Characterization of U2932 cell line subpopulations and evaluation of their sensibility to a chemotherapeutic drug.

Master of Science Thesis

Alana Miranda Pinheiro

Medicine with Industrial Specialisation- Biomedicine

Aalborg University

Supervisors:
Professor Karen Dybkær, MSc, PhD
Department of Haematology, Aalborg University Hospital, Denmark
Senior Researcher Julie Støve Bødker, MSc, PhD
Department of Haematology, Aalborg University Hospital, Denmark

31-05-2018

Acknowledgment

This master thesis was developed during the 9th and 10th semesters, initiating on September 2017 and finalized on May 2018. All the procedures were developed at Aalborg University Hospital, Aalborg, Denmark, under the supervision of Professor Karen Dybkær and co-supervision of Julie Støve Bødker. The student academic expenses were entirely covered by INNO+ Scholarship in Medicine with Industrial Specialisation. The scholarship is associated to National Experimental Therapeutic partnership (NEXT), granted by the Danish Innovation Fund for Research, Technology and Growth. The School of Medicine and Health at Aalborg University and Aalborg University Hospital are partners of the NEXT partnership focusing on early clinical trials of novel drugs.

Firstly, I would like to express my gratitude to my supervisor, Karen Dybkær, who provided me access to the laboratory facilities where this master thesis was developed and assisted me throughout the project in a very helpful and friendly manner. I would like also to thank my co-supervisor Julie Støve Bødker who patiently instructed me about the transcriptional profile analysis and gave me extra attention during this thesis writing process. Both supervisor and co-supervisor constantly motivated me and gave me valuable advices.

Secondly, I would like to thank the laboratory technicians Zuzana Valnickova Hansen and Helle Høholt for instruct me about the cell lab procedures and give me technical assistance concerning to drug screen, DNA and RNA purification for microarray analyses. I am also thankful for the Post Doc Alexander Schmitz who introduced me to the flow cytometry, guiding me throughout the experiment set up and data analyses. In addition, I am especially grateful for Ditte Jespersen and Hanne Due for constructive advices and help. Finally, I would like to thank the NEXT partnership and the Danish Innovation Fund for Research, Technology and Growth for provide me financial support for my master studies.

Abbreviations

ABC	Activated B cell Lymphoma
AID	Activation-induced cytidine deaminase
ANOVA	Analysis of Variance
BAGS	B-cell associated gene signatures
BCL2	B-cell lymphoma 2
BCL2A1	BCL2-related protein A1
BCL6	B-cell lymphoma 6
BCR	B cell receptor
Blimp-1	Beta interferon regulatory domain
BrdU	Bromodeoxyuridine
cDNA	Complementary DNA
CSR	Class switch recombination
CSCM	Cancer stem cell model
CT/ BEAC	BCNU, etoposide, araC, cyclophosphamide regimen.
DHAP	dexametasone, cisplatin, araC regimen
DLBCL	Diffuse large B cell lymphoma
DSMZ	German collection of microorganism and cell cultures
EBV	Epstein-Barr virus infection
EPOH	Etoposide, prednisone, vincristine, cyclophosphamide, doxorubicin regimen
FACS	Fluorescence activated cell sorting
FDC	Follicular dendritic cells
FDR	False discovery rate
FLIP	FLICE-like Inhibitory Protein
FOXP1	Forkhead box protein P1
GADD45A	Growth Arrest and DNA Damage Inducible Alpha
GC	Germinal center
GCB	Germinal center B- cell Lymphoma
GEP	Gene expression profiles
GI	Growth inhibition
Hemaclass	Online One-By-One Microarray Normalization and Classification of Hematological Cancers for Precision Medicine
HIF-1	Hypoxia-Inducible Factor
HL	Hodgkin's Lymphoma
IFN	Interferon
IPI	International Prognostic Index
IRF4	Interferon Regulatory Factor 4
MHCII	Major histocompatibility complex II
MAPK	Mitogen Activated Protein Kinases
MDR1	Multidrug resistance 1
MEF2B	Myocyte Enhancer Factor 2B
MIME	Methyl-GAG, ifosfamide, methotrexate, etoposide regimen
MM	Multiple myeloma
MYC	MYC Proto-Oncogene
MYD88	Myeloid differentiation primary response 88

NC	Non- classified
NF-kB	Natural factor kappa B
NHL	Non- Hodgkin Lymphoma
OS	Overall survival
PAX 5	Paired box protein 5
PC	Plasma cells
PFS	Progression-free survival
PRDM1	PR Domain Containing 1, With ZNF Domain
RA	Rheumatoid arthritis
RANKL	Receptor activator of nuclear factor kappa-B ligand
R-CHOP	Rituximab, cyclophosphamide, doxorubicin, vincristine and prednisone combined regimen
REGS	Resistance gene signature
RT Q- PCR	Real- time quantitative PCR
SHM	Somatic hypermutation
SLE	Systematic lupus erythematosus
STAT3	Signal transducers and activators of transcription
TNF	Tumour necrosis factor
TP53	Tumor protein p53
UC	Unclassified
XBP1	X-Box Binding Protein 1

Summary

1. Abstract

DLBCL presents a variety of mutations throughout the B cell differentiation process. Genetic alteration at specific B cells state classify this tumor as ABC or GCB subtype. DLBCL subtypes demonstrate straight relation with patients' prognostic and overall survival where ABC is considered the most aggressive and with elevated resistance level to R- CHOP regimen. Recently, two subpopulations demonstrating different and steady immunophenotype for 35 days combined with altered expression of genes regulating the B cell differentiation were identified on U2932 cell line, an ABC- like DLBCL. Elevated heterogeneity and plasticity is a marked characteristic of malignant cell in a carcinogenic environment favoring the disease progression, drug resistance and culminating into metastasis. Thus, here we hypothesis that U2932 is a highly heterogenous cell line composed of different subpopulations at distinct B cell states of differentiation marked by a different phenotype and drug resistance. Aiming to confirm or deny this assumption, we characterized U2932 cell line and its subpopulations immunophenotype, transcriptional profiling and sensibility to doxorubicin. Flow cytometry analyses demonstrated the coexistence of different subpopulations on U2932 cell line, each of them marked by a specific phenotype and genetic alterations. Due to their consistent expression during the first month of analyses, P4, expressing CD38/ CD20- high, and P5, expressing CD38/ CD20- low, are the main subpopulations analyzed in this study. To precisely determine the state of differentiation and molecularly comprehend the proliferative and drug sensibility features of each subpopulation, gene expression and pathways enrichment were determined by microarrays analyses at day 36. For P4, the phenotype resembled extrafollicular plasma cells demonstrating *PRDM1*/*Blimp-1*-low, *XBP1*s-low, CD38/CD20 -positive, *PAX5*-high, *BCL6*-high and *IRF4*-positive expression. On the other hand, P5 phenotype resembled germinal center plasmablast/plasma cell, demonstrating *PRDM1*/*Blimp-1*- positive, *XBP1*s- positive, CD38/CD20 -negative, *PAX5*-high, *BCL6*- intermediate, *IRF4*- positive and EBV expression. This data is supported by hemaclass BAGS classification of P4 and P5 as plasmablast when only normalized with 17DLBCL + 6U2932, and the ABC classification for all cohorts utilized for normalization. In addition, P4 showed lower proliferation and higher resistance to doxorubicin compared to the other subpopulations and original cell line. This data is associated with the up regulation of HIF- pathway for P5 resulting from P4 vs P5 pathway enrichment, and the highly positive expression of CD20 for P4. Accordingly, the hemaclass REGS classification of U2932, P4 and P5 as doxorubicin resistant are common for the distinct normalization cohorts. Autoimmune disturbances and/or infectious conditions are common

pathways demonstrated between U2932 vs P4 and P4 vs P5 enrichment. Finally, all U2932 subpopulations demonstrate immature B cell like phenotype marked by CD19+, CD38+/-, CD10+, IgM+, CD5- and CD27- at day 107, indicating a process of dedifferentiation. At the same date, P3 an P4 demonstrate an heterogenous expression for CD43 suggesting that the malignant cells are still in a transiting state. After day 100, all subpopulations demonstrated similar resistance to doxorubicin in accordance with the common phenotype. In all cases, CD19 expression was overall positive even though a minor fraction of all subpopulations demonstrated low expression of this marker, indicating a process of internalization or the progression of the cells in the B cell differentiation.

2. Introduction

2.1. B cell differentiation

B cell is a fundamental component of the adaptive humoral response providing a higher specified reaction against antigens (Ag) through the productions of respective antibodies (Ab), an essential function of the immune system. The initial phase of B cell development is centralized in the formation of the B cell receptor (BCR) constituted by two identical heavy chain and two identical light chain, presenting a carboxyterminal constant (C) and an amino terminal variable (V) fragment. The BCR formation is orchestred by the variable (V) diversity (D) joining (J) recombination of genes segments encoding the variable region of heavy chain, and V(J) recombination for the light chain, which might consist of a kappa or lambda isotype. The chains are composed of immunoglobulin polypeptides connected by disulphide bridges. In addition, the distinct possibilities of heavy and light chain pairing determine the antigen that BCR interact with [1-4]. Immature B cell BCR firstly only express IgM. Still in the bone marrow, the cells are submitted to the central B cell selection where immature B cell reacting to antigens expressed by itself are selected and conducted to apoptosis or receptor editing [4]. Once presenting functional antigen receptor, foreign antibody reacting and with V fragments are able to conduct a protein translation, the B lymphocyte precursor leave the bone marrow as a mature, naive B cell. After receiving an antigen signal it migrates to the germinal center (GC) initiating the B cell differentiation process [1-4] (Figure 2.1).

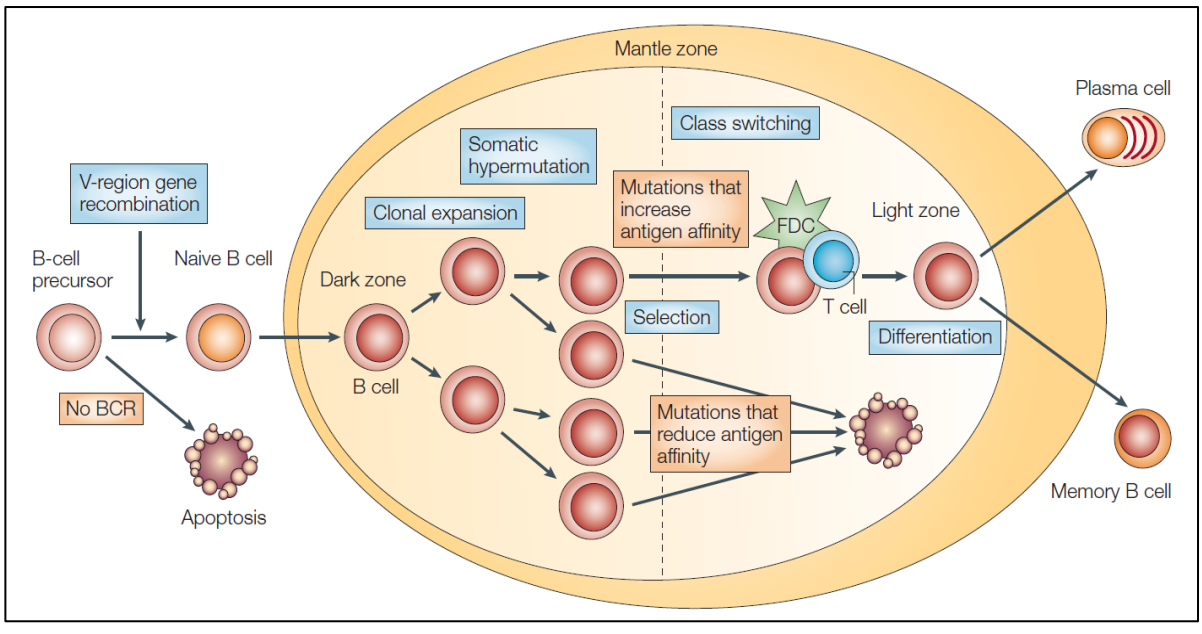

Figure 2.1*: B cell early development, germinal center differentiation and terminal differentiation.* Naive antigen-activated B cells receive signals from 'T-cell help' and migrate to primary B-cell follicles in secondary lymphoid where the germinal centres are established. At this point, proliferating GC B cells displaced outside of the follicle form a mantle zone around the GC. Proliferating GC B cells are predominant in the GC dark zone, where the process of SHM is activated. This process introduces a high rate of mutations into the rearranged Ig variable (V)-region genes of the B cells. At the GC light zone, the mutations from the SHM are evaluated as advantageous or disadvantageous. B cells with disadvantageous mutations undergo apoptosis and are submitted to class switch recombination (CSR). For the GC B cells which mutations favour the antigen affinity are positively selected. At the selection process, GC B cells interact with CD4+ T cells and follicular dendritic cells (FDCs). Finally, the GC B cell leave the microenvironment after it has differentiated into memory B cells or plasma cells (Source: Kuppers *et al*. 2005) [1].

The B cells differentiation is an irreversible mechanism due to the occurrence of DNA structure alteration mediated by immunoglobulin gene remodelling processes, somatic hypermutation (SHM) and class switch recombination (CSR). SHM and CSR guarantee B cell capacity to express high-affinity antibodies for a variety of isotype classes finally resulting on plasma and memory cells. The GC is a histological formation resulting from follicular aggregation in secondary lymphomas structures, mainly characterized by dark and light zone regions. The dark zone majorly presents proliferating or clonal exposing GC B cell, submitted to mutations, short deletions, or insertions at the Ig V region genes, a process known as SHM. At the light zone, cells presenting disadvantageous mutations reducing antigen affinity to BCR are negatively selected and undergo apoptosis [1, 2, 5]. The cells positively selected, centrocytes, establish close contact with CD4+ T cells and follicular dendritic cells (FDCs). At this site, B cells test their antigen receptor (BCR) against antigen displayed on FDCs, competing for limited T follicular helper cells, submitted to CSR and posteriorly originating centroblasts [6]. CSR consists on sequential switch of regions from IgM heavy chain originating IgG, IgA, or IgE. Finally, GC B cell differentiate into plasma or memory cell (Figure1). The entire process conducted on the germinal center are regulated by a variety of transcriptional networks, such as PAX5 and BCL6 [1, 2, 5, 6].

Besides plasma cells originated from the germinal center, there are also the extrafollicular plasma cells which formation correspond to a rapid T- cell dependent response, marked by the higher strength of interaction between these cells BCR and antigen compared to the GC plasma cells. This process culminates into rapid Ab production posterior to Ag interaction [paus et al 2008]. It is still not completed elucidated the non- GC B cell differentiation. However, it is known that B cell interaction with antigen in the blood or marginal zone direct the cells to the splenic T zones, where they are induced to growth not compulsorily depending on CD4+ T cell help. After approximately two cell cycles, the cells migrate as plasmablast from the T zone to

the medullary cord of lymphnodes or to the *foci* in the red pulp of spleen where they become plasma cells. A considered fraction of extrafollicular plasma cells are short lived cells, highly expression *BCL6* which represses B lymphocyte- induced maturation protein (Blimp-1), a transcription regulator fundamental for the terminal GC differentiation into plasma cells. The complete differentiation from plasmablast to plasma cells is dependent to CD11high dendritic cells [7]. Due to the inexistence of SHM at the extrafollicular plasma cell formation, these plasma cells only present presenting Ab specificities encoded within the primary repertoire [8].

2.2. Fundamental transcriptional factors for B cell differentiation

Germinal center differentiation is a process highly regulated by a diversity of transcription factors aiming to ensure the appropriate formation of plasma and memory cells. Those transcription factors functions are highly interrelated. Consequently, the inappropriate expression of one transcription factor compromise the homeostasis of the entire process. Here we focus on five of the fundamental transcription factors for the B cell differentiation: *BCL6*, *PAX5*, *IRF4*, *MYC* and *PRDM1*/Blimp-1.

