/A	N/A	4/4	4/4
Tissue Origin	Lymph node		Gastric, nasopharynx, soft tissue of head and neck, Axilla, lymph node, tonsil, femur, kidney	

APPENDIX H: PCA RESULTS

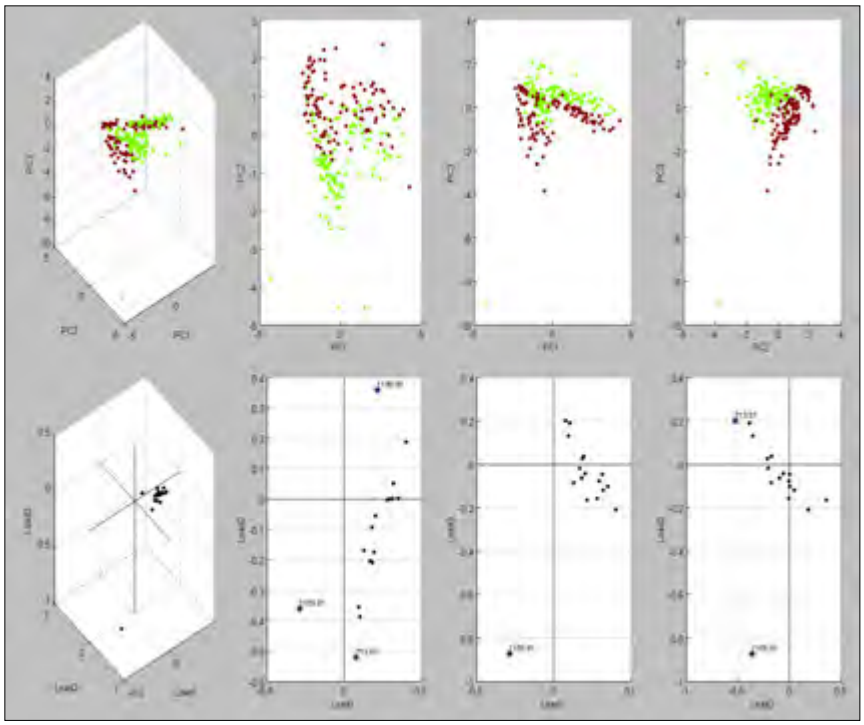


A



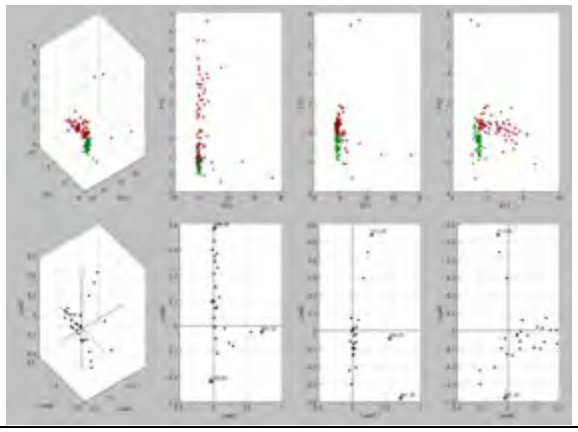
B

PCA analysis of reactive nodes (green) and DLBCL (red) cases in HIV negative (Panel A) and HIV positive patients (Panel B).

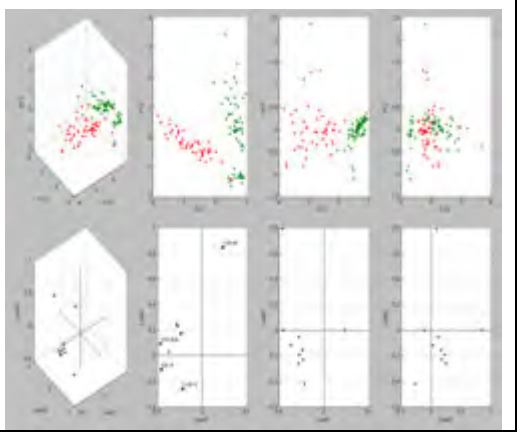


PCA of DLBCL cases from HIV negative (green) and HIV positive (red) patients.