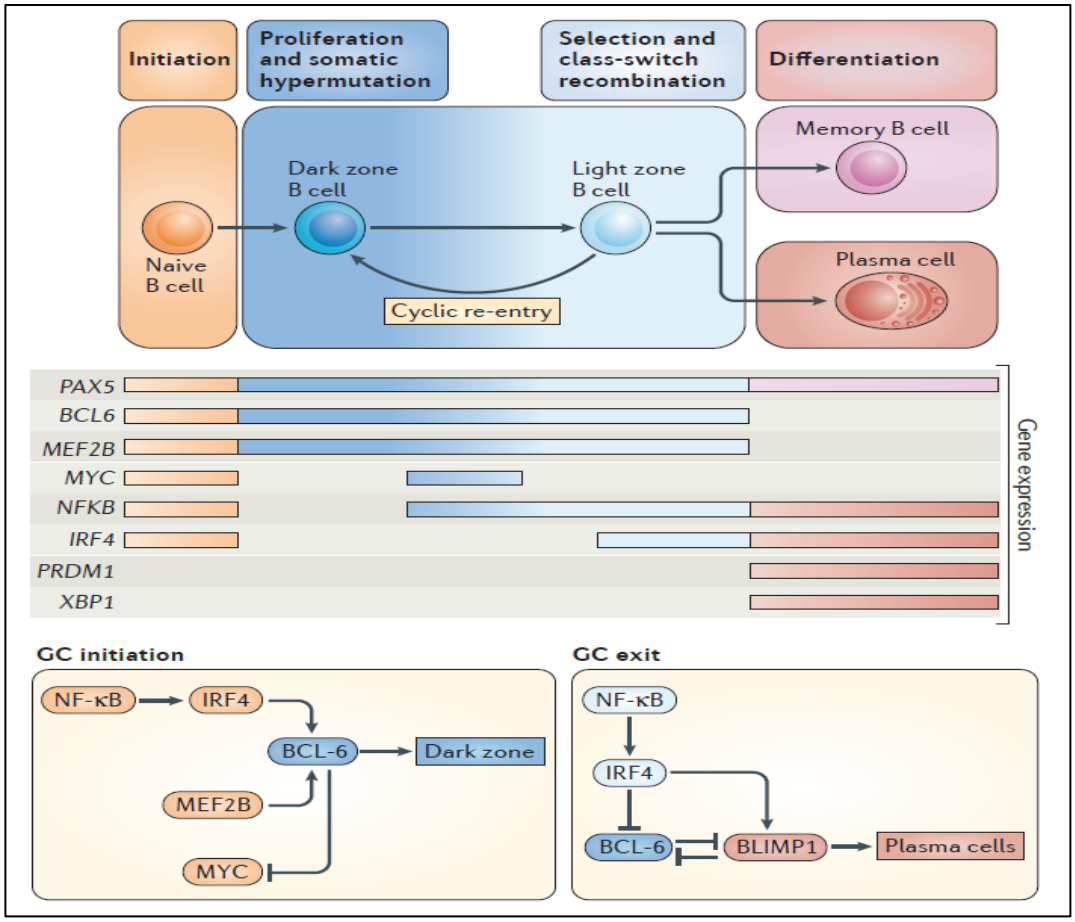

***Figure 2.2**: Transcriptional networks regulating the GC differentiation process.* A variety of transcription factors participate of the regulation of the initial, development and terminal GC differentiation activities. *PAX5* is expressed continually throughout mature B cells existence and not expressed on cells driven to plasma cell differentiation. On the other hand, *BCL6* and *MEF2B* are only expressed during GC stage. Although *MYC* is not expressed on the majority of GC B cells, it is requested for B cell initiation and reconducted into the dark zone from the light zone. For the GC initiation and GC light zone, *NF- kB* and *IRF4* expression are required, but not during GC dark zone phase. For the cells driving into terminal plasma cell differentiation, *PRDM1*, which encodes Blimp-1, and *XBP1* expression is requested. Besides regulator of GC differentiation, these transcription factors modulate each other activities. *NF- kB* activates *IRF4*, promoting *BCL6* expression during GC initiation, and inhibiting *BCL6* and *PRDM1* expression at terminal GC differentiation. *BCL6* downregulates *MYC* and *PRDM1* expression (Basso et al. 2015) [9].

2.2.1. BCL6

BCL6 is considered the major transcription factor regulating the early stage of germinal center differentiation process. *BCL6* is a transcription repressor avoiding immature B cells proliferation, activation and differentiation which would compromise the plasmablast and memory cells accuracy [1, 2, 5]. *BCL6* is highly expressed in part of the naive B cell which BCR properly engaged by antigens. *BCL6* regulates the SHM and CSR processes by establishing a favourable environment for DNA mutations related to immunoglobulin gene remodelling processes [5, 9]. To accomplish its main functions, *BCL6* affects multiple signalling pathways involved with GC differentiation. To promote a higher threshold for DNA damage responses *BCL6* interferes on transcriptional networks of interest, for example by directly regulating TP53. To avoid immature B cells activation, *BCL6* regulates activation-induced cytidine deaminase (AID), promoter of somatic mutation and translocation occurring at SHM and CSR. In addition, it impairs terminal differentiation progress by downregulating Blimp- 1, a plasma cell master regulator. The high expression of *BCL6* is maintained during the germinal center differentiation process until light zone ending phase where *IRF4* mediates it repression at RNA and protein level through CD40 signaling. At post transcriptional level, *BCL6* expression is also repressed through MAPK which mediates the protein degradation as a response to BCR signaling and acetylation mediated inactivation [2, 5, 9]. Due to the *BCL6* repressor feature, its downregulation in the final light zone phase is crucial for post GC differentiation properly occur.

2.2.2. PAX5

PAX5 is considered the main regulator of B cell identity, being expressed during the entire GC differentiation process. It targets B cell on myriad of DNA sites and presents a dual function varying from repressor to stimulator according to the stage of differentiation and the gene it interacts with. For immature B cells, it participates to the lymphoid commitment, selecting and designating the progenitor cells to the differentiation process. For mature B cells, it stimulated the expression of genes essential for the development of fundamental B cell characteristics, such as genes composing Ig heavy chain and CD79A (BCR), interferon regulatory factor *(IRF)4*, fundamental for the regulation of initial phases of GC differentiation and initiation of terminal plasma cell differentiation, respectively. As a repressor, *PAX5* also inhibits the expression of lineage inappropriate genes at PCs. Moreover, *PAX5* directly represses the expression XBP1, a master regulator of later stages of GC differentiation. Due to the function of *PAX5* as a stimulator and regulator of the B cell transcriptional programme, its down-regulation is compulsory for GC B cell differentiation into PCs properly expressing Ab [5, 10, 11].

2.2.3. MYC

MYC identify and process cellular signals producing a transcriptional feedback and consequently a cellular metabolism stimulation [12]. *MYC* expression during GC differentiation is marked by induction or amplification of transcriptional factors culminating into GC B cell growth and proliferation. For normal B cells, *MYC* expression vary according to the stage of differentiation process. At GC initiation, its expression rates increase, followed by a transcriptional suppression during SHM at dark zones, and a crescent re- expression at light zone stage. *MYC* suppression at dark zones result from its interaction with *BCL6*, while *MYC* re-expression in part of early stage of light zone B cells is associated with *BCL6* suppression at this phase of differentiation process. *MYC* expression at light zone occur on B cells redirected to dark zone for further antigen affinity development. The light zone B cell demonstrating appropriate antigen affinity downregulate *BCL6* and *MYC* and exit the germinal center to become early plasmablast or memory cells [9, 13].

2.2.4. IRF4

IRF4 is expressed on initial GC B cell, light zone B cells and upregulated at the final PC differentiation. This expression variation of *IRF4* is determined by the strength of BCR signaling, varying based on the affinity between BCR and antigen, low at the initial phase of GC differentiation and high at PC differentiation. The variable motifs of *IRF4* interaction also allow its participation on divergent functions. At initial stage of B cell differentiation, *IRF4* contributes to the *BCL6* expression participating of GC differentiation process. At final stage of B cell differentiation *IRF4* expression is fundamental for Ig class switch recombination, Ig production and secretion by PC. Consequently, *IRF4* deficient animals have already demonstrated AID deficiency and impaired formation of mature PC caused by low XBP1 expression. Hence, *IRF4* contributes to Blimp- 1 expression as well as represses *BCL6* extinguishing GC differentiation process [5, 11].

2.2.5. PRDM1/Blimp-1

B lymphocyte- induced maturation protein (Blimp- 1) is expressed by the *PRDM1* gene, corresponding to a fundamental regulator of GC differentiation and a hallmark for the terminal plasma cell differentiation. It expression repress *MYC, BCL6* and *PAX5* gene. By inhibiting the *MYC* and BCL6 expression it concludes the GC differentiation avoiding typical B cell process occurring at this state, such as antigen affinity improvement. In addition, Blimp-1 inhibition of *PAX5* culminates into the loss of B cell identity and initial plasma cell characterization through the expression of XBP1 [14, 15]. Blimp-1 mutation or repression avoid the cell cycle end and its abnormal expression cause phenotype modification on antibody secreting cells [14]. The main function of Blimp-1 is ensure the plasma cell physiological conditions for Ab expression [15].

2.3. Lymphoma

Lymphomas comprise a set of malignant lymphocytes of more commonly B cell, but also T cells and natural killer cells. It results from acquired capacity of cellular growth misregulation and replication substantiated by a reprogramed energy metabolism. The lymphomagenesis process is marked by constant proliferation signaling, inhibition of cellular growth repressors and stimulation of anti-apoptotic factors. Consequently, the cells become immortally replicable

favouring angiogenesis, invasion and culminating into a metastasis. These genomicaly unstable tumours paradoxically develop tumor promoting inflammation function which simultaneously produces immune antitumoral response and contributes for the tumor environment formation. [16-18].

Due to the prominent many- sided lymphoma characteristic, this tumour diagnostic and classification require a complex combination of morphological, histological, clinical, and molecular features evaluation [16, 18]. The main lymphoma classes are the Hodgkin's Lymphoma (HL) and non- Hodgkin Lymphoma (NHL), corresponding to approximately 10% and 90% of lymphoma cases, respectively. As the major lymphoma class, NHL has a diversified histological and clinical feature, thereafter, being classified into subtypes, diffuse large B-cell lymphoma (DLBCL), follicular lymphoma, mantle cell lymphoma, splenic marginal zone B-cell lymphoma, and chronic lymphocytic leukemia [19, 20].

The DLBCL is the most recurrent subtype of NHL, corresponding to about a third of this lymphoma class cases. Although, it corresponds to NHL sub classification, DLBCL is still marked by a molecular heterogeneity, variable gene expression and phenotype, culminating into a variable level of aggressiveness and clinical outcome. Alteration on specific stages of GC differentiation is the main cause of DLBCL development. The GC processes promoting Ig high affinity and isotype diversity are also the moments of high DNA alteration where transcriptional and protein misregulation may originate malignant B cells [19]. DLBCL cells are marked by erroneous regulation of GC differentiation activities. *PAX5, BCL6* and *PRDM1* expression at transcriptional level determine the occurrence of naïve B-cells, germinal center B-cells, and plasmablast, respectively. Hence, these factors act as essential indicator of malignant cell differentiation stage and consequently its most appropriate classification.

2.3.1. DLBCL

Diffuse large B cell lymphoma highly compromises patient's life, corresponding to 31% of NHL cases in Western Countries and 37% of B-cell tumours worldwide [21]. It is registered 25,000 cases of DLBCL annually over the world, and data indicates that 3-4 of 100.000 European Union habitants are diagnosticated each year. Only in Denmark 400 DLBCL diagnostics are registered per year [22- 24]. This DLBCL is more commonly recurrent among elderly patients 60-80 years old. Higher annual incidence of 26.6 of 100.000 cases is registered

for patients age 80- 84 in Europe, lower disease prognostic is also recurrent among higher age patients [21, 24, 25]. This group of lymphoid malignancies is featured by large cells presenting vesicular nucleoli and basophilic cytoplasm, elevated proliferation rate is also a mark of these cells growth pattern [21]. The DLBCL are located on lymph nodes, Waldeyer's tonsillar ring or extra nodal tissues on approximately 60%, 10% and 30% of cases, respectively. Waldeyer's tonsillar ring represents the origin of 11% of DLBCL with patient's overall survival of 5 years in 77% of cases [22, 26, 27]. DLBCL is usually the result of histological transformation of a primary or secondary lymphoma culminating into *de novo* tumour formation. In addition, it distinct tissue origin is favourable for the heterogeneity of this lymphoid malignancy, pathology and clinical outcome [21, 26, 28]. It also includes cases deriving from clinical evolution of various less aggressive B-NHL, for example follicular lymphoma and chronic lymphocytic leukaemia [29, 30].

2.3.2. Molecular subtypes of DLBCL

Based on gene expression analyses, Germinal center B- cell Lymphoma (GCB) and Activated B- cell Lymphoma (ABC) are the main DLBCL molecular subtypes. Each subtype is marked with specific or predominant genetic alteration indicating distinct stage of differentiation for the cell of origin, directly influencing on the pathway of this malignancy development [5, 30]. However, both biological entities present common pathogenic characteristics, such as alteration on chromatin modifiers and *BCL6* dysregulation. The frequently registered genetic alteration on both DLBCL subtypes comprises on rearrangement of immunoglobulin genes and Ig isotype diversity. In this sense, DLBCL originates from GC B-cell, centroblasts, centrocytes, or B- cell at later stage of differentiation, plasmacyte, plasmablast [5, 9, 31]. GCB originates from ordinary GC B- cell while ABC arise from B- cell interrupted during plasmocytic differentiation. Finally, the genetic distinction among GCB and ABC results in different chemotherapy response and outcome, in other words, distinct biological entities [5, 31].

GCB DLBCL. The main event determining the GBC DLBCL identity is the misregulation of SHM and CRS and confinement of B cells in the GC. In parallel, GCB development is supported by hyperactivation of antiapoptotic system [2, 9]. In addition, mutation on the negative auto regulatory *BCL6* first noncoding exon occurrent on this DLBCL subtype causes *BCL6* overexpression. Moreover, somatic mutations on B-cell specific enzyme AID, a SHM and CSR mediator, results on mutations on *MYC* genes [2, 9, 31, 33]. In this sense, *BCL6* and

c- MYC are among the main genes overexpressed on GCB. GCB analyses has recently indicated the occurrence of aberrant activity of transcriptional repressor for *PRDM1* and *IRF4*, facilitating GC proliferation and impairing post-GC differentiation [9, 31].

ABC DLBCL. The main misregulated event determining the ABC DLBCL identity is the blockage of terminal GC B cell differentiation into plasma cells, a process substantiated by constitutive activation of NF-kB. The blockage of B cell terminal differentiation is marked by the *PRDM1* inactivation and the downregulation of the plasma cell master regulator, Blimp-1, as a principal effect. This process is mediated by a combination of mutation and altered pathways more often observed at this DLBCL subtype, inactivation of *PRDM1, BCL6* dysregulation by chromosomal translocations and alteration on transcription factor that can form a complex with *IRF4.* For the occurrence of NF-kB hyperactivation, mutations on CD79A/ CD79B avoid BCR endocytosis and inhibit this receptor negative regulator causing a chronic BCR signaling. Consequently, it promotes hyperactivation of CAD11 and its transduction complex, activating NF-kB. In addition, mutations on the gene encoding myeloid differentiation primary response protein (MYD88) modify this protein function culminating into NF-kB activation. Finally, negative regulators of NF-kB are inhibited. In this sense, ABC-DLBCLs is favoured by the NF-kB transcriptional programme pro-survival characteristic, maintaining the B- cells in erroneous differentiation stage [2, 9]. The overstimulation of NF-B pathway promotes the differently expression of its target genes *IRF4* on activated B cell lymphoma cell lines. The gene expression profile of ABC DLBCL is characterized with a constitutive expression of *IRF4, FLIP* and *BCL2* [30, 31, 34]. *IRF4* high expression allows the irregular proliferation of the tumor cells. *FLIP* and *BCL2* mRNA overexpression promotes antiapoptotic function [31]. Besides, the FOXP1 mRNA overexpression resulting from trisomy 3 aberration is also a hallmark of this DLBCL subtype [30, 33]. These genes compose the genetic ABC DLBCL "signature", and their constitutive expression is not observed in GCB DLBCL [34].

2.3.3. BAGS and REGS classification

The DLBCL subtypes classification is based on the characteristic of the lymphoma cell of origin, considering that relevant biological information of ABC/GCB cells is obtained from its non-transformed cellular progenitor. In this sense, this classification is substantiated on features of normal GC cells, centrocytes and centroblasts, or *in vitro* analyses of activated B-cells from

peripheral blood naïve and memory cells. This ABC/GCB classification is limited due to restricted knowledge about the extent of molecular and physiological similarities between normal B cell and malignant B cells [31]. For this reason, a more concise classification combining fluorescence activated cell sorting (FACS), gene expression profiles (GEP), and subset-specific B-cell associated gene signatures (BAGS) of naïve, centroblasts, centrocytes, memory, and plasmablast B-cells from normal tonsil tissue obtained from statistical modelling were develop by our group [35].

The BAGS determination is based on data corresponding to DLBCL from five different clinical cohort's. GCB DLBL were mainly classified as centroblasts and centrocyte in 30% and 50% of cases, respectively. On the other hand, ABC DLBCL were not classified as any subset-specific of BEGS, presenting 22% of samples as UC. The BAGS classification, independently of cell of origin classification and IPI, presented relevant prognostic relation with OS and PFS. This data indicates the BAGS classification introduces pathogenic and prognostic information not previously taken in consideration. However, BAGS analyses demonstrated enhanced correlation with prognose on GCB DLBCL cases, 23% of patients assigned as centrocytes presented prognostic improvements while 14% of GCB patients assigned as centroblasts demonstrated adverse prognose [35].

For resistance gene signature (REGS) classification, 26 malignant B cell lines were submitted to drug screen resulting in *in vitro* drug response. Sequentially, GEP before drug screen was related to the degree of growth inhibition dose for each cell line. Finally, drug resistance probability of chemotherapeutic composing the first line treatment of DLBCL, cyclophosphamide, doxorubicin, and vincristine, were determined for individual samples [36]. REGS indicated higher resistance to vincristine and doxorubicin by GCB-centroblasts than GBC-centrocyte samples, supporting survival analysis. BAGS classification promoted drug specific predictive information for GCB samples. This data indicates that GEP enable cell of origin classification for DLBCL, relating malignant B cell to normal B cell function and differentiation stages. In addition, samples classified in distinct BAGS subtypes present different clinical outcome according to pathogenesis and resistance specific mechanisms [35]. This classification allows improvement on disease management and individualized therapy.

2.4. Treatment and Prognosis

2.4.1. DLBCL treatment, the R-CHOP regimen

Currently, the first line treatment for DLBCL is the R-CHOP regimen, composed of five chemotherapeutic drugs, rituximab, cyclophosphamide, doxorubicin, vincristine and prednisone. The R- CHOP drugs interact in a way to promote the tumour cells degradation or malfunction. Rituximab is a chimeric monoclonal antibody which interacts with the CD20 antigen, highly expressed in the DLBCL. CD20 participates in the regulation of intracellular calcium, cell cycle, and apoptosis. Assuming the high expression of CD20 DLBCL phenotype, it inhibition induces the cells to direct apoptosis though lysis. Rituximab is considered the most efficient drug of the R- CHOP regimen [37, 38]. Cyclophosphamide adds alkyl radicals into DNA strands avoiding DNA replication through the DNA cross-linkage formation [39]. Doxorubicin presents two main pathways in the cancer treatment: intercalation into DNA and disruption of topoisomerase-II-mediated DNA repair and the radical free production [40]. Vincristine promotes the microtubules degradation and consequently compromises the occurrence of mitosis, interrupting the cellular multiplication [41, 42]. Finally, prednisone is utilized to reduce inflammation and suppress the body's immune response. It is commonly ingested previously to the other drugs in the R-CHOP treatment to prevent drug hypersensitivity [43].