A

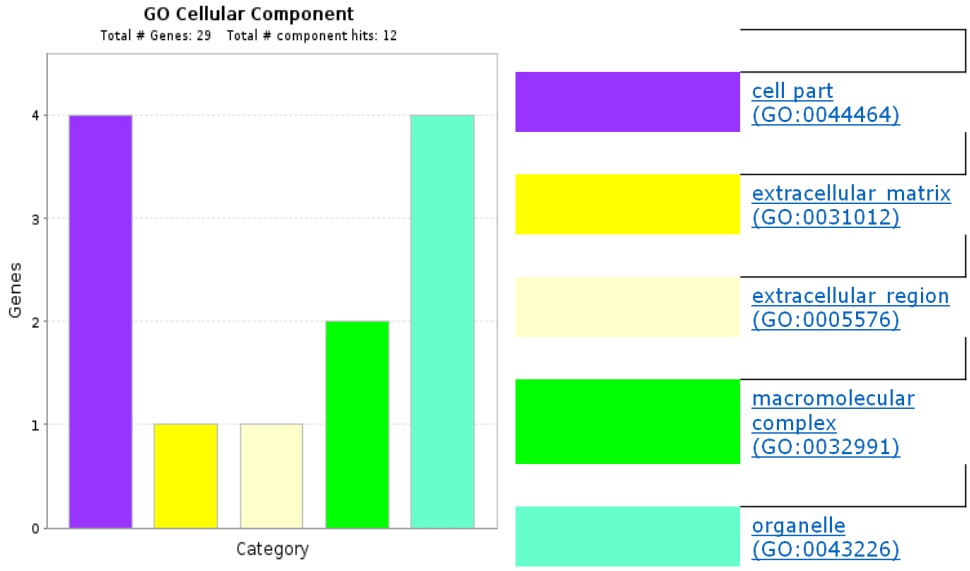


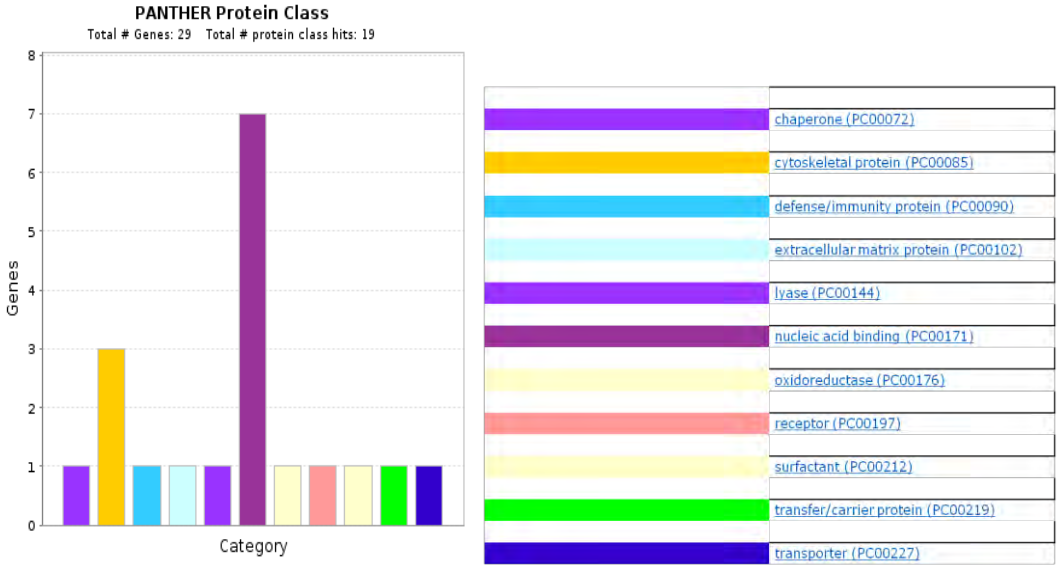
B

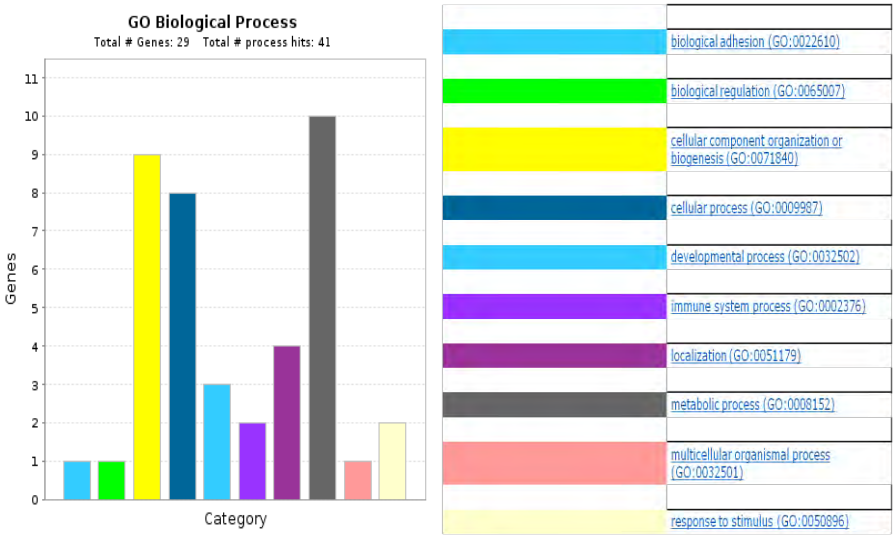


PCA analysis of germinal centre (GC) (green) and non- GC (red/maroon) DLBCL subtypes in HIV negative (Panel A) and positive (Panel B) cases.