Intending to improve the DLBCL treatment and outcome, rituximab was added to the CHOP regimen on 1990 [44]. From this point, R-CHOP has registered significant increase of 5 years OS in approximately 70% of patients compared to 15% registered to CHOP [45]. Despite the stimulating data, patients submitted to this regimen are still relapsing in 50%-60% of cases which associated with refractory diseases results in the cause of patient’s death in 40% of cases [21, 31, 46]. 85% of GCB DLBCL have achieved an OS of 3 years after R-CHOP treatment, whereas, 69% ABC DLBCL registered the same indicator data. It points to the fact that R-CHOP efficiency is affected by the specific molecular characteristics of DLBCL subtypes, being more effective among GCB DLBCL patients [36, 47]. Although DLBCL is characterized by extensive inter- and intra-tumour heterogeneity differing in drug response and treatment outcome, patients presenting distinct DLBCL subtype diagnostic are routinely treated with the same first-line chemotherapy [31, 48]. Considering that tumours subpopulation evolves under selective forces, it is natural to assume that tumour heterogeneity can facilitate therapy-induced selection of a drug resistant tumour population present in the primary tumour, which will later

result into treatment failure [48, 49]. That fact alarms the necessity of alternative treatments more accurately related to the lymphomas molecular features.

2.4.2. Doxorubicin

Doxorubicin (trade name adriamycin or rubex) is the main drug of the anthracyclines, a group of antibiotic and anticancer drugs [40, 50]. Its natural source is *Streptomyces peucetius var. caesius* from which it was first extract in the 1970's. Since then, doxorubicin was utilized as a treatment for a variety of cellular malignancies in different organs such as breast, lung, gastric, ovarian, thyroid, non-Hodgkin's and Hodgkin's lymphoma, multiple myeloma, sarcoma, and paediatric cancers [51, 52]. Doxorubicin is administered via intravenous infusion and presents cardiotoxicity as a marked side effect. This chemotherapeutic drug promotes cytotoxic effects through two main pathways, the disruption of topoisomerase-II-mediated DNA repair and oxidative stress [40, 50].

The main mechanism of action of doxorubicin is mediated by topoisomerase II, a DNA gyrase presenting high activity on proliferative cells. Doxorubicin penetrates the nucleus and target the DNA during the replication of the DNA helix. At this stage of replication, there must be reversible rotation around the replicate fork preventing innately immobilization of the replicated DNA molecule during mitotic segregation. Topoisomerase II mediates the rotation around the axis by cutting the DNA strands and closing the ruptures again. Doxorubicin DNA intercalation stabilizes the DNA- topoisomerase II complex after the filaments have been cut off, disrupting DNA replication at this point. Doxorubicin inhibits DNA and RNA synthesis compromising the cell replication [40, 50, 53]. In parallel, doxorubicin converted into a semiquinone through an oxidation process. Semiquinone is an unstable metabolite which returns to the doxorubicin structure by realising reactive oxygen species. The reactive oxygen species cause lipid peroxidation, membrane damage, DNA and proteins damage, oxidative stress, and finally activating the cell death by apoptotic pathways [40, 54, 55].

The main adverse event caused the doxorubicin treatment is cardiotoxicity. Dysrhythmias and heart failure are this drug side effect considered as a possible consequence of free radical generation. The cardiac damage caused by doxorubicin is proportional to the dose consumed, in this sense, presenting the total cumulative dose as the only criteria currently used to predict its toxicity [53, 54]. However, due to indications of doxorubicin distinct pathways for anti-

tumour activity and cardiotoxicity, better understanding of this drug mechanism of action may result into treatments with equal efficacy and reduced toxicity [40].

2.4.3. Resistance

The occurrence of relapsed and refractory disease in 40% of DLBCL patients after R- CHOP treatment is a clear indicator that this chemotherapy regimen is still presenting limitation [58]. Drug resistance is a potential cause of the elevated percentage of R- CHOP unsatisfactory outcome, culminating into disease progression and consequently patient life quality impact. Treatment resistance can be categorized as intrinsic and acquired, based on the initial positive response to drug therapy. Intrinsic resistance is characterized by an initial negative response to the treatment. It is related to specific molecular characteristics of the tumor identified at the diagnostic which enable the proliferation of malignant cells even in patient under drug treatment. In this sense, intrinsic resistance results into refractory diseases. On the other hand, acquired resistance is marked by an initial positive tumor response to the treatment which become limited due to growth of chemo resistant subpopulations [59-61]. Detection of biomarkers is fundamental for not only predict the drug treatment effectiveness, but also to indicate treatment resistance or refractoriness.

2.5. Biomarkers

According to the National Institute of Health, biomarkers are specific indicators of biological and pathologic processes as well as pharmacological responses. Biomarkers are classified into diagnostic, prognostic and predictive based on the type of information it aims to provide about the mechanism that has been investigated [62]. To detect a disease or an inaccurate physiological mechanism, highlighting the hallmarks of a mechanism under homeostatic and sick conditions the diagnostic biomarkers are applied. To evaluate the overall survival or progression free- survival for patients submitted to the same therapy regimen the prognostic biomarkers are utilized. Finally, predictive biomarkers are used to predetermine the patients response to a specific drug treatment. Independent of the classification, elevated specificity and sensibility are characteristic requested for an appropriated biomarker [63]. Currently, the Early Detection Research Network has developed guidelines for the identification of diagnostic biomarkers, aiming to establish biomarkers with predictive and prognostic potential. For this the possible biomarkers must be submitted to preclinical analysis, tested in retrospective and

prospective cohort. It must distinguish malignant and not malignant samples, identify pre-cancer indications and demonstrate benefits as an early diagnostic for the analysed population [64]. Hallmarks antigens for immature B cells, germinal center B cells and plasmas cells as well as regulatory molecules for B cell differentiation are commonly utilized for DLBCL characterization, subtypes classification and prognosis [65, 66].

2.5.1. CD19

The human CD19 antigen is a 95 kd transmembrane glycoprotein. It composes the immunoglobulin superfamily and it is classified as a type I transmembrane protein [67, 68]. CD19 is considered a hallmark for B cell differentiation and maturation process due to the marked surface density variation of this antigen at B cell lymphopoiesis and mature B- cells. At the terminal plasma cell differentiation, CD19 expression become absent. In this sense, CD19 is commonly used as a biomarker for normal and malignant B cells studies. CD19 is first expressed during B cell development concomitant with the immunoglobulin gene rearrangement. However, the highest expression of CD19 is only reached at mature B cells, where its expression is 3- fold higher than on preceding B cell stages [67]. Although it is still not clear the directly CD19 contribution for the B cell malignancy, this antigen expression is maintained at normal or high level on 88% of B cell lymphomas, including DLBCL [67, 69].

Overall, CD19 modulates BCR dependent and independent signaling in the presence of PAX5, developing BCR co- receptor function. The co- receptor CD19 functionally alters BCR signal transduction increasing signals through pre-BCR/BCR. For the B cell activation, CD19 develop function as an adaptor protein to gather cytoplasmic signaling proteins at the membrane. In addition, CD19 colligated with BCR act as a signal subunit for the CD19/CD21 complex. CD19 colligation with surface Ig results on the increase of B cell activation [67, 68]. CD19 complex modulates intrinsic and receptor-induced signals, diminishing the threshold for receptor dependent signaling. Hence, CD19 modulate B-cell fate at multiple stages of development [67, 68, 70].

2.5.2. CD20

The human CD20 antigen is a transmembrane protein of approximately 35 kDa apparently traversing the membrane four times. CD20 is considered a putative calcium channel and B lymphocyte specific cell surface molecule [71-73]. CD20 is high expressed during multiple stages of B cell development and differentiation encompassing the immature and mature B cell phases. It is predominantly expressed after Ig heavy chain rearrangement. CD20 expression is extinguished at the terminal differentiation into plasma cells, not being expressed by plasma cells. Thus, CD20 is also a hallmark for B cell lymphocyte, and appropriate biomarker for this cell studies on normal and cancerous conditions [72,73]. High majority of B cell lymphomas present CD20 expression and 98% of DLBCL demonstrated CD20 expression, indicating CD20 favorable function on tumor B cell maintenance [34]. Physiological function of CD20 is still not yet elucidated. However, it is believed that CD20 participates of B-cell activation, proliferation, and calcium transport process [73, 74]. CD20 have also demonstrated pro-apoptotic effects mediated via tyrosine kinase-dependent signalling pathways as well as apoptotic suppression at germinal center cells [73, 75].

2.5.3. CD38

The human CD38 antigen is a single chain glycoprotein of 45 kDa, classified as type II transmembrane protein and presenting both enzymatic and receptor functions [76, 77]. CD38 is highly expressed on germinal center B cell and mature plasma B cells in humans. It expression increases accordingly with the B cell maturation. In this sense, CD38 is not only an appropriate biomarker for B cells development, differentiation and maturation but also it has different function for each of these phases. Malignant B cell conserve CD38 expression on human progenitors B cells [73, 74]. Phosphorylation of substrates, intracellular calcium release, and increased expression of molecules involved in proliferation or apoptosis are the physiological effects promoted by the CD38. This antigen presents opposite functions according to the B cell stage, repressing or stimulating this cells proliferation. On bone marrow immature B cells, CD38 crosslinking produces a negative signal causing the growth arrest of the cells [77]. On germinal center B cell, CD38 interaction with monoclonal Ab promotes a positive signal culminating into cellular survival and resistance to pro apoptotic factors activities [76].

2.5.4. Other biomarkers

Lambda and kappa. During the early development of B cell, a sequence of allelic exclusion events on the light chain of BCR occur to ensure a variable and joining gene segments. Consequently, mature B cell light chain present only one class of immunoglobulin, rather kappa or lambda. It is not completely elucidated exactly how the allelic exclusion is regulated or which factor determine the light chain kappa or lambda formation. Aiming to explain this process two models were developed, the regulated and the stochastic models. For regulated model, kappa rearranges are the first light-chain option and only in case of non-productive or self-reactive kappa rearrangements the lambda occurs. The stochastic model proposes that isotypic exclusion results from high incidence of non-productive rearrangements and by a different recombination frequency of kappa and lambda light-chain genes, favouring kappa rearrangements. The occurrence of double- light chain expression as well as light chain restriction (expressing only kappa or lambda) is common among B cell malignant [79, 80] Plasma cell neoplasm expressing exclusively kappa or lambda are considered more aggressive and present a shorter survival for patients [81].

CD45. CD45 antigen presents a multiple isoform with three exons in the extracellular domain and molecular weight varying from 180kDa to 240KDa. CD45 is commonly expressed during B cell development, but not fundamental for B cell maturation which had been demonstrated by CD45 deficient animal studies [82]. As its main function, CD45 negatively regulates BCR signaling in immature B cell through modulation of signal transduction and BCR threshold [82, 83].

CD10. CD10 antigen is a single-chain glycoprotein with a molecular weight of 100-kd. This antigen is under normal conditions expressed on pro- B cells and mature GC B cells. CD10 is highly expressed on malignant B cells including 20%- 30% of *de novo* DLBCL cases and it elevated expression has been associated with shorter overall survival. On cancerous cells, it is assumed that CD10 inhibit regulatory peptides function, enabling progression of malignant cells differentiation [84, 85].

IgM. The IgM presents immunological memory function and is a marked characteristic on early B cell receptor and memory B cells. Antibodies produced at germinal center result from first antigen interaction, presenting high specificity and mainly correspond to IgG isotype. IgM

memory B cells results from a T- cell independent interaction and compose the first line defence of our body against infections, presenting antibacterial specificity [4, 86, 87]

CD27. The human CD27 is a type I glycoprotein, composing the tumour necrosis factor (TNF) receptor family [88-90]. This antigen is differently expressed on B cell subpopulations, presenting positive expressing on health adults and high expression on auto immune disease patients. In this sense, B cell subpopulations may present high or low expression of CD27 which put in question that antigen efficiency as a biomarker [88]. However, CD27 increased expression with age, immunoglobulin production, localization within the marginal zone, the presence of mutations in Ig V region genes, and their improved ability to differentiate to plasma cells indicate that CD27+ B cells are memory B cells. Moreover, studies on CD27+ B cell indicate their contribution for the plasma cell differentiation. Oligoclonal expansion of naive CD27 B cells is mediated by activated T cells helpers. During this process, SHM is registered in part of B cell resulting on CD27 surface expression and finally differentiation into memory B cells. Activated T cells expression CD70 interacts with CD27 on memory B cells culminating into plasma cells generation [90]. For this reason, CD27 is currently utilized as a biomarker to differentiate naïve B cell and memory B cells.

CD5. The antigen CD5 is 67-kDa surface glycoprotein, expressed by specific B cell subpopulations [91]**.** CD5 is commonly associated with B cell malignancies, favouring cancerous transformations by producing multi-specific antibodies. In normal conditions, CD5 recognize autoantigens cross reacting with a variety of bacterial antigens. Early B cell expressing CD5 are induced to apoptosis by CD72. However, the B cells which cognately interact with T cells, loose the CD5 expression and advance to the germinal center [92].

CD23. The antigen CD23 is an integral membrane glycoprotein participating on the regulation of IgE synthesis. It composes the IgE complexes promoting elevated antigen interaction with T cells and it also regulates IgE production, enabling B cell survival on germinal center by interacting with CD21. In addition, CD23 facilitates the antigen interaction from B cells to T cells. CD23 is highly expressed during B cells differentiation, favouring this cells conversion into IgE-secreting plasma cells. However, CD23 expression is downregulated by mature B cell after isotype switch and memory B cell differentiaton [93, 94].

CD43. The human antigen CD43 is a heavy transmembrane glycoprotein with 235 amino acid extracellular domain expressed in a variety of hematopoietic cells, including B cells at normal and pathologic conditions. This antigen expression is marked at early B cells after BCR stimulation, and mature B cells where it regulates cell proliferation through PKC mechanism. However, it is not present on the transitional phase of this cell. In addition, CD43 is also expressed on plasma cells [95-97]. This antigen regulates apoptosis and B cell differentiation and immune homeostasis. In cancerous cells, abnormal expression of CD43 inhibit B cell G1 arrest prolonging B cell survival. In this sense, CD43 favours unregulated B cell proliferation and improved survival capacity, markers of malignant cells development. Approximately 25% of DLBCL cases express CD43 and this antigen has been considered a bad prognostic factor [93, 98].

2.6. Tumor heterogeneity

The tumor formation results from clonal evolution of malignant cells marked by a myriad of stochastic mutations where only the alteration conferring competitive advantage are maintained [99, 100]. Hence, cancer tumours are characterized by a high cellular heterogeneity which results from the malignant cells organization and specific tumor features. The cancer stem cell model, a model for cancer development and tumor composition, proposes that cancers are organized into malignant stem cells subpopulations and non-malignant stem cells subpopulations hierarchically organized. The subpopulations parallelly conduct tumor growth and illness progression, potentially culminating into drug resistance and metastasis. Hence, the cancer stem cell model (CSCM) partially clarifies the phenotypic and morphologic heterogeneity on cancer cells composing the same tumor. Although these mutations cannot immediately confer epigenetic modification (Figure 2.6a). However, it is still not elucidated the percentage of the malignant cell on the tumor following this model [101]. It also necessary to highlight the cancer cell plasticity, allowing these cells to change into different states. In this sense, they can obtain new mutations on cells phases prone to cellular alteration and genetic modifications, simultaneously maintaining the phenotypical equilibrium for the tumor. Gene expression is the main regulator of cell-state decisions. The phenotypically balance is maintained by cell proliferation mediated through intercellular signals until reach the equivalent proportions of cells in different states or cell transition into different states. These observations are based on Markov model which mainly proposes that cells transited stochastically between states [102].

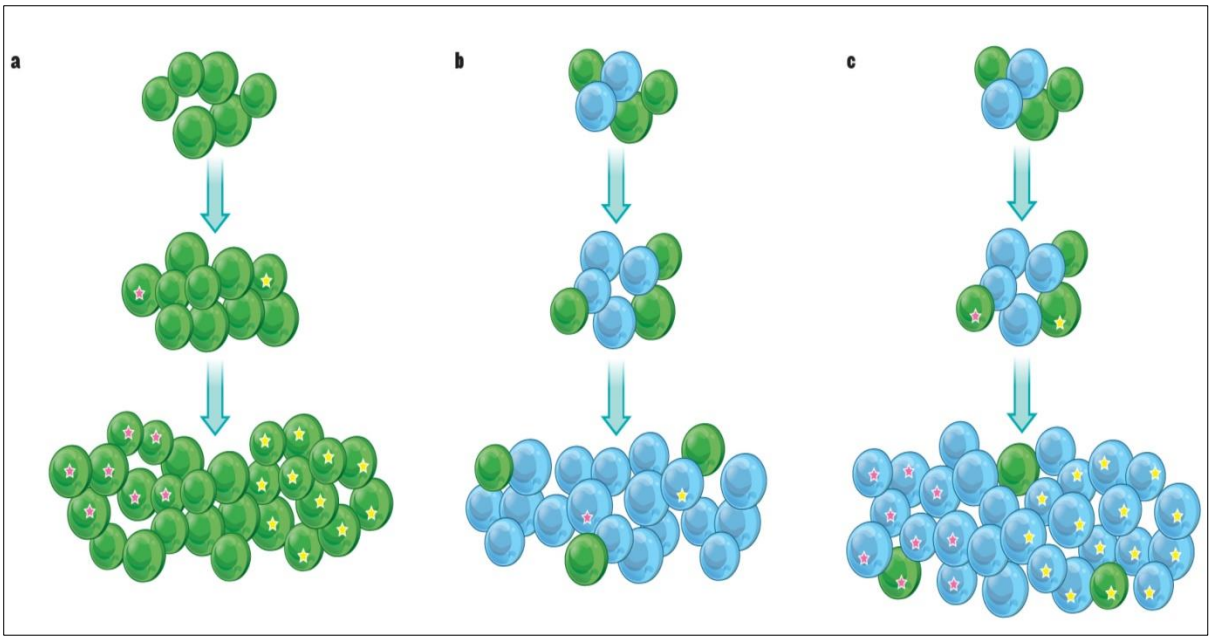

Figure 2.6a: *Cancer heterogeneity for cells following the stem cell model.* Subpopulation evolution and differentiation of tumorigenic cells into non-tumorigenic cells is an independent, or associated processes. Different mutations are characterized by different star colours. (a) New mutations which promote phenotype or function alteration elevate heterogeneity on tumours. (b) Tumorigenic cells differentiation results into non-tumorigenic progeny, establishing tumour heterogeneity. In this case, non- tumorigenic cell mutations are only propagated to restore tumorigenic potential. (c) In case of mutations in tumorigenic cells, tumor heterogeneity results from clonal evolution and differentiation of tumorigenic into non-tumorigenic. Thus, phenotypic and functional alteration cannot immediately result into epigenetic differences between tumorigenic and non-tumorigenic cells, indicating the high relevance of genetic alteration for the tumour heterogeneity (Source: Meacham and Morrison *et al.* 2013) [101].

Chemotherapeutic drugs exposition combined with genetic alteration also contribute for tumours heterogeneity. Due to cancer therapies higher efficient against specific cellular states, chemotherapeutic drugs regimen might favour selective change of phenotype for specific tumor subpopulations [103-106]. In this sense, drug resistance is also a plastic feature of cancer cells as tumor subpopulation might demonstrate resistance or sensitive response to a drug according to the cellular state on the moment of drug exposition [107]. Moreover, cells presenting high genetic heterogeneity composing the same tumor can generate new tumours with elevated genetic heterogeneity, distinct phenotype and consequently different functional characteristics. Experiments demonstrate that minor subpopulations become the dominant subpopulation after submitted to drug therapy, indicating that genetic alteration promoted by chemotherapy exposition contribute to drug resistance. It indicates variable probability of survival among the different cells determined through genetic differences between subpopulations. Dominant subpopulations which remain dominant after relapse will potentially develop *de novo* mutations. Therapy resistance might be acquired through relapse- specific mutations (Figure

2.6b) [101]. Tumour heterogeneity also represents an extra challenge for predictive and diagnostic biomarkers due to the evasion or small fraction of the subpopulation affecting the treatment outcome in the moment of diagnostic [49].

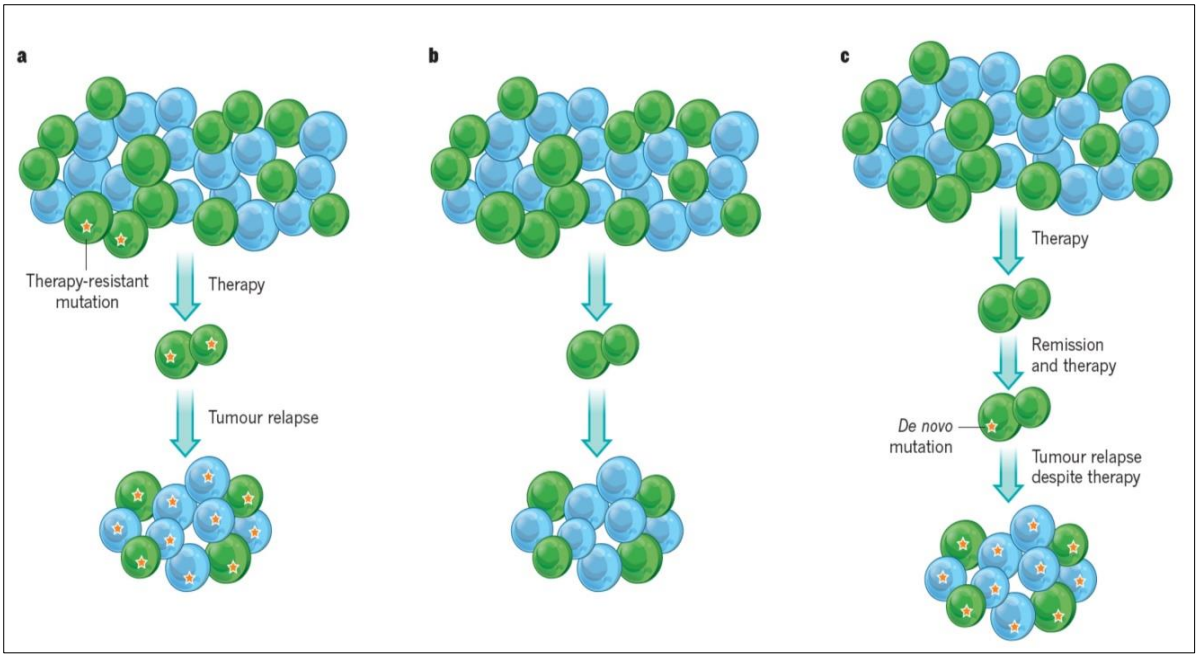

Figure 2.6b: *Genetic alterations and tumorigenic cells characteristics can independently, or associated result into therapy resistance*. Different mutations are characterized by different star colours. (a) Genetic alterations as main cause of therapy resistance. In this sense, the cells presenting specific genetic alteration are resistant to the therapy exposition culminating into tumor relapse. (b) Tumorigenic cells characteristics as cell plasticity allow certain cancers inherently resistance to specific therapies. (c) Tumorigenic cells characteristic and genetic alteration obtained after therapy exposition favour disease relapse. The acquisition of *de novo* mutations promoted by the dominant subpopulation after relapse improve therapy resistance, favouring disease progression (Source: Meacham and Morrison *et al.* 2013) [101].

2.6.1. Previous works of importance for this study

Considering the genetic heterogeneity caused by subclones evolution, genetic and epigenetic alteration as a marker feature among different types of tumours, Quentmeier *et al* 2013 proposes an *in vitro* model for cancer heterogeneity study utilizing the DLBCL cell line U2932. This cell line demonstrated the coexistence of two distinct flow-sorted subpopulation, R1 and R2. The subpopulations present different phenotype identified through cell surface markers CD20 vs CD38 expression (Figure 2.6.1). Each subpopulation manifested a stable immunophenotype 36 days after sorting, not regaining the original cell line phenotype after 100 days of sorting. In this study, U2932 subpopulations were traced as subclones of the original tumor presenting clone-specific immunoglobulin IgVH4–39 hypermutation patterns. *BCL6* and *BCL2* was overexpressed in R1 whereas *MYC* and *BCL2* was overexpressed for R2.

Immunoglobulin hypermutation and cytogenetic analysis indicate that both subpopulations originate form a mother clone with genomic *BCL2* amplification. Each subpopulation differently obtained secondary rearrangements resulting into *BCL6* or *MYC* overexpression. Microarray analysis also demonstrated different gene expression for the transcriptional targets of the aberrantly expressed oncogenes in each subpopulation. Epigenetic alterations were also indicated by DNA methylation analysis [108].

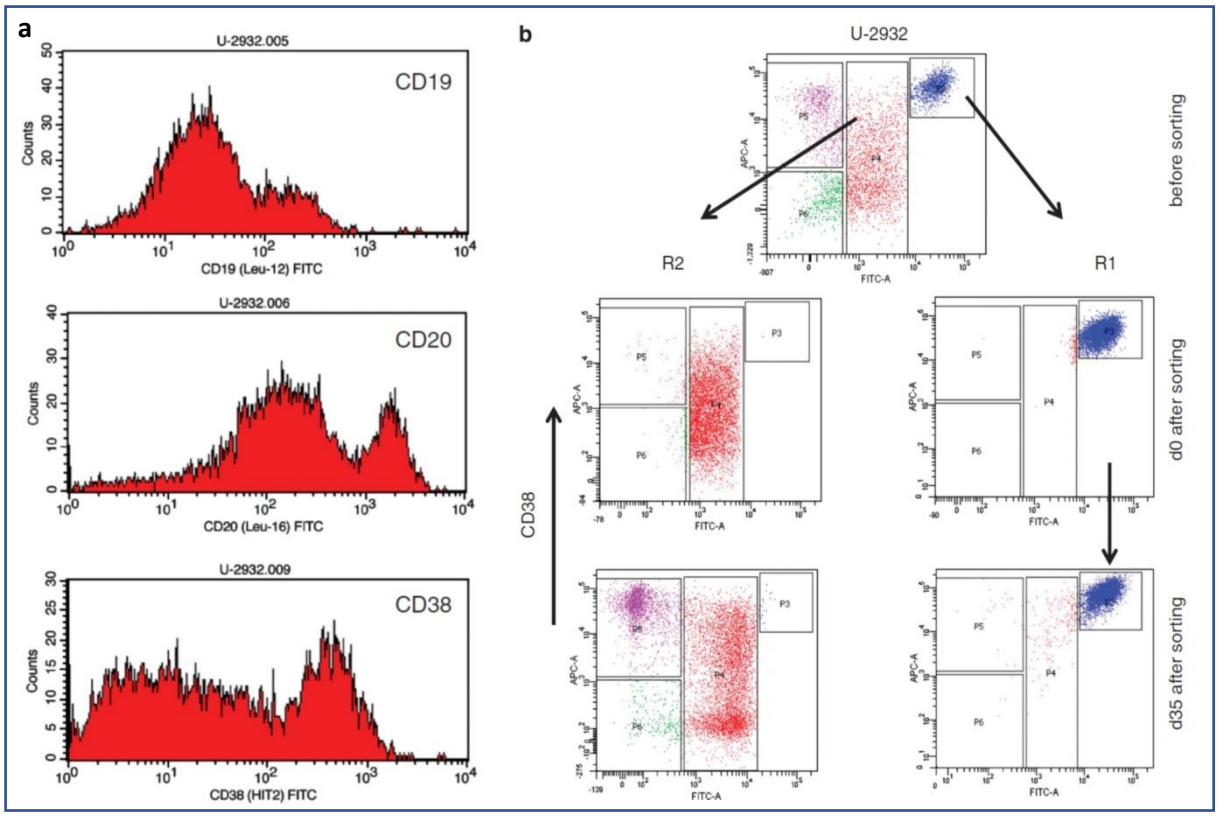

Figure 2.6.1: *U2932 cell line immunophenotype*. (a) Flow cytometry expression of CD markers was analysed demonstrated double peaks for CD19, CD20 and CD38. (b) Only CD20 and CD38 demonstrated phenotypic subpopulations for U-2932. R1 expression of CD20 and CD38 is high and phenotypically stable. R2 is composed of three sort gates for which pure and stable subpopulations sorting was not considered possible. After 35 days, R1 and R2 CD20/CD38 expression remains stable and distinct from each other, confirmed by 100 days reanalysis (Source: Quentmeier *et al*. 2013) [108].

Nevertheless, an initial trial developed by our group aiming at characterizing the subpopulations in the U2932 cell line identified two distinct subpopulations based on CD19. On Quentmeier *et al*. 2013 analyses, CD19 population did not maintain their expression, regaining the original cell line phenotype after 35 days. On the other hand, our first trial has determined the stable expression of CD19 high and CD19 low in duplicate, after 29 days (details on appendix 1, figure 1). Therefore, our experiment showed heterogenous expression of CD19 in parallel to CD20 and CD38. Further analyses of the CD19 subpopulations were conducted aiming to determine their growth rate and drug screen sensibility to vincristine. The

growth curve and drug screen analysis showed no statistical difference in the growth rate between the two sorting groups.

3. Hypothesis, Aim & Objectives

Background:

DLBCL is a highest incident NHL diagnose worldwide marked by mutations in different states of B cell differentiation resulting in the distinct types of this tumor, GCB and ABC, which has been directly related to patients' prognostic and overall survival. According to the cancer stem cell model, tumours heterogeneity characterized by the collaboration of the distinct subpopulations on the same tumor environment favours cell proliferation and illness progression, potentially culminating into drug resistance and metastasis. Moreover, Markov model proposes that elevated cancer cell plasticity allows it to assume new cell states mediated through alteration on specific gene expression. In this sense, cancer cells can stochastically transitate between states until the tumor reach phenotypic equilibrium. Additionally, chemotherapeutic drugs demonstrate higher efficient against specific cellular states, favoring selective change of phenotype for specific tumor subpopulations. Interestingly, Quentmeier *et al* 2013 identified the coexistence of different subpopulations in the U2932 cell line, a DLBCL cell line which has been previously submitted to the R-CHOP regimen. Each subpopulation demonstrated a steady phenotype for 35 days combined with altered gene expression of regulators for the B cell differentiation.

Hypothesis:

The U2932 cell line is a highly heterogenous cell line composed of different subpopulations at distinct B cell states of differentiation marked by a different phenotype and drug resistance.

Aim:

Characterize U2932 cell line and its subpopulations based on immunophenotype, transcriptional profiling and sensibility to the drug: doxorubicin.

Objectives:

I. To determine the existence of subpopulations in the U2932 cell line based on CD markers expression in order to compare and evaluate their phenotype stability as well as determine the B cell subset that each subpopulation comprehend to.

II. To characterize the drug sensibility of the U2932 cell lines and its subpopulations to Doxorubicin, molecularly justifying possible response variances.

III. To determine the expression of genes regulating B cell differentiation, establishing the pathway enrichment for U2932 cell line and its subpopulations based on their gene alterations.

4. Materials and Methods

The methodology utilized in this project is grouped in four main sets of experiments (Figure 4). The first set of procedures verified the coexistence of subpopulations in the U2932 cell line (obtained from the German collection of microorganism and cell cultures, DSMZ) based on surface on CD markers expression (CD19, CD20, CD38). Subsequently, purification of the defined subpopulations from the original cell line by FACS resulted in four subpopulations, P3, P4, P5 and P6. (Figure 4.1A). The second set of procedures investigated the immunophenotype stability of U2932 cell line and subpopulations from day 1 to 36 by flow cytometry (Figure 4.1B1). Simultaneously, the growth pattern of each subpopulation was characterized by growth curve analysis (Figure 4.1B2). The third set of procedures focused on characterizing the transcriptome profile of U2932 cell line and subpopulations by microarray on day 36 after (Figure 4.1C). Finally, U2932 cell line and subpopulations sensibility to doxorubicin, a chemotherapeutic drug utilized in DLBCL therapy, was accessed after 36 days from sorting by drug screen analyses (Figure 4.1D). Doxorubicin is a chemotherapeutic drug utilized in DLBCL therapy). The immunophenotype characterization and drug screen analyses were repeated on day 93 and 100 after sorting.

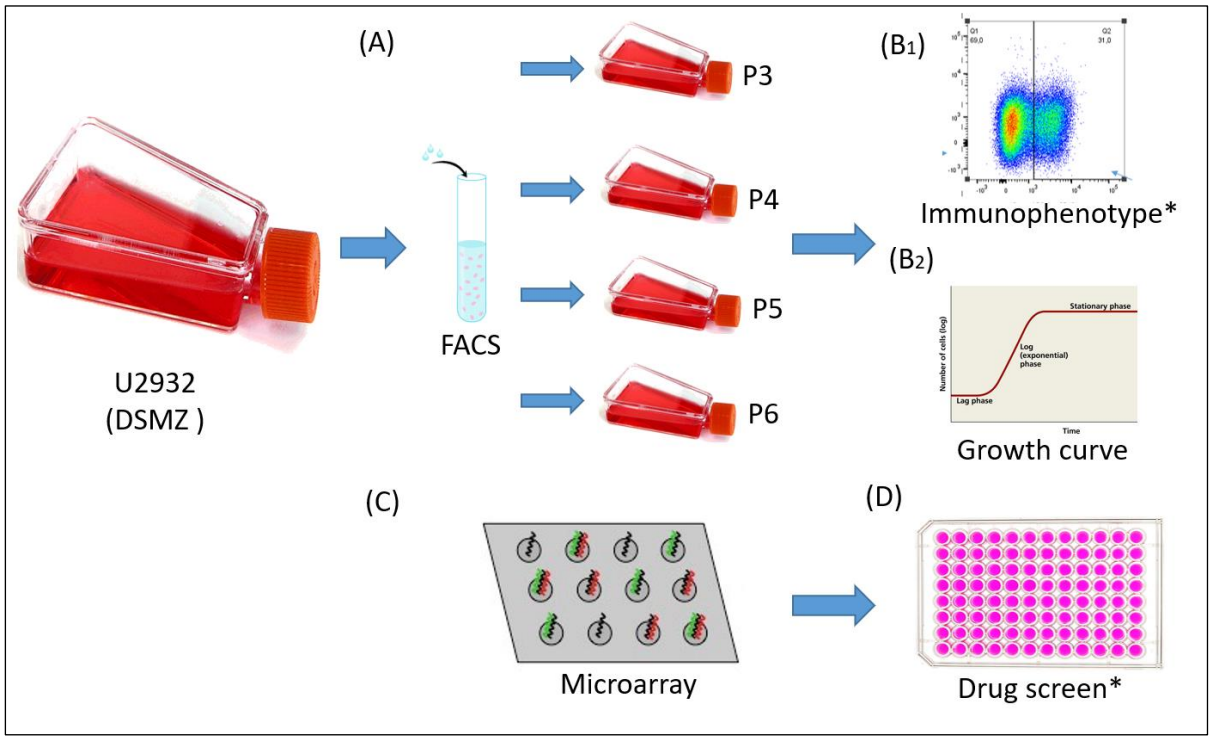

Figure 4: *Overview of project procedures.* U2932 was submitted to FACS based on distinct CD markers expression of cell populations composing the cell line (A). The stability of the subpopulations CD markers expression was accessed assessed through repetitive immunophenotype analyses by flow cytometry during 36

days after sorting (B1). Parallel to this procedure a growth curve of each subpopulation and the original cell line were established (B2). Microarray analyses of the original cell line and subpopulations on day 36 after sorting (C). Drug screen analyses of the U2932 and subpopulations and the original cell line (D). *At 93 to 100 days after sorting the stability of the subpopulation CD markers expression and drug screen analyses were repeated utilizing the same parameters of previews experiments.