APPENDIX I: PANTHER ANALYSIS






GO Biological Process
Total # Genes: 29 Total # process hits: 41
Genes
Category
0
1
2
3
4
5
6
7
8
9
10
11
biological adhesion (GO:0022610)
biological regulation (GO:0065007)
cellular component organization or biogenesis (GO:0071840)
cellular process (GO:0009987)
developmental process (GO:0032502)
immune system process (GO:0002376)
localization (GO:0051179)
metabolic process (GO:0008152)
multicellular organismal process (GO:0032501)
response to stimulus (GO:0050896)

APPENDIX J: IHC RESULTS FOR ANTI-HSP70 ANTIBODY

Patient #	HIV status	DLBCL subtype	Hsp70	p24	Clinical outcome
1	neg	RN	neg	neg	N/A
2	neg	RN	neg	neg	N/A
3	pos	RN	neg	neg	N/A
4	pos	RN	pos*	pos	N/A
5	neg	ABC	pos	neg	deceased
6	neg	ABC	pos*	neg	alive
7	neg	GC	pos*	neg	deceased
8	neg	GC	neg	neg	alive
9	pos	ABC	neg	neg	deceased
10	pos	GC	pos	neg	deceased
11	neg	RN	pos*	ND	N/A
12	neg	RN	pos*	ND	N/A
13	neg	ABC	pos	ND	relapsed
14	neg	ABC	neg	ND	alive
15	neg	ABC	pos*	ND	deceased
16	neg	ABC	pos	ND	deceased
17	neg	ABC	pos*	ND	relapsed
18	neg	ABC	neg	ND	deceased
19	neg	ABC	pos	ND	relapsed
20	neg	GC	neg	ND	alive
21	neg	GC	pos	ND	relapsed
22	neg	unknown	pos*	ND	deceased
23	neg	unknown	neg	ND	relapsed
24	neg	unknown	neg	ND	alive
25	neg	unknown	pos*	ND	deceased
26	neg	unkown	neg	ND	alive
27	neg	unkown	neg	ND	relapsed
28	neg	unkown	neg	ND	deceased
29	neg	unkown	pos	ND	deceased
30	neg	unkown	neg	ND	deceased
31	pos	ABC	neg	ND	deceased
32	pos	ABC	pos*	ND	relapsed
33	pos	ABC	pos	ND	alive
34	pos	ABC	neg	ND	alive
35	pos	ABC	pos	ND	deceased

36	pos	GC	pos	ND	deceased
37	pos	GC	neg	ND	deceased
38	pos	GC	neg	ND	alive
39	pos	unknown	neg	ND	deceased
40	pos	unknown	neg	ND	deceased
41	pos	unknown	pos	ND	deceased
42	pos	ABC	neg	ND	deceased
43	neg	unknown	neg	ND	alive
44	neg	RN	pos*	ND	
45	neg	GC	neg	ND	deceased
46	neg	ABC	neg	ND	deceased
47	pos	GC	neg	ND	deceased
48	neg	ABC	pos*	ND	alive
49	neg	ABC	neg	ND	alive
50	neg	GC	pos*	ND	alive
51	neg	ABC	pos*	ND	alive
52	pos	unknown	neg	ND	alive

Note: The highlighted cases were used in the MALDI IMS analysis

APPENDIX K: THE PEPTIDES IDENTIFIED IN THE STUDY

MALDI IMS ion	nLC -MS ion	Peptide Sequence	Identification (Accession)
703.516			
704.51	704.4049	K.GVPLYR.H	ENOA_HUMAN
713.55	714.35	R.TLYGFGG.-	H4_HUMAN
757.61	758.48	K.LILFPR.K	RL13_HUMAN
788.727	788.4805	R.KLPFQR.L	H31T_HUMAN
789.7218	789.4083	K.KFGGPGAR.A	RS16_HUMAN
801.739			
805.739	805.43	K.VGVNGFGR.I	G3P_HUMAN
806.729	806.4464	K.YNQLLR.I	ENOA_HUMAN
816.729	816.45	R.EIQTAVR.L	H2B1K_HUMAN
836.64	836.43	R.GPAGPQGP.R	CO1A1_HUMAN
848.716	848.4622	R.FVEQVAR.R	CO6A2_HUMAN
849.72			
850.761	850.5275	R.HLQLAIR.N	H2A3_HUMAN
861.28			
898.81	898.4790	K.QGYVIYR.I	RL15_HUMAN
927.84	927.5104	K.AGPIWDLR.L	ALBU