4.1. U2932 cell line

The U2932 cell line was purchased from the German collection of microorganism and cell cultures (DSMZ). This cell line is derived from the ascites of abdominal area of a 29-year-old female patient diagnosed with DLBCL.

4.1.1 Patient historic

The patient was initially diagnosed with Hodgkin lymphoma nodular sclerosis type 1 based on lymph node biopsy, being the mediastinum, lung hili and pleural effusions the initial areas affected. After 14 years from the original diagnosis, an enlarged tonsilla biopsy showed an DLBCL. The patient was again submitted to radio and chemotherapy including EPOCH (etoposide, prednisone, vincristine, cyclophosphamide, doxorubicin), MIME (methyl-GAG, ifosfamide, methotrexate, etoposide) and DHAP (dexametasone, cisplatin, araC) followed by high-dose CT with BEAC (BCNU, etoposide, araC, cyclophosphamide). After 15 months from the initiation of the DLBCL treatment the patient presented a relapse which advanced to ascites followed by the patient death. The cell line was established from the ascites of the last relapse of DLBCL tumor and presented a small round morphology with cells growing singly but mainly in clusters in suspension. This DLBCL cell line was classified as an ABC- like lymphoma and cells were described to overexpress *BCL2* and *BCL6* [109]. Detailed information about U2932 cell line in appendix 1, figure 2.

4.1.2. Tissue culture of the U2932 cell line

A total of $2.5e10^6$ ascites tumor cells purchased from the DSMZ company were centrifuged on 1200rpm per 5 min, resuspended in RPMI 1640 medium (Gibco) containing 20% fetal calcium serum (FCS, Gibco), and 10% antibiotic (penicillin), maintained at 37 ^{o}C in a 5% CO_2 humidified incubator. After the first seven days, the FCS concentration in the media were

adjusted to 10%, the other parameters were maintained. Due to its high proliferation, the cell line soon reached an elevated number of cells, and it was transferred to a $75cm^2$ culture flask, receiving 25ml of media (RPMI-1640 + 10% FCS + 1% penicillin) after one week from the initiation of the cell line growth procedures. For 21 days, the cells were passaged with 72h of interval maintaining a density of $0.5e10^6$ cells/ml with a 48- 50h doubling time and 98% viability. The original U2932 cell line were split seven times before submitted to any further procedures, being able to reach a growth rate stability. The subpopulations later obtained from the original cell line sorting were maintained under the same parameters.

4.2 Multicolor Flow Cytometry & Fluorescence-activated cell sorting (FACS).

The U2932 cell line and subpopulations were labelled with combination of differentiation markers (Florescence labelled monoclonal antibodies) which had previously demonstrated to characterize distinct stages of B cells differentiation and B cell identity. All samples were centrifuged 540g per 5min, resuspended in 100µl of stain buffer (SB), and incubated with the markers for 30 min in dark. For U2932 initial labelling, $18e10^6$ cell were stained and utilized for subpopulation identification. The markers (all from BD Bioscience, San Jose, California) utilized in the initial monoclonal Ab panel for U2932 subpopulation identification were: 30 µl per test of CD20 (V450 clone 2H7, Lot 6342902), 30 µl per test of CD38 (APC-H7, clone HB7, Lot 5341728), 30 µl per test of CD19 (APC, clone SJ25C1, Lot 86207), 30 µl per test of CD25 (PE-CY7, clone M-A251, Lot 88260), 30µl per test of CD59 (PE, clone p282, Lot 7012666/ 38956), 30 µl per test of CD80 (PE, clone L307).

For each subpopulation and the original cell line, $2e10^6$ cells were stained and phenotypically characterized at each test. The markers utilized in the monoclonal Ab panel for U2932 sorting and phenotype characterization of each subpopulation were: 5µl per test of CD20 (V450 clone 2H7, Lot 6342902), 5µl per test of CD38 (APC-H7, clone HB7, Lot 5341728), 5µl per test of CD19 (APC, clone SJ25C1, Lot 86207). A final monoclonal Ab panel were established, aiming to provide a more detailed immunophenotype characterization of U2932 and subpopulations. The markers (all from BD Bioscience, San Jose, California) utilized in the final monoclonal Ab panel were: 5µl per test of CD20 (V450 clone 2H7, Lot 6342902), 5µl per test of CD38 (APC-H7, clone HB7, Lot 5341728), 5µl per test of CD19 (APC, clone SJ25C1, Lot 86207), 20 µl per test of IgM (FITC, G20-127), 5µl per test of CD27(PcPCy5, clone 0323), 20 µl per test of CD10 (PE, clone HI10a), 5µl per test of CD5 (PcPCy5, clone UCHT2), 20 µl per test of

CD23 (PE, clone IBVCS 5), 5µl per test of CD43 (APC, clone 1G10), 5µl per test of lambda (FIT, clone 1552), 5µl per test of kappa (PE, clone TB28-2) and 5µl per test of CD45 (HV500, clone 2D1). FACS Aria II cell sorter or FACS Canto 2 (both from Becton Dickinson, Heidelberg, Germany) with 3- laser (violet, blue, red) standard configuration were utilized for the immunophenotype analyses of U2932 cells labelled with the mentioned CD markers.

After incubation, all samples were washed with SB and centrifuged at 540g for 5 min. The resulting pellet was washed and resuspended with 200-700 µl phosphate buffered saline (PBS). For the analysis preparation, the cytometer was submitted to a "fluidic startup" set up in the BD FACSDiva software. This procedure is recommended by the manufacturer's avoiding the tool malfunction during the samples analyses and the contamination of the samples with previous biologic material analyzed. Sequentially, FACSDiva software (BD Biosciences) promoted samples acquisition and the compensation values were determined based on a control samples stained with the same monoclonal Ab utilized in the experiment.

For the sorting, a total of 50,000 events per tube were acquired to define the gating regions of U2932 subpopulation. CD38/CD20 population density was the main parameter for gating and identify the distinct subpopulations of U2932 cell line in a psedocolor plot using the embedded software of the instrument, BD FACSDIVA software. Immediately after sorting each subpopulation was resuspended on RPMI growth media (containing 10% FCS and 1% Penicillin), the total volume of the media varied according to the total number of the cell obtained for each subpopulation. The U2932 cell line, source of the cells for the sorting procedure, was still maintained on the appropriate conditions previously described. For this experiment, no bulk population was established due to the inexistence of difference between a bulk population and the original U2932 cell line based on previous trials. For the immunophenotype characterization, cell population density and frequency of distribution were the main parameters utilized. In each analysis, 500.000 events were evaluated per subpopulation. The department specialists developed all FACS procedures. The FLOWJO 7.0 (BD Bioscience, San Jose, California) was the software utilized for the data analyses.

4.3. Growth Curve

To characterize growth rate of the U2932 cell line and its subpopulations a growth curve experiment was conducted. For all samples counting was done in triplicates and in two concentrations. The U2932 cell line presents a doubling time of approximately 50h. For this reason, the growth curve for this cell line was done in $0.5e10^6$ cell/ml and $0.25e10^6$cell/ml. The

lower concentration was included to certify that the cells growth decline after the initial 5 days of experiment was not caused by the "overpopulation" of cells. The procedure set up occurred on day 17 after sorting, all subpopulations and original cell line were on passage 14 and 15, respectively. In a 6 well plate, 5ml of each concentration were added in three wells of the plate. In this sense, each six well plate presented sample of only one subpopulation avoiding contamination. The day of platting was considered day 0, the counting was only initiated 24h after platting and repeated on the following days at the same time.

The cells were counted daily for a period of 10 days to collect the data covering the lag, exponential and stationary phases. The counting was done manually using a hemocytometer (Bürker-Türk counting chamber) and a fluorescent microscope (Nikon Eclipse, TS100). To allow a better visualization of cells and to distinguish between the live and dead cells, a dye exclusion test was performed. 40µl of cell suspension from each well was mixed with 40µl of trypan blue 0.4% (Gibco, Ref15250- 061) resulting in 2-fold dilution. Trypan blue penetrates the damaged membrane of the dead cells providing a blue coloration for the dead cells cytoplasm. Due to manipulation error, the growth curve procedure was repeated for the P3 cell line 55 days after sorting (passage 23), under the same parameters previously described.

4.4. Pictures

To demonstrate the cellular morphology and distribution for U2932 and subpopulations, pictures were taken of each sample on day 12, 36 and 100 after sorting. The samples visualization was done utilizing a fluorescent microscope (Nikon Eclipse, TS100) with Nikon lens 10x amplification. To capture the images a camera (Nikon Instruments, Europe, B.V.) and camera NS element F400.06 software were utilized. All imagens presented 1000px and were posteriorly edited on NS element F400.06 software allowing a brightness adjustment.

4.5. RNA expression analyses

4.5.1 RNA purification

Dried pellet containing $5e10^6$ cells were prepared by spinning the cells on 1200rpm for 5 min and resuspending the samples on trypsin on day 36 after cell sorting for the U2932 cell line and its subpopulation. The samples were stocked on 180°C until for two weeks until continuation of procedures. RNA purification was performed by combining the protocols of the TRIsure reagent (Invitrogen) and mirVana miRNA isolation kit, as previously described by our group

[110]. In details, following TRIsure reagent protocol, isolation. The samples were homogenized and incubated at 30 °C allowing RNA dissociation of nucleoprotein complex. The cells were resuspended in 1mL TRIzol reagent and incubated for 5 minutes at room temperature. Sequentially, 200µl chloroform was added and samples were homogenized. The resulting samples were centrifuged at 12000x g for 15 minutes at 4°C. After 2 minutes of incubation at room temperature and the RNA aqueous phase was isolated. For RNA isolation procedure, the *mir*Vana miRNA isolation kit (Thermo Fisher, Waltham, Massachusetts, USA) was utilized. The 100% ethanol was added in a 1.25 x volume of the isolated phase. The resulting lysate/ethanol mixture were filtered through a Filter Cartridge. The samples were added on purification spin column and washed once with miRNA wash solution and twice with wash solution 2/3. Finally, RNA elution occurred after samples being submitted to 50µl 95°C RNase-free water. The RNA purification procedure was done by the laboratory technician. The samples resulting from the previously described procedures were submitted to the microarray analyses.

4.5.2 Microarray analyses

Microarray is an efficient method to determine the entire transcriptional content of biological sample of interest in health or disease. This methodology has been applied in a myriad of DLBCL studies allowing identification of entire GEP of B cell lymphomas and the comparation between different samples, facilitating the characterization of disease subtypes and samples classification [31]. The procedure consists in the conversion of RNA into complementary DNA (cDNA) through reverse transcription. In addition, samples are individually labeled with distinct colors of florescent probe. Sequentially the cDNA molecules are submitted to hybridization to complementary oligonucleotide probes on the microarray. Finally, the microarray is scanned, measuring single gene expression through fluorescent signal determination.

Here the microarray was utilized to determine the global gene expression profile of DLBCL, U2932 cell line and its subpopulations, P4 and P5, enabling the posterior comparison among the different samples data. The analyses of each sample were conducted in duplicate. For GEP determination, Affymetrix GeneChip HG-U133 Plus 2.0 array was the microarray platform utilized. It classified DLBCL cell line and primary clinical samples through algorithms probabilities analyses as ABC/GCB, BAGS, and REGS [34]. HG-U133 platform enables the

coverage of 25,000 protein coding genes which might originate more than 47,000 mRNA transcripts translated as protein in this analysis. The entire microarray analysis was conducted by a laboratory technician. Nanodrop and Bioanalyzer analysis guarantied appropriate conservation of samples. The analyses of microarray CEL files were conducted on Partek Genomic Suits 7.0 software and Online One-By-One Microarray Normalization and Classification of Hematological Cancers for Precision Medicine (Hemaclass.org).

4.6. Drug screen

To identify the efficiency of currently available drug treatment for DLBCL in the sorted subpopulations compared to the original cell line, drug screen analyses with doxorubicin was conducted. The original cell line and three subpopulations, P3, P4, P5, were submitted to four doses of drug in distinct concentration. Two days after split each subpopulation was manually counted using a hemocytometer (Bürker-Türk counting chamber) and a fluorescent microscope (Nikon Eclipse, TS100), certifying the appropriate cell line density, $0.5e10^6$ cell/ml, for the cell suspension. For the doxorubicin analyses, the dose concentration was nominated concentration 0, 1, 2 and 3 (C0, C1, C2, C3). The C2 (0,16µg/ml) corresponded to the GI50 value of doxorubicin to the U2932 cell line previously determined in the department. The C3 (0,32µg/ml) and C2 (0,075µg/ml) correspond to the double and half concentration of C2, respectively. The C0 corresponded to the samples suspension submitted to isotonic saline only.

The procedure was developed in a 96 well plate where each dose of each sample was plated in triplicate, the procedure was done in duplicates on day 43 and 50 after sorting and repeated on day 100 and 107 after sorting. Wells were seeded with 120µl cell suspension per cell line and 30µl of drug or isotonic saline. To avoid evaporation and bias caused by environment interaction, the columns and lines on the plate extremity contained cell suspension and isotonic saline which data were not considered for analyses. Two columns containing medium and isotonic salt water were used for background correction. Dose-response measurements were conducted after drug exposure at 0h, and 48h. Two hours before absorbance analyses the samples were submitted to CellTiter 96 Aqueous One Solution Reagent. Absorbance of light was recorded at 492 nm using the microplate reader Fluostar Optima (BMG LABTECH, Ortenberg, Germany). The data was analyzed as described in 3.11.

4.7. Barcoding

The barcoding analysis allow the identification of DNA specificity utilizing short DNA sequence from a standard part of the genome as a marker, enabling the detection of common or distinct genetic material [111]. A dried pellet of the original cell line and each of its subpopulations were prepared on day 12, 26, 36 after sorting. Those samples were submitted to barcoding analysis to avoid the possibility of U2932 contamination with another cell line

4.8. Samples storage

After finalizing the lab procedure all sorted subpopulations established during this project were stored in liquid N2 at -196^{o}C, allowing it to be easily accessed for future experiments. Cell aliquots of 10e10^{6} cells per cell line were prepared. The high number of cells is necessary because during freezing a fraction of the cells is commonly damaged. The cells were stored in freezing solution containing 70% medium, 10% dimethyl sulfoxide and 20% FCS.

4.9. Statistical Analyses

For the growth curve analyses, one- way ANOVA test with Tukey multiple comparison post-test were utilized to determine the divergence on the growth pattern between U2932, P3, P4, P5 and P6. Only the data corresponding to the five first days of analyses of each subpopulation and original cell line were statistically evaluated as it corresponded to the exponential phase of the cells growth. The one- way ANOVA test was applied to each day individually, with no matching or paring and assuming a Gaussian distribution. The multiple comparison post- test compared the mean of each column with the mean of every other column, with Tukey as statistical hypothesis testing for multiple comparison reporting multiplicity adjusted P value for each comparison. Values bellow 0.05 were considered statistically significant (P value <0.05). All analyses were done on GraphPad Prism 7.0.

For the drug screen analyses, one- way ANOVA test with Tukey multiple comparison post-test were utilized. The statistical analyses determined the effect of different doses of doxorubicin in the same subpopulations and indicated occurrence of distinct sensibility among the subpopulation when exposed to the same dose of drug. The one- way ANOVA test was applied to each dose concentration individually evaluating the sensibility of the different subpopulation. In addition, one- way ANOVA test was applied to each subpopulation individually evaluating the effect to the four distinct doses. For one-way ANOVA analyses no

matching or paring were done and Gaussian distribution were assumed. The multiple comparison post- test compared the mean of each column with the mean of every other column, with Tukey as statistical hypothesis testing for multiple comparison reporting multiplicity adjusted P value for each comparison. Values bellow 0.05 were considered statistically significant (P value <0.05). All analyses were done on GraphPad Prism 7.0.

For the microarray analyses, one- way ANOVA was utilized as a streamline to determine the differentially expressed genes between U2932 and its subclones. The analyses were conducted separately for paired groups analyzed (U2932 vs P4, U2932 vs P5 and P4 vs P5). Only the genes with fold change higher than 2 and lower than -2 with P value bellow 0.001 (P value < 0.001) were considered for analyses. For the pathway enrichment analyses, Fisher's exact test was applied utilizing *Homo sapiens* specificity. Both analyses were conducted on Partek Genomic Suit 7.0. For specific genes of interest one- way ANOVA with Tukey multiple comparison post- test was utilized to evaluate its expression among different subpopulation. For one-way ANOVA analyses no matching or paring were done and Gaussian distribution were assumed. The multiple comparison post- test compared the mean of each column with the mean of every other column, with Tukey as statistical hypothesis testing for multiple comparison reporting multiplicity adjusted P value for each comparison. Values bellow 0.05 were considered statistically significant (P value <0.05). All analyses were done on GraphPad Prism 7.0.

5. Results

5.1. Immunophenotype of the starting cell line.

The cells labelled with CD59, CD80 and CD25 demonstrated a homogeneous expression for the initial U2932 cell line. Only the cells labelled with CD20, CD38 and CD19 expressed heterogeneity for U2932. Thus, only the CD markers which indicated cells heterogeneity expression were sequentially used for the sorting procedures. CD20 and CD38 high and low population density of expression are the main parameter for gating and identify the distinct subpopulations of U2932 cell line in a psedocolor plot. For the gating, only the areas of high density of each subpopulation were considered avoiding the inclusion of cells in the intermediate area between distinct populations (Figure 4.1a). In this sense, the U2932 cell line resulted in four subpopulations, P3, P4, P5 and P6. The total of cells obtained for each subpopulation varied based on the distinct percentage of representation of each subpopulation composing U2932 cell line. The P3, P4, P5 and P6 presented respectively $3e10^5$, $9e10^5$,$1.8e10^6$ and $6e10^5$ cells.

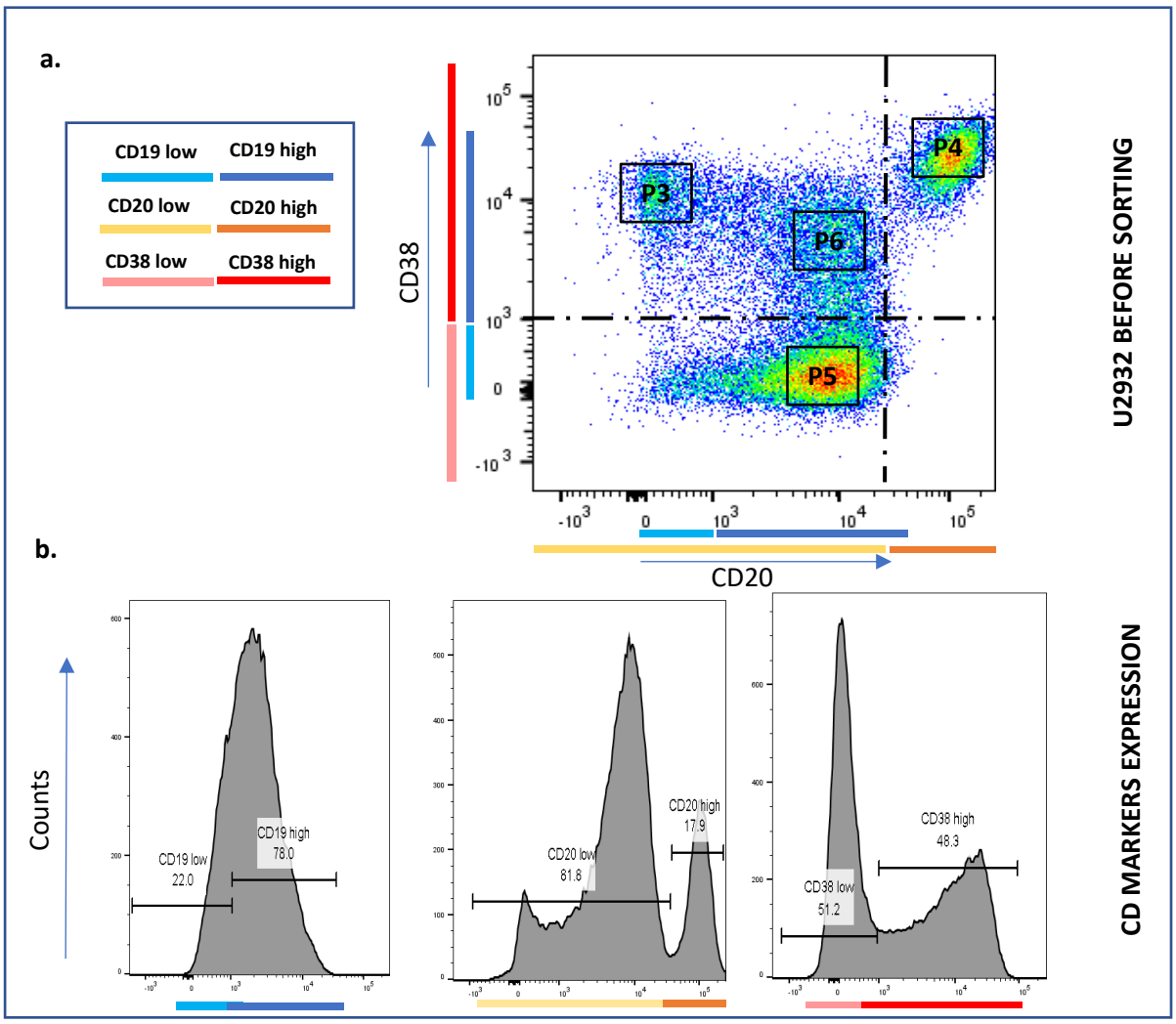

Figure 5.1: *U2932 cell line on its first day of immunophenotype analyses by BD FACSDIVA software*. (a) The B cell markers, CD20 and CD38, phenotypic expression by flow cytometry. Areas of high density, orange and yellow, and areas of lower density, green and blue, represent the cell line expression for CD38 and CD20 on the pseudo color plot. Different populations present a unique high- density area surrounded by low density area. For the sorting, only the areas of high density of P3, P4, P5 and P6 were considered for gating. (b) Histogram charts of CD markers expression demonstrating double peaks for CD20 and CD38.

The U2932 cell line flow cytometric analysis demonstrates a heterogenic immune phenotypic expression for the B cell markers, CD20 and CD38 (Figure 5.1). Overall, this cell line present predominant low expression of CD20 (81.8% of total cells analysed, table 1; appendix 2, table 1) and low expression of CD38 (51% of total cells analyzed, table 1, appendix 2, table 1), considering the cut off for the high expression of each marker at $1.5e10^4$ and $1e10^3$, respectively. However, the cell line clearly shows four subpopulations presenting distinct phenotype for CD20 and CD38, an uncommon feature for an immortalized cell line. (Figure 5.1a). Through a histogram chart representation of CD makers expression on U2932 cell line, it is possible identify double peaks for the CD20 and CD38 markers (Figure 5.1b). This data supports the identification of subpopulations with different immune phenotype expression on the same cell line. The cell line does not present a double peak expression for CD19, even though the cells are distributed between 0 to $1e10^4$ considering that the cut off for the high expression of CD19 is $1e10^3$. In this sense, the CD19 expression is an indicative of a differentiation or internalization process. The two main U2932 subpopulations, P4 and P5, demonstrate a reverse Abs expression characterized by its distribution on opposing extremities of the CD38/CD20 expression chart. At day 0, P4 presents a combined high expression for CD38 and CD20, whereas P5 show a combined low expression of CD38 and CD20. On the other hand, the two smaller U2932 subpopulations, P3 and P6, present predominant low expression for CD20 and high expression for CD38. The phenotype expression of each subpopulation is summarized at the table 1.

5.1.1. Immunophenotype of the sorted subpopulation for 36 days.

The U2932 cell line is sorted by FACS, following its subpopulation specific CD markers expression for gating the original cell line. The four subpopulations P3, P4, P5 and P6 correspond respectively to 5.79%, 18%, 49.5% and 10.8% of the original U2932 (Figure 5.1.1a). Immediately after sorting, re-analyzes of the aliquots of the sorted cells is developed to verify the purity of the sorting procedure (Figure 5.1.1b). All subpopulations demonstrate high level of purity (99% purity), presenting only a minor fraction of the sorted samples expressing CD markers not corresponding to the subpopulation immune phenotypic features

previously identified on the original cell line. The only exception is the P6 which present 93% of purity. Subsequently, frequently during the period of 36 days, aliquots of the sorted subpopulations expanded under the same conditions as the sorted subpopulations are submitted to flow cytometry analyses, evaluating CD marker expression level of stability compared to the previously demonstrated on the original cell line.

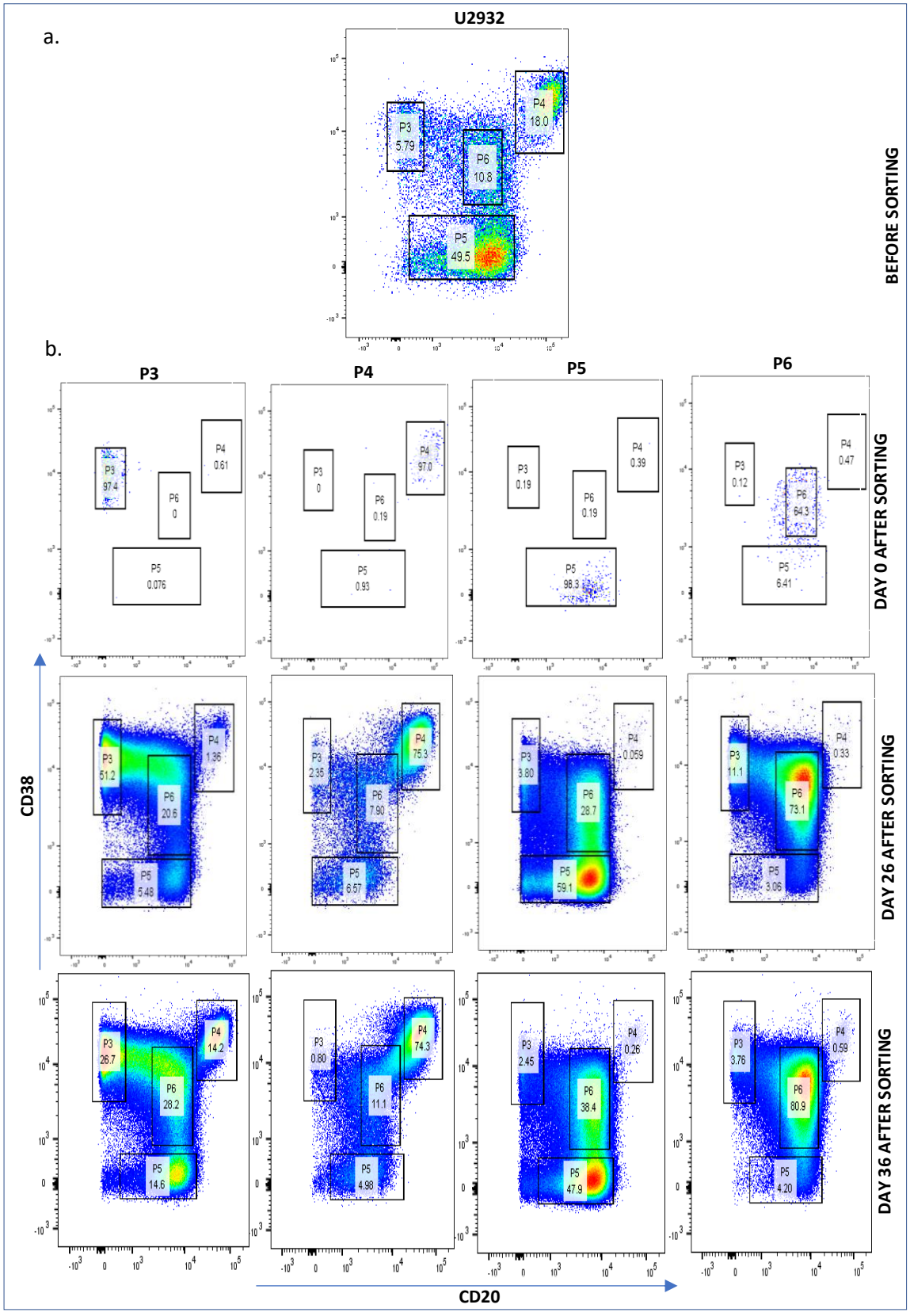

***Figure 5.1.1**: Florescent activated cell sorting of DLBCL U2932 cell line and immunophenotype analyses of its subpopulation for 36 days.* (a) Analyses of U2932 cell line immunophenotype by flow cytometry and determination of its subpopulation based on distinct CD38/CD20 marker expression. (b) P3, P4, P5 and P6 are the four U2932 subpopulations determined by phenotypically stable expression of CD20 and CD38 for 36 days. P3 and P6 are CD38 high/CD20 low, P4 is CD38 high/ CD20 high, P5 is CD38 low/CD20 low. After 35 days, the CD20/CD38 expression patterns of these four subpopulations maintained their phenotypic features clearly distinct from the unsorted cell line.

At day 26 and 36 after sorting, all subpopulations present immunophenotypic change compared to the subpopulation immune phenotypic features on day 0 after sorting. In detail, the major percentage of P4 (75.3% and 74.3% on day 26 and 36, respectively), P5 (59.1% and 47.9% on day 26 and 36, respectively) and P6 (73.1% and 80.9% on day 26 and 36, respectively) maintain their original CD marker expression for CD20 and CD38. Although a percentage of P5 and P6 present CD markers expression similar to P3, P5 and P6, these populations does not demonstrate more than 1% CD marker expression similar to P4, not presenting high CD20 expression. For P5, a very significant percentage of this subpopulation (28.7% and 38.4% on day 26 and 36 respectively) present CD marker expression resembling P6, indicating a change in the direction of the P6 phenotype expression. P3 is the subpopulation which present less concise CD maker expression. 57% of the P3 demonstrate an immune phenotype resembling the expression of other subpopulations studied (Table 1; appendix 2, table 1). Following the P3 expression chart from day 0 to 36, it is possible observe the change of P3 CD markers expression firstly presenting a similar immunophenotype to P6, and thereafter presenting similar expression to P4 and P5.

After 36 days of analyses, P3 and P6 are characterized by a low expression of CD20 and a high expression of CD38, P4 is characterized by high expression of CD38 and CD20, P5 is characterized by a low expression of CD38 and CD20. After 35 days, the CD20/CD38 expression patterns of these four subpopulations maintain their phenotypic features clearly distinct from the unsorted cell line. At day 36 of immune phenotype analyses, all subpopulations continue to express an immunophenotype distinct from the original cell line. The selected subpopulations stably express the sorting phenotype for a prolonged time, instead of regaining the original expression pattern of the unsorted cell line. However, a percentage of each subpopulation CD markers expression resemble to two or three of other subpopulations. In addition, the immunophenotype expression of a relevant percentage of P3 and P5 resemble to the P6 CD markers expression (Table 1; appendix 2, table 1).

5.1.2. Immunophenotype of the sorted subpopulation at day 100.

U2932 subpopulations and the unsorted cell line are submitted to flow cytometry analyses 93 days after sorting. The original cell line present distinct immunophenotype compared to the observed on its 36 days after sorting, demonstrating low expression of CD20 on 99% of its cells and high expression of CD38 on 79% of its cells. Surprisingly, the change on the immunophenotyping of U2932 cell line gain more emphasis after 100 days of analyzes, when the unsorted cell line present high expression for CD20 on 99% of its cells and high expression for CD38 on 81.7% of its cells (Figure 5.1.2A, Table 1, appendix 2, table 1). Thus, the unsorted U2932 demonstrate an emphatic drop on CD20 expression parallel to a crescent rise of CD38 expression, showing an immunophenotype change of the original cell line. In addition, the P6 represents now the major part of the cell line among the subpopulations composing the U2932 cell line, contrasting the previous immunophenotype analyses where the P5 is predominant.

At day 93 after sorting, P5 immunophenotype was clearly the most similar to unsorted cell line. For P5, the low expression of CD20 is predominant on 99.5% of cells, while 49% of its cells present high expression of CD38. At 100 days after sorting, the similarities among P5 and the unsorted cell line increase as CD20 low expression is maintained on 99.4% of P5 cells, while 51.7% of its cells expresses CD38 high. P3, P4 and P5 are still expressing CD20 high on 8.21%, 11.4% and 3.4% of their total population. However, CD20 high expression dropped to 5.22%, 9.66% and 2.14% for P3, P4 and P5, respectively. The same subpopulations present a rising of CD38 high expression from 93 to 100 days of analyses. The CD 38 high expression varied from 50.6%, 43.8% and 45.6% to 54.2%, 60.5% and 62% for P3, P4 and P6, respectively (Appendix 2, table 1).

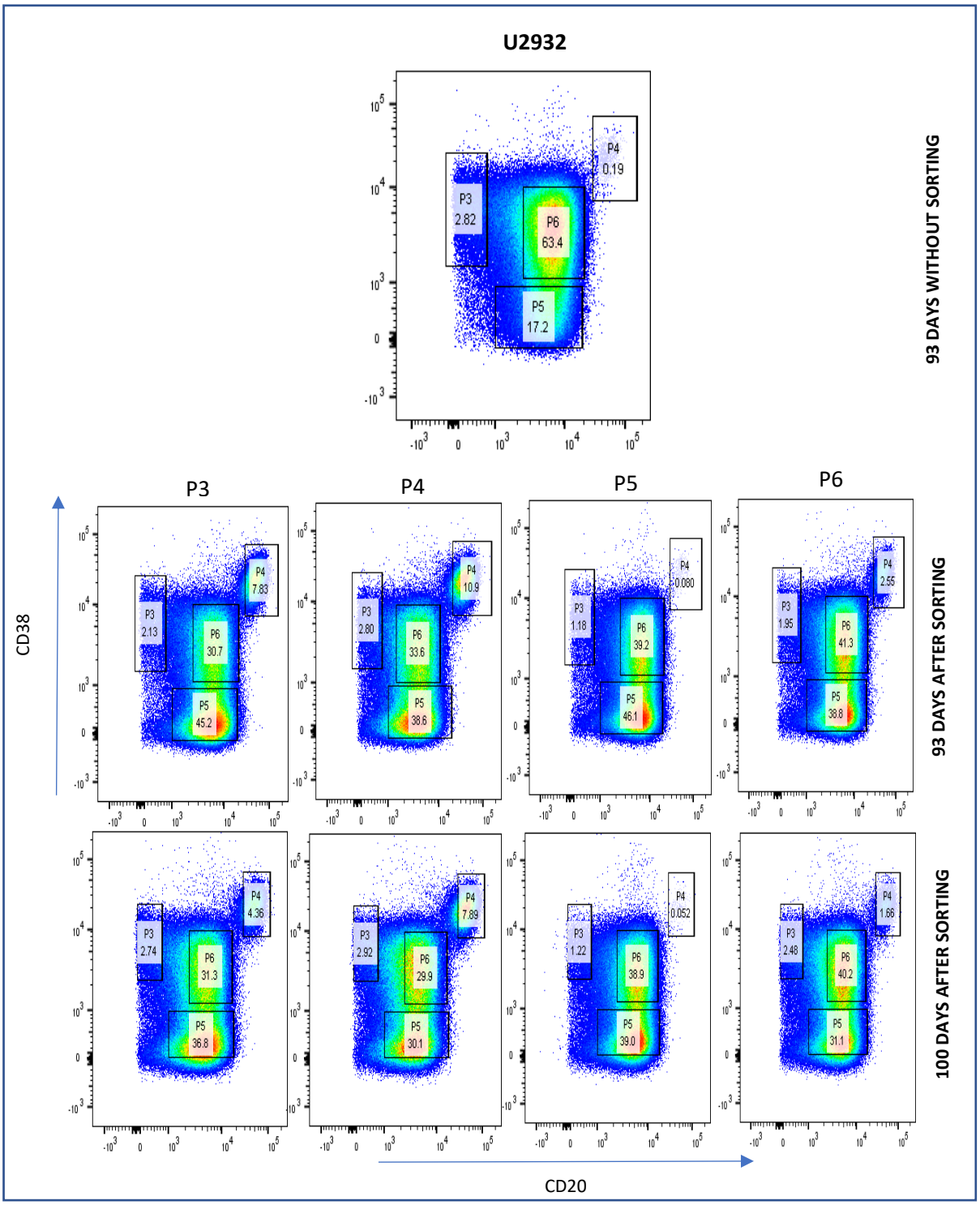

Figure 5.1.2A: *Florescent activated cell sorting of DLBCL cell line U2932 and immunophenotype analyses of its subpopulation on 93 and 100 days.* (a) Analyses of U2932 cell line immunophenotype by flow cytometry on day 93, based on distinct CD38 and CD20 markers expression. (b) U2932 subpopulations, P3, P4, P5 and P6, demonstrate immunophenotypically not steady expression of CD20 and CD38 at days 93 and 100. All subpopulations regain the original cell line CD markers expression, characterized by the common expression of CD20 low and CD38 high on majority of cells composing each subpopulation.