928.805			
929.809			
944.860	944.5346	R.AGLQFPVGR.V	H2A3_HUMAN
945.845			H2A3_HUMAN
957.85	957.5291	R.TIGISVDPR.R	RL13_HUMAN
966.845			
971.76			
976.81	976.45	K.AGFAGDDAPR.A	ACTB
978.8			
988.73			
1032.93	1032.5988	R.YRPGTVALR.E	H31T_HUMAN
1039.88	1039.5168	K.EGIPPDQQR.L	RL40_HUMAN
1105.93	1106.5737	R.SEIDLFNIR.K	ANXA5
1199.17	1199.6801	K.DAGTIAGLNVLR.I	HSP7C_HUMAN
1287.15	1286.6723	R.EPQVYTLPPSR.D	IGHG1_HUMAN
1350.15	1353.7517	K.NKTGAAPIIDVVR.S	RL15_HUMAN
1378.18	1378.63	R.GGGGNFGPGPGSNFR.G	ROA2_HUMAN
1506.33	1507.85	R.APIIAVTRNPQTAR.Q	KPYM_HUMAN
1792.39	1790.9083	K.SYELPDGQVITIGNER.F	ACTG

APPENDIX L: THE VIRAL DATABASE SEARCH HITS

MALDI IMS ion	nLC -MS ion	Peptide Sequence	Identification (Accession)
788.727	788.4805	K.KMPLQR.F	R1AB_BYVU [Replicase polyprotein 1ab (Beet yellow virus)]
944.5346	R.AGLQFPVGR.V	R.LEVSAGGRR.N	074L_FRG3G [Uncharacterised protein 074L (Frog virus 3)]
976.81	976.45	K.AGFAGDDAPR.A	FGR_FSVGR [Tyrosine protein kinase transforming protein Fgr (Feline sarcoma virus)]

APPENDIX M: REAGENT SUPPLIERS

Name	Supplier
Trifluoroacetic acid	Sigma
Acetonitrile	Merck
Water, proteomics grade	Fluka
HCCA matrix	Bruker
Ethanol Absolute	Millipore
H2PO4N	Sigma
Trypsin, modified 100 µg	Promega
Zip tip C18 column	Merck
Water, molecular biology grade	Sigma
Acetonitrile	Sigma
Trypsin, modified 100 µg	Promega
ITO slides	Sigma
HCCA	Bruker
Poly-L-Lysine	Sigma
IGEPAL	Sigma
Ammonium Bicarbonate	Sigma
ITO coated slides	Sigma
Xylene	Millipore

Ethanol, Absolute	Illovo
Ethanol, proteomics grade	Millipore
Water, proteomics grade	Fluka
Xylene	Lasec
Ethanol (crude)	Kimix
Entelin DPX	Merck
Tris	Sigma
EDTA	Merck
PBS tablets	Quantum Biotechnologies
Tween	Sigma
Hydrogen peroxidase	Merck
Envision anti- mouse HRP labelled polymer kit	Dako
Liquid DAB chromogenic system	Dako

APPENDIX N: FISCHER'S EXACT CORRELATION TEST FOR HSP70

Correlation of hsp70 expression with outcome in ALL DLBCL cases

Hsp70	alive	deceased	relapsed
neg	14	16	4
pos	1	7	3

Fischer's exact test, P= 0.096

Correlation of Hsp70 expression with outcome in HIV negative DLBCL cases

Hsp70	alive	deceased	relapsed
neg	11	9	3
pos	0	3	3

Fischer's exact test, P= 0.033

Correlation of Hsp70 expression with outcome in HIV positive DLBCL cases

Hsp70	alive	deceased	relapsed
neg	3	7	1
pos	1	4	0

Fischer's exact test, P= 1.0

Correlation of Hsp70 expression in plasma cells (pos*) with outcome in ALL DLBCL cases

Hsp70	alive	deceased	relapsed
neg	10	12	2
pos	1	7	3
Pos*	4	4	2

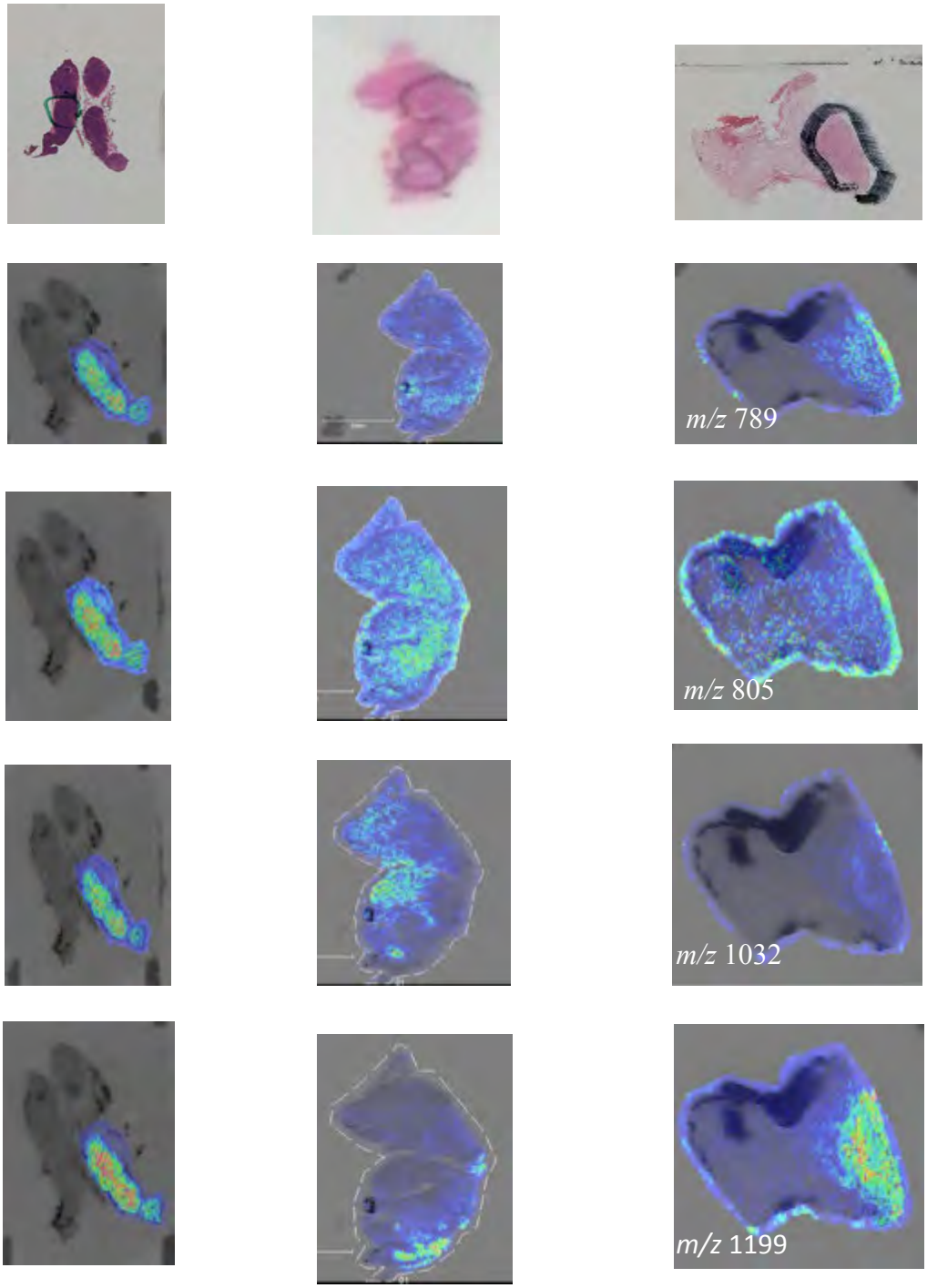
Fischer's exact test, P= 0.228

Correlation of Hsp70 expression in plasma cells (pos*) with outcome in HIV negative DLBCL cases

Hsp70	alive	deceased	relapsed
neg	7	5	2
pos	0	3	3
Pos*	4	4	1

Fischer's exact test, P= 0.191

APPENDIX O: DISTRIBUTION OF PEPTIDES IN OTHER TISSUES



The distribution of the selected m/z ions in RN (first panel), non-GC subtype (middle panel) and GC subtype (last panel) of HIV positive DLBCL cases.

APPENDIX P

Hereditary Genetics: Current Research

Phillips, et al., Hereditary Genet 2015, 4:2
http://dx.doi.org/10.4172/2161-1041.1000143

Research Article Open Access

MicroRNA 21 Expression Levels in HIV Negative and HIV Positive Diffuse Large B Cell Lymphoma

Phillips P[1], Mograth E[1], Sekar D[2], Sookhayi R[1], Govender D[1], Mohamed Z[3], Zerbini LF[2] and Naidoo R[1]

[1]*Division of Anatomical Pathology, Department of Clinical Laboratory Sciences, Faculty of Health Sciences, University of Cape Town/National Health Laboratory Service, Groote Schuur Hospital, South Africa*

[2]*Cancer Genomics Group, International Centre for Genetic Engineering and Biotechnology (ICGEB), South Africa*