Surprisingly, at day 100 day of growth/expansion, all subpopulations demonstrate a not steady immunophenotype compared to the previous 36 days of analysis, regaining the unsorted U2932 cell line immunophenotype on the same analyses date (Figure 5.1.2A, Table 1, Appendix 2, table 1). The elevated reduction of cells number expressing CD20 high, which previously composed the P4, on the unsorted cell line indicates the death or the admission of a new phenotype by the cells composing this subpopulation. The P5 which have never expressed CD20 high or an immunophenotype resembling the P4 subpopulation maintained this characteristic. The unsorted cell line and all subpopulations expressed CD20 low and CD38 high, resembling P5 and P6 immunophenotype on day 36. Thus, P3, P5 and P6 demonstrate steadier phenotype compared to P4 during the 100 days of analyses. P3, P5 and P6 only demonstrate significant expression change for CD38 whereas P4 demonstrates expression change for CD38 and CD20. All subpopulations and the original cell line maintained positive expression for CD19 (Figure 5.1.2B). That data points to different stage of differentiation among the distinct subpopulations composing U2932.

Table 1: Immunophenotypic expression of U2932 and subpopulations at day 0, 36 and 100. ++ and -- indicates high or low expression above 80% or below 20%, respectively. +/- indicates high or low expression approximate to 50%.

Day 0	CD19	CD20	CD38
U2932	+	--	+/-
P3	+	--	++
P4	++	++	++
P5	+	--	--
P6	+	--	++
Day 36	**CD19**	**CD20**	**CD38**
U2932	+	--	+/-
P3	++	--	++
P4	+	++	++
P5	+	--	--
P6	+	--	++
Day 100	**CD19**	**CD20**	**CD38**
U2932	++	--	++
P3	+	--	+/-
P4	+	--	+
P5	+	--	+/-

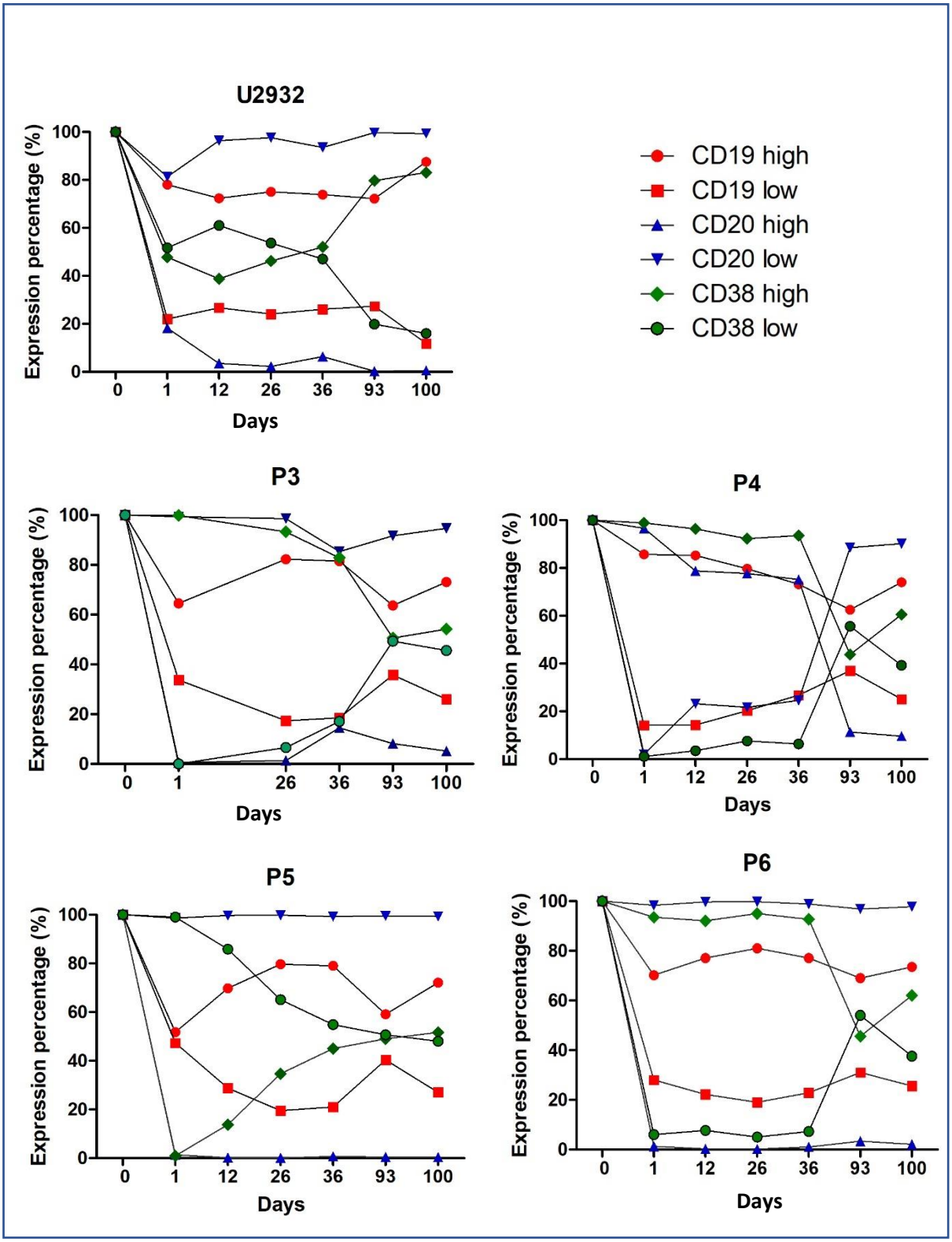

Figure 5.1.2B: *Percentage of cells presenting CD markers high and low expression for each subpopulation and U2932 cell line per day of flow cytometry analyses.* The sum of each CD marker high and low expression corresponds approximately 100% of the cells composing each population. Data obtained from SSC-A/ CD marker analysis on FLOWJO. *P3 had not reached appropriated number of cells at day 12 and it was not submitted to flow cytometry analyses at that date.

5.2. Immunophenotype of original cell line for 100 days.

Due to the death and/or decrease of cells expressing P4 and P5 phenotype and simultaneous growth of cells presenting P6 phenotype, U2932 unsorted cell line demonstrates alteration on its CD markers expression in different moments of flow cytometry analysis (Figure 4.3). In this sense, it only presents clear similarities with all subpopulation on day 93 and 100, but already demonstrating similar CD marks expression with specific subpopulation in previous time point. From day 0 to 26, the unsorted cell line expresses CD38 and CD20 low, same phenotype of P5, major phenotype expressed by the cells composing U2932 unsorted. At day 36, CD38 expression level starts to rise and the unsorted cell acquires CD20 low and CD38 high expression, resembling P6 immunophenotype. Finally, all subpopulations present an immunophenotype resembling the unsorted cell line on day 100.

The unsorted cell line is frequently labeled with CD19, CD20 and CD38 markers throughout the 100 days of analysis. The SSC-A/CD plot allows the visualization of CD19 steady expression contrasting the variable expression of CD20 and CD38 by the cells composing the U2932 unsorted cell line (Figure 5.2.1). CD19 maintains an expression rate from 0 to 10^4 for the entire experiment, presenting a unique population characterized by one area of high and middle cell density distribution (red and orange colour correspond to areas of high cell density distribution, yellow is the middle range). For CD19, majority of cells present a constant expression of CD19 high, varying from 76.3% on day 0 to 87.5% on day 100 (Appendix 2, table 1).

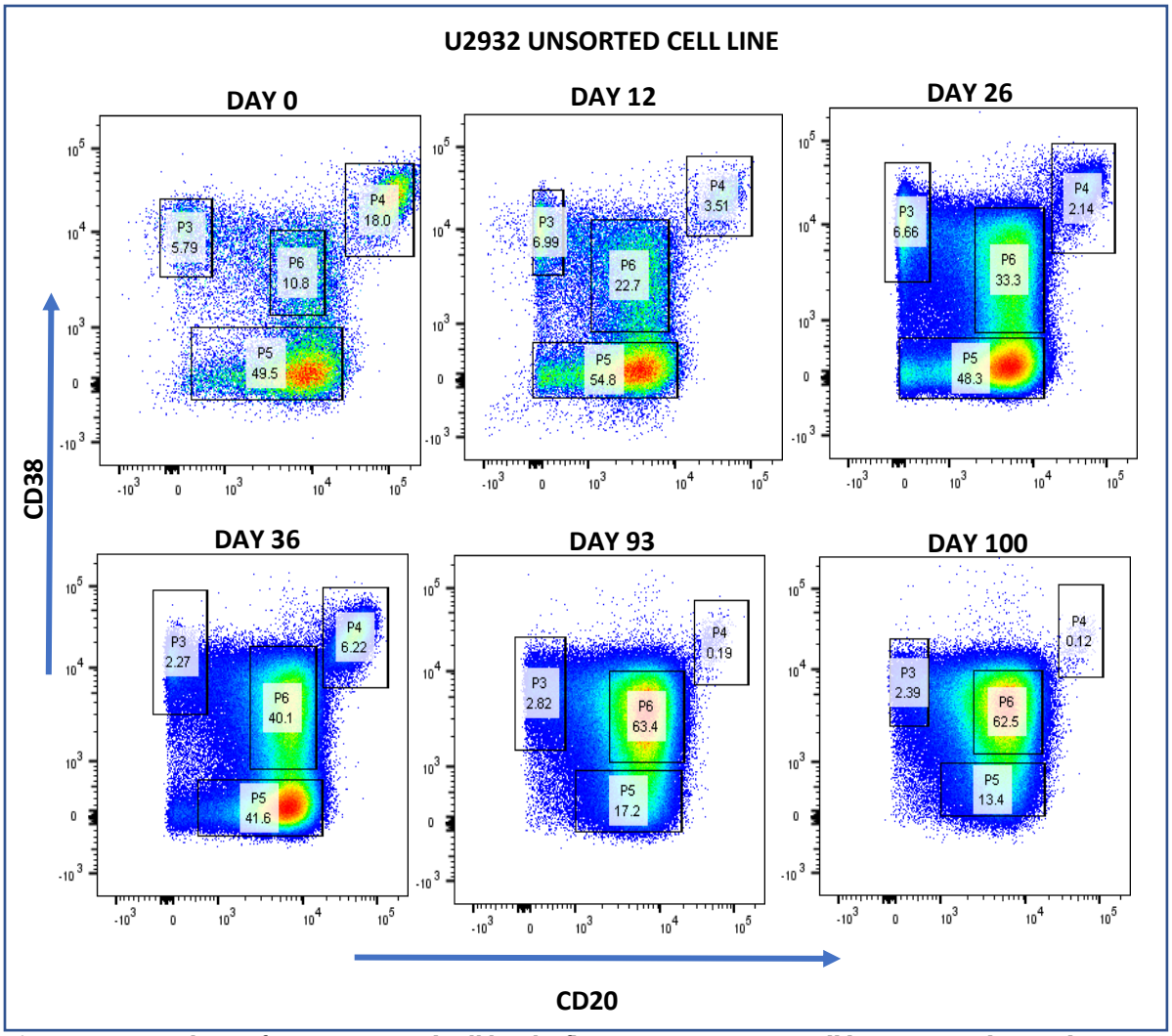

***Figure 5.2.1**: Analyses of U2932 unsorted cell line by flow cytometry.* U2932 cell line expressed CD38 low/CD20 low from day 0 until day 26. From day 36 until day 100, the cell line expresses CD38 high/CD20 low. The immunophenotype alteration might represent a reduction of P3 and P4 representation in the U2932 from day 12. In addition, the P6 representation on U2932 increases becoming the predominant subpopulation on this cell line.

At day 0, CD20 density expression plot presents an expression rate from 0 to 10^5, composed of one subpopulations with low CD20 expression, 81.3% of cells in total, and one subpopulation with high CD20 expression, 18.3% of cells. In addition, the cell expressing CD20 low appears to start to develop a third subpopulation at day 0, which significantly reduce at day 36 and completely disappeared at day 100. At day 36, the subpopulation expressing CD20 high is reduced to 6.4% and its area of high cell density became a low cell density characterized by the blue and green colours. At day 100, only one area of high cell density can be identified with an CD20 rate of expression from 0 to 10^4, featuring 99.5% of cells with an CD20 low expression. CD38 is composed of two distinct subpopulations with a more balanced distribution and presenting an expression rate from 0 to 10^5 on day 0 and 36. The subpopulation

expressing CD38 low varied from 51.7% on day 0 to 47.6% at day 36, while the subpopulation expressing CD38 high varied from 47.8% on day 0 to 52% at the same time point. In addition, the cell expressing CD38 high appears to start to develop a third subpopulation on day 36. At day 100, 83.1% of cells demonstrate CD38 high expression including the high- density cell area, the cells distribution and CD38 expression characterized only one population, with expression rate distribution reduced to 0 to 10^4 (Figure 5.2.2).

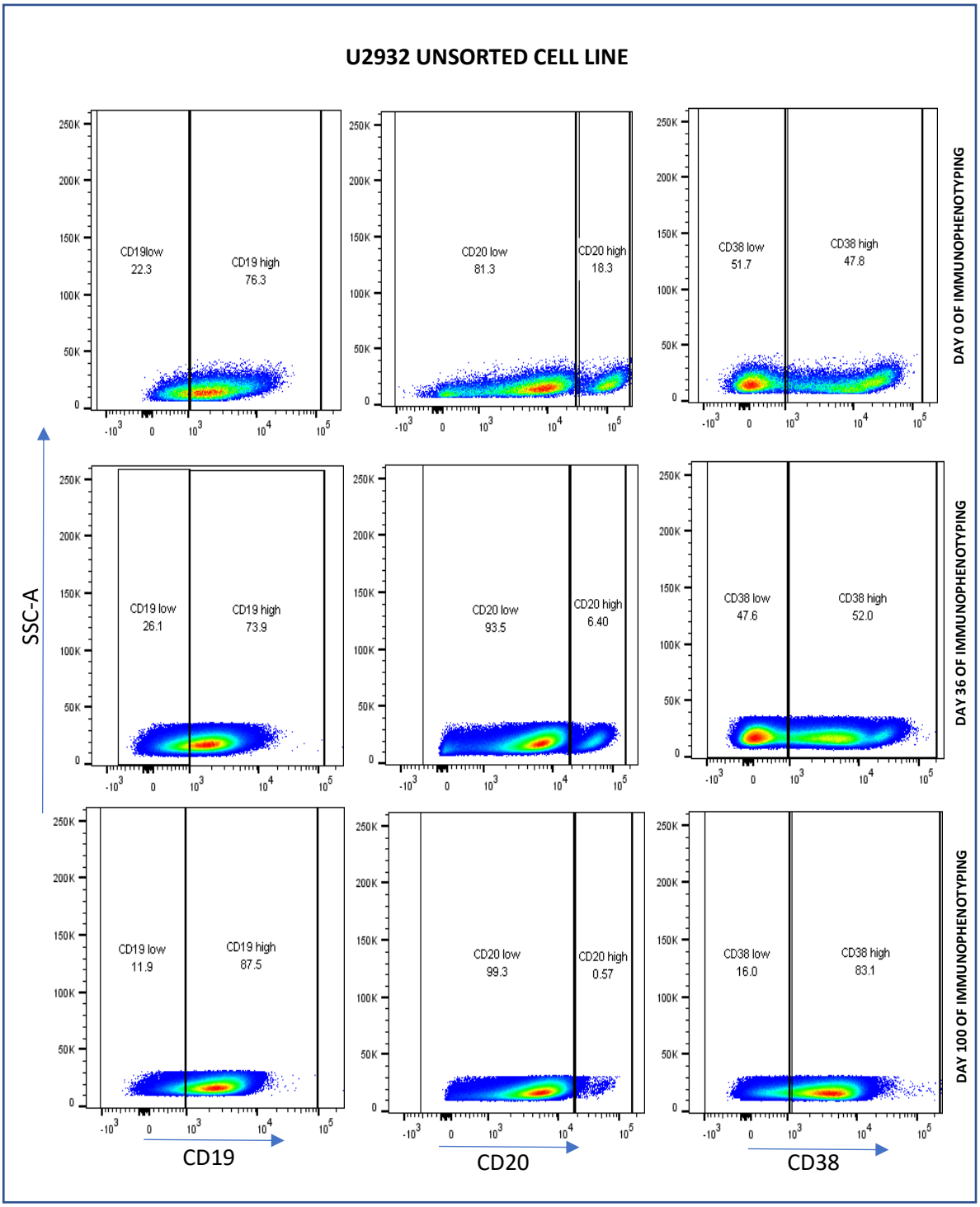

Figure 5.2.2: *U2932 unsorted cell line analysis through SSC-A/CD markers expression by flow cytometry*. The cell line is analyzed in three distinct time points, day 0, 36 and 100. The cells maintain a stable expression for CD19 with a unique population and a predominant CD19 high expression throughout the experiment. CD20 and CD38 presented two subpopulations on day 0 and three populations on day 36. Only on day 100, U2932 unsorted population cells distribution and CD20 and CD38 expression characterized only one population.