[3]*Department of Radiation Oncology, University of Cape Town, Groote Schuur Hospital, South Africa*

***Corresponding author:** Richard Naidoo, Division of Anatomical Pathology, Department of Clinical Laboratory Sciences Faculty of Health Sciences University of Cape Town Anzio Road, Observatory7925, Cape Town, Western Cape, South Africa, Tel: 27214047615, Fax: 27214047611, E-mail: Richard.Naidoo@uct.ac.za

Rec Date: February 23, 2015, **Acc Date:** March 11, 2015, **Pub Date:** March 14, 2015

Abstract

Diffuse large B cell lymphoma (DLBCL) is a heterogeneous disease with various morphological and molecular subtypes. It commonly occurs in the elderly as well as in people infected with HIV. According to the Hans' algorithm DLBCL can be classified prognostically into the favourable germinal centre (GC) subtype and the unfavourable non-GC subtype. The disease tends to be more aggressive in HIV infected individuals. Research on DLBCL has been largely conducted on HIV negative samples and it is still unclear how HIV affects the molecular mechanisms of the disease. We conducted a study to compare the expression levels of miR-21 in HIV negative and HIV positive DLBCL patients. Our results suggest that miR-21 expression levels are higher in HIV positive cases of DLBCL. We did not observe any statistically significant difference in miR-21 levels between DLBCL subtypes in both HIV negative and positive cases. Favourable prognosis in HIV negative patients was associated with high miR-21 expression and the reverse was true in HIV positive patients. In conclusion, our results indicate that HIV infection may affect the expression of miR-21 in DLBCL. This highlights the need for further studies on HIV positive DLBCL as the prognostic significance of miR-21 expression level may differ depending on the HIV status.

Keywords: B-cell lymphoma; Micro RNA; Prognosis; HIV; miR-21; DLBCL

Introduction

Diffuse large B cell lymphoma (DLBCL) is the most common subtype of non-Hodgkin lymphoma (NHL) comprising of 30% of all NHL [1]. The disease shows marked heterogeneity, with morphological and molecular variants [2,3], making it difficult to treat. DLBCL is divided into two histological subtypes with prognostic significance according to the Hans' algorithm. The non-germinal centre (non-GC) subtype has a worse prognosis than the germinal centre B cell like (GC) subtype [4].

Recent data suggests that HIV gp120 induces B cells expressing mannose C type lectin receptors (MCLRs) to undergo immunoglobulin class switching by increasing the expression of activation-induced cytidine deaminase [5]. Chronic B-cell activation in HIV infection is driven by the production of B-cell stimulating cytokines which are secreted by gp120 bound monocytes [6]. These cytokines act by upregulating MCLRs on the surface of B cells, thus creating chronic B cell activation [5]. Therefore, people infected with HIV are at an increased risk of developing B cell NHL [7]. DLBCL is the most common NHL diagnosed in HIV infected individuals [8,9].

MicroRNAs are non-coding RNA molecules that silence the expression of target genes [10]. Gene silencing is achieved by imperfect or perfect pairing of microRNA with3'UTR of target mRNA thereby resulting in inhibition of translation or degradation, respectively [11,12]. They are transcribed as long transcripts and processed by the enzymes Drosha and Dicer into microRNA duplexes. The functional strand of the duplex is loaded into the RNA- induced silencing complex enzyme (RISC) for gene silencing [10]. MicroRNAs play a role in diverse biological processes including development, cell growth and apoptosis [11,13,14]. Their expression levels are usually disrupted during malignancy [12], making them crucial regulators of the cell cycle and potential biomarkers for cancer [15].

MicroRNA 21 (miR-21) is overexpressed in most solid malignant neoplasms such as carcinomas of the breast, pancreas, lung, stomach and prostate, and glioblastomas [11,16-18]. miR-21 is involved in tumourigenesis [17], tumour progression [19], and metastasis [20]. Data generated by Thapa and colleagues reveal that the levels of miR-21 are elevated in circulating B cells of HIV positive individuals who eventually develop NHL in less than three years [15]. An above normal increase in the expression of miR-21 has been observed in cell lines, tissue [21] and serum [22,23] of DLBCL patients. Given the high rate of HIV infection in Southern Africa and the increased risk of acquiring DLBCL in HIV positive individuals; there is little research on miR-21 expression in DLBCL of HIV positive individuals. Therefore, we conducted a study to compare the expression levels of miR-21 in DLBCL in HIV positive and HIV negative individuals, and also to determine its prognostic potential in HIV positive DLBCL since no data is currently available in this particular cohort.

APPENDIX Q

RN vs. DLBCL in HIV negative samples		
Mass (m/z)	**Average spectral counts**	**P Value**
848.6824904	1.640581802	3.25488E-10
	4.621165528	
966.8115383	1.781379787	1.77266E-10
	7.848620903	
849.6895863	2.536168069	1.26326E-08
	4.448123243	
713.5894992	1.566549939	4.81689E-09
	3.764913579	
1199.140513	42.04226872	7.46274E-10
	16.41734753	
944.8332528	12.94403384	2.00549E-10
	57.9598311	
945.8189248	8.555499662	2.79281E-08
	23.40624022	
1105.902484	3.567334432	0.001222306
	9.516735867	
704.5296597	4.433981487	3.30131E-10
	0.958931649	
703.5300794	9.704732021	2.00549E-10
	1.425739457	
816.6952902	1.903373504	3.99985E-06
	3.144055688	
850.7294258	1.932036404	2.79049E-05
	3.070037899	
927.8205705	5.212617163	3.25488E-10
	2.356424317	
801.7069476	2.048129317	1.30734E-06
	3.684781149	
1032.92104	18.63873375	2.73687E-08
	6.674150391	
1039.851656	6.610977244	2.53088E-05
	9.995413492	
928.7982625	3.462690006	1.72394E-07
	2.17224338	

805.7201214	4.202738401	0.0003027
	3.055520977	
789.6996394	6.973311073	2.79049E-05
	4.024570292	
800.6984832	3.02172035	4.85845E-05
	1.958786847	
1287.180979	13.05314673	0.000729309
	8.307209538	
1506.353642	15.781436	0.001770492
	10.48666049	
898.7809871	2.292654573	0.000795565
	2.99261064	
861.3485862	1.478005994	0.000348822
	4.983441684	
1792.36239	19.40193012	0.00119135
	14.01941369	
957.8393698	3.897666709	0.001921746
	4.659769973	
788.7046317	13.61064674	8.52279E-37
	11.88406465	
940.7983161	3.680547966	0.001921746
	3.968477108	
1350.072	8.229488832	0.002098595
	8.336123317	
971.8538135	3.432872197	0.002836113
	3.463089742	
RN vs. DLBCL in HIV positive samples		
Mass (m/z)	**Average spectral counts**	**P Value**
1199.264838	22.57148023	5.83654E-08
	47.3583536	
928.8523233	3.53048065	2.57665E-09
	7.470231581	
929.8606054	2.730678081	2.67795E-07
	4.948219418	
913.870908	2.308478534	6.95571E-08
	3.885498974	
1039.939009	7.808761889	2.01347E-05

	12.79867059	
955.9166261	3.340607119	2.12542E-05
	5.173299318	
1032.995113	12.47804688	2.48328E-05
	24.77230317	
940.873734	3.66371794	0.00033059
	5.522471635	
957.9176726	3.773143406	0.000827826
	5.877385793	
944.9139686	37.43925397	0.023544877
	53.18901007	
945.901411	19.32856228	0.096060132
	25.5430401	
976.8613064	6.174796667	0.641477291
	4.416680602	
901.8573617	2.82523104	0.020909117
	3.6762631	
806.7813637	2.744360136	0.020909117
	3.519799067	
1350.236044	12.32194462	0.005096781
	15.14775719	
849.767046	5.033143867	0.005526735
	4.204505114	
788.7801803	10.10359859	0.035297695
	13.5359925	
805.7879371	5.212136549	0.118286364
	6.312920751	
801.803695	3.972383905	2.87988E-16
	2.905240387	
789.7732526	4.881943575	0.118286364
	6.03710889	
713.6937518	3.575978626	6.85905E-14
	2.680069239	
816.7791454	3.233234767	0.018723773
	3.655313048	
850.809073	4.049494886	0.014078119
	3.607894566	
848.7712425	3.317296703	0.503438499

	3.471410596	
927.8931057	4.539180318	0.020442803
	4.651097341	
DLBCL - vs. DLBCL +		
Mass (m/z)	**Average spectral counts**	**P Value**
848.6744654	4.332166842	6.75016E-14
	2.391057541	
849.6830213	4.968031645	3.24363E-12
	3.09719882	
1199.158789	16.09044641	2.02709E-11
	33.16941282	
966.807324	8.277738486	3.24363E-12
	5.052465512	
1032.906715	6.940176357	2.22657E-11
	15.1788371	
816.6916499	3.498571481	3.59153E-06
	2.581095714	
944.8367432	61.54123199	0.340965291
	47.43398047	
805.7182447	3.560944407	0.128475832
	4.429705026	
850.7211837	3.549771263	0.018687064
	3.036758273	
957.8299739	4.389037793	0.048655014
	5.065131029	
789.7025559	4.475229145	0.048655014
	5.194636623	
1105.853231	10.6876101	0.009558436
	13.74151664	
945.806785	25.75740601	0.390092408
	21.81732467	
1039.845578	10.99789716	0.068868156
	10.18944963	
940.7993617	3.988628465	0.044143044
	3.496476818	
788.7127264	12.51059334	0.690012314
	11.83853688	

GC vs. non-GC in HIV negative samples		
Mass (m/z)	**Average spectral counts**	**P Value**
836.6443385	5.198087328	2.16956E-23
	1.577030151	
898.7052997	4.491995532	2.16956E-23
	2.089025364	
852.6430848	3.348912967	4.48957E-20
	1.370248327	
837.641179	3.136661144	3.18523E-16
	1.624835295	
850.6666988	4.061525805	8.89692E-15
	2.061777584	
1105.835641	15.75522758	3.42633E-11
	7.456601925	
816.6416667	3.964127571	9.91374E-10
	2.465323664	
1199.061796	7.309773584	4.42146E-08
	20.77178604	
1032.820831	4.116179079	0.000177643
	7.239172787	
1127.827176	7.068504182	0.000309293
	5.448507232	
802.6470153	2.515063014	0.000676255
	2.041782231	
966.7438184	5.731880095	0.001197052
	8.184522349	

957.763857	3.799186508	0.00128222
	4.951484216	
945.7453563	14.83723585	0.002949321
	24.59288367	
944.7675888	31.23194311	0.003071558
	62.35475003	
971.7734707	3.052610511	0.025848861
	3.751043953	
1792.211581	14.80173834	0.029116639
	11.68849907	
940.7149721	2.67508276	0.104815427
	3.479708905	
861.2860225	4.569758427	0.110176874
	3.019012809	
788.6521874	7.404979294	0.179332437
	10.18406471	
849.6225426	4.508143137	0.265722764
	4.095063966	
713.5660022	3.479005901	0.394100478
	3.892589762	
757.6219046	2.955278254	0.419146526
	3.220660652	
801.65776	3.296574302	0.449588336
	3.495802456	

1039.773724	9.624692087	0.662995253
	10.08619471	
789.6415841	3.525210168	0.781520977
	3.35819505	
848.6139741	4.926367149	0.843283757
	5.046528244	
845.6503385	3.272838889	0.872566889
	3.375483166	
805.6547989	2.812182556	0.872566889
	2.752888386	
GC vs. non-GC in HIV positive samples		
Mass (m/z)	**Average spectral counts**	**P Value**
1105.431116	24.61210873	1.22509E-31
	5.313585109	4.35763E-30
1199.245457	17.67092282	1.27058E-24
	50.91851233	1.27058E-24
1032.988102	7.540316488	1.27058E-24
	23.56449408	1.27058E-24
944.884397	37.7799762	1.25578E-12
	57.58455126	1.25578E-12
788.7635164	9.003112793	1.35664E-12
	14.50136772	1.35664E-12

CONFERENCE PROCEEDINGS AND PUBLICATIONS

PUBLICATIONS

Phillips P, McGrath E, Sekar D, et al. MicroRNA 21 expression levels in HIV and non-HIV related diffuse large b cell lymphoma. Hereditary Genetics 4:2, 2015 http://dx.doi.org/10.4172/2161-1041.1000143

CONFERENCE PROCEEDINGS

1. Pumza Phillips, Raveendra Sookhayi, Zac McDonald, Richard Naidoo. (2013, April). A proteomics assay using formalin fixed paraffin embedded tissue from HIV and non- HIV related diffuse large B cell lymphoma. Laboratory Medicine Congress, Cape Town.
2. Pumza Phillips, Richard Naidoo, Raveendra Sookhayi, Durairaj Sekar, Luiz Zerbini. (2013, September). The levels of microRNA 21 in diffuse Large B cell subtypes. UCT/CLS research day, Cape Town.
3. Pumza Phillips, Raveendra Sookhayi, Richard Naidoo. (2013, Septermber). Biomarker identification in diffuse large B cell lymphoma. NHLS scientific day, Cape Town.
4. Pumza Phillips, Durairaj Sekar, Raveendra Sookhayi, Dhiren Govender, Zainab

Mohamed, Luiz Zerbini, Richard Naidoo. (2014, October). MicroRNA 21 expression levels in HIV and non HIV related diffuse large B cell lymphoma. International Academy of Pathology, Thailand.

5. Pumza Phillips, Raveendra Sookhayi, Richard Naidoo. (2014, September). Biomarker identification in diffuse large B cell lymphoma. NHLS scientific day, Cape Town.
6. Pumza Phillips, Raveendra Sookhayi, Richard Naidoo. (2015, September). Biomarker identification in diffuse large B cell lymphoma. NHLS scientific day, Cape Town.