5.3. Major characterization U2932 cell line and its subpopulations at day 107.

The previous immunophenotype results demonstrate that not only U2932 subpopulations regain its original cell line expression, CD38 high /CD20 low, after 100 days of analysis, but also that unsorted U2932 cell line does not present a stable CD markers expression. U2932 cell line phenotype changed from CD38 low/ CD20 low to CD38 high/ CD20 low. To better characterize state of U2932 cell line and its subclones on the B cell development and differentiation process, all samples are submitted to the analyses of nine new markers on day 106 of analyses. For 8 of the markers analyzed, U2932 cell line and its subpopulation demonstrate a homogenous immunophenotype, characterized by the marker expression as one main population. U2932 and its subpopulation present high expression for lambda, CD45, CD10, IgM and low expression of kappa, CD27, CD5 and CD23 (Table 2). CD43 does not demonstrates and homogenous expression, presenting two high density areas each for P3, P4 and P6, however, the CD43 high expression was predominant in all cases (Figure 5.3, Table 2; Appendix 1, figure 3).

Table 2: Major immunophenotypic expression of U2932 and subpopulations at day 107. ++ and -- indicates high or low expression above 80% or below 20%, respectively. +/- indicates high or low expression approximate to 50%.

Day 107	U2932	P3	P4	P5	P6
CD19	++	+	+	+	+
CD20	--	--	--	--	--
CD38	++	+	+	+	+
CD5	--	--	--	--	--
CD10	+	+	+	+	+
CD23	--	--	--	--	--
CD27	--	--	--	--	--
IgM	++	++	++	++	++
CD45	++	++	++	++	++
CD43	++	+/-	+/-	+/-	+/-
Lambda	+/-	+/-	+/-	+/-	+/-
Kappa	+/-	+/-	+/-	+/-	+/-

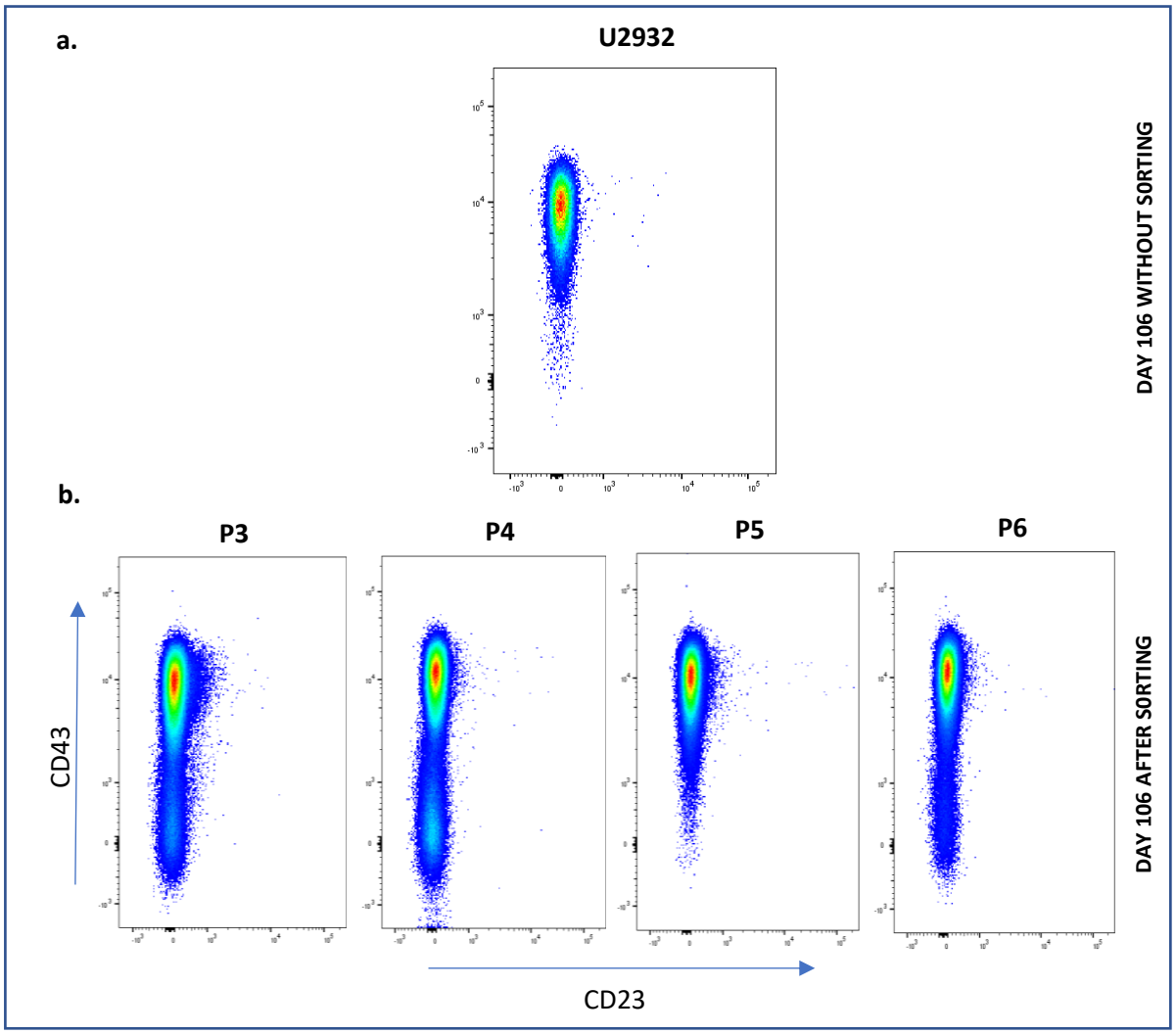

Figure 5.3: *U2932 cell line and subpopulations on day 106 of immunophenotype analyses by flow cytometry (a) Homogenous CD43 and CD23 expression for the unsorted U2932 cell line.* (b) CD43 and CD23 phenotypic expression by flow cytometry for the U2932 subpopulations. In all cases CD43 and CD23high expression is predominate, however, P3, P4 and P6 also expressed CD43 low.

5.4. Growth rate characterization of U2932 cell line and subpopulations

To characterize the growth pattern of U2932 cell line and its subpopulations, a growth curve for each of the subpopulations were generated over ten days (Figure 5.4). All samples present a lag phase from day 0 to day1. However, U2932, P3 and P5 reach the plateau at day 4 whereas P4 and P6 reached a plateau at day 5. To identify distinction on the growth pattern between the different cell lines, one- way ANOVA analyses with Tukey multiple comparison test (P value < 0.05) were separately performed per day, from day 1 to day 5. One- way ANOVA indicate significant difference present on day 1 (P value= 0,0005), day 2 (P value <0,0001) and day 3 (P value= 0,0004). Tukey multiple comparison test show significant difference between P4 and

the other cell lines on day 1, 2 and 3. At day 1, P4 demonstrate significant difference on growth pattern related to U2932 (P value = 0.0032), P3 (P value = 0.0044), P5 (P value = 0.0005) and P6 (P value = 0.0014). At day 2, P4 demonstrate significant difference on growth pattern related to U2932 (P value = <0.0001), P3 (P value <0.0001), P5 (P value = 0.0002). On day 3, P4 demonstrate significant difference on growth pattern related to U2932 (P value =0.0004), P3 (P value = 0.0012), P5 (P value = 0.003). P6 also show a different pattern of growth when compared to U2932, P3 and P5. On day 2, P6 demonstrate significant difference on growth pattern related to U2932 (P value =0.0001), P3 (P value = 0.0016), P5 (P value = 0.0047). At day 3, P6 demonstrate significant difference on growth pattern related to U2932 (P value =0.022). At day 4 and 5, no significant difference was identified among the growth pattern of U2932, P3, P5, P5 and P6. Thus, the statistics analyses show that P4 presents the slowest growth pattern among the cell line analyzed on the exponential phase of growth. Complete one-way ANOVA analysis with Tukey multiple comparison of the growth curves can be accessed on appendix 2, table 3.

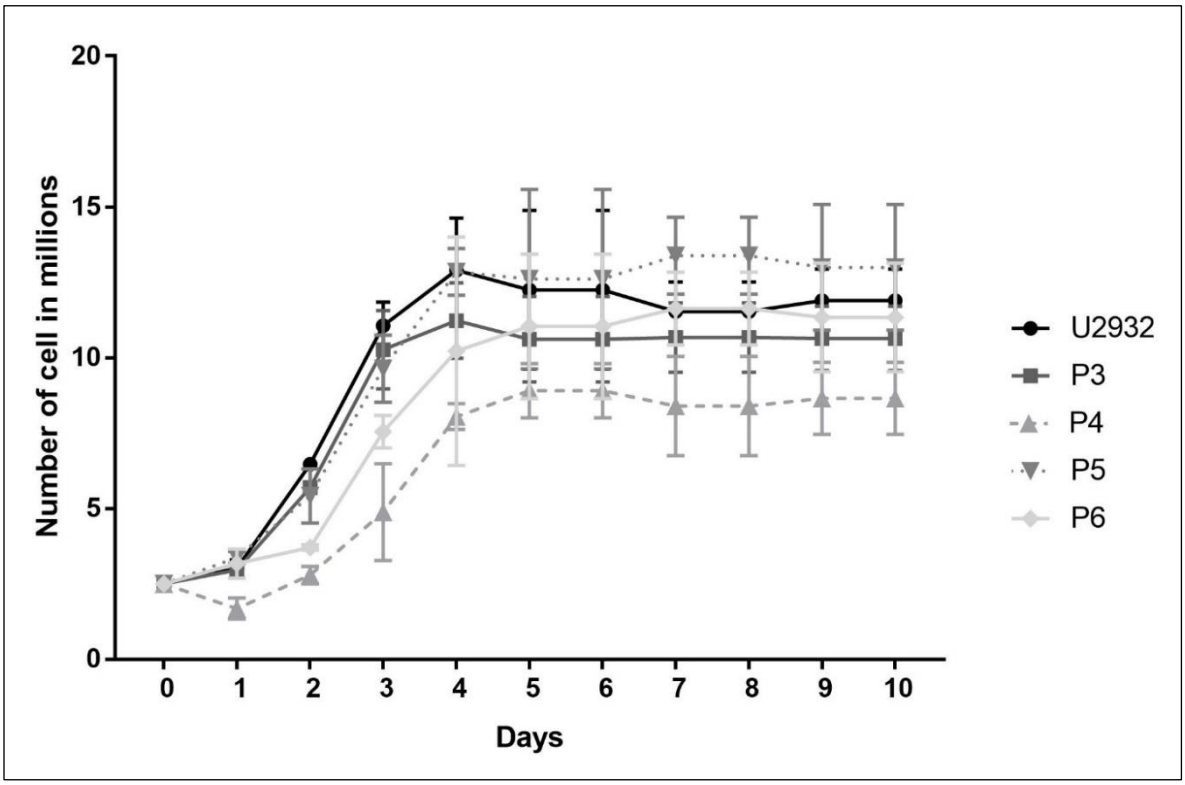

Figure 5.4: *Growth curve of U2932 cell line and its subpopulations for 10 days*. One-way ANOVA separately performed per each day of exponential phase demonstrate significant difference on growth pattern of P4 vs U2932/ P3/ P5 / P6 at day 1 of analysis and significant difference on growth pattern of P4 vs U2932/ P3/ P5 at day 2 and at day 3 of analysis. It also demonstrated significant difference on growth pattern of P6 vs U2932/ P3/ P5 at day 2 and significant difference on growth pattern of P6 vs U2932 on day 3 (P value > 0.5; Prisma GraphPad 7.0).

5.5. Morphologic characterization of the U2932 cell line and subpopulations.

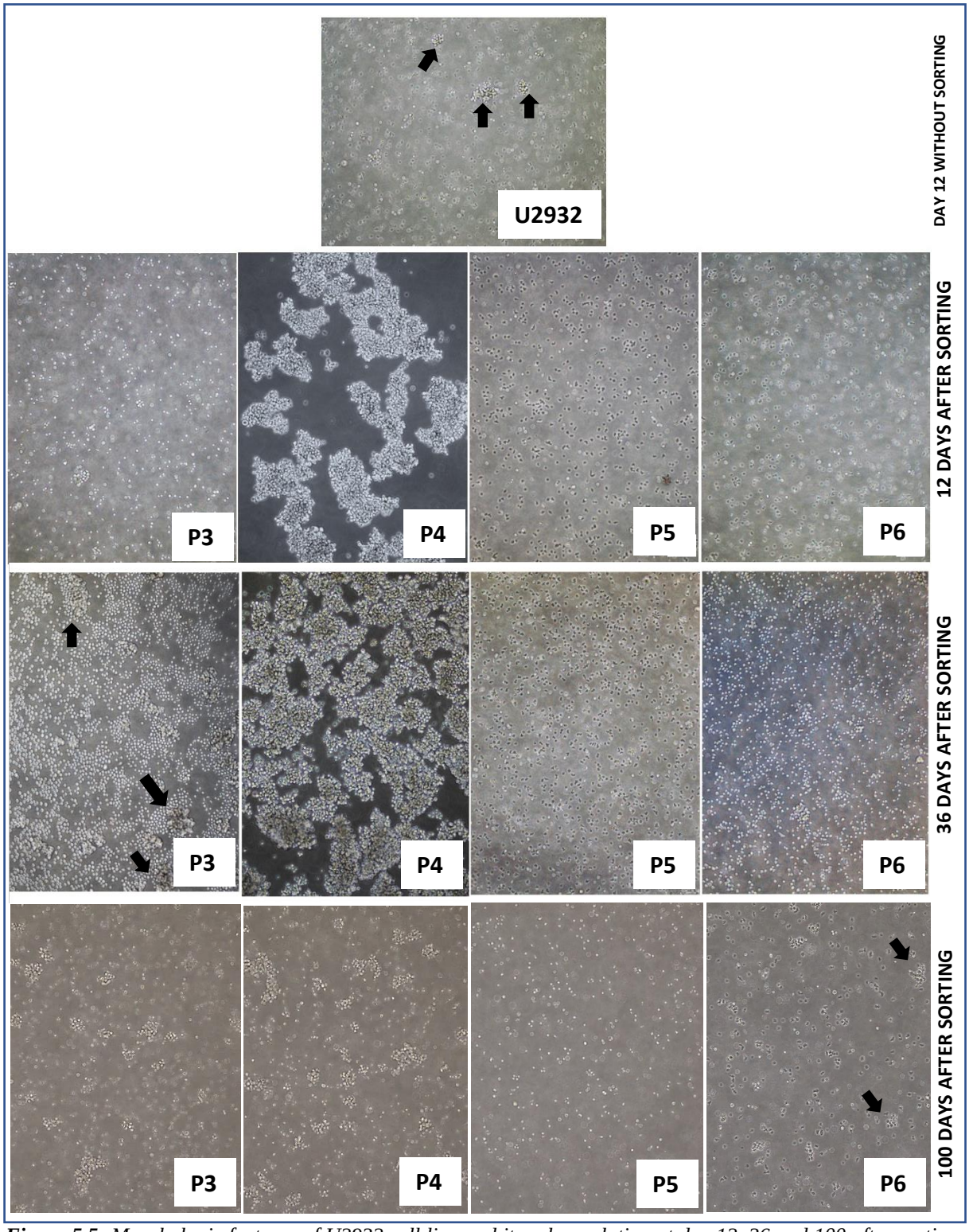

Figure 5.5*: Morphologic features of U2932 cell line and its subpopulation at day 12, 36 and 100 after sorting.* All cell lines present small round cells distributed individually on the cell line or in clusters. The presence of big cluster formations on P4 at day 12 and 36 indicate that this morphological disposition is related with the phenotypic characteristic of this subpopulation. Small cluster formations are indicated with arrows.

To provide a morphologic description of the U2932 cell line and its four distinct subpopulations, all cell lines are photographed at day 12, day 36 and day 100. At day 12, it is possible to identify small cells singly growing and small cluster formations on the U2932 cell line (Figure 5.5 - indicated by arrows), however, U2932 cells are mainly disposing spread in the media. On the same date, P3, P5 and P6 also showed small cells with spread distribution on media, but small clusters formation as observed on U2932, are not detected. Nevertheless, P4 demonstrates a cell disposal on big clusters clearly contrasting U2932, P3, P5 and P6 morphological features. The presence of small clusters on the original cell line is appropriate as this cell line is composed by three subpopulations presenting singly growing cells and one subpopulation growing in clusters. In this sense, it is possible to assume that the small cluster of U2932 corresponds to its P4 subpopulation.

At day 36, P4 cells presents even bigger clusters formation compared to day 12, emphasizing the distinct morphologic organization of this subpopulation compared to the others. On the same date, P3, P5 and P6 are still disposing singly growing cells, however, P3 presents small cluster formations of cells (Figure 5.5-indicated by arrows). Assuming that the small clusters corresponds to P4, the presence of small clusters on P3 support the previous flow cytometry analyses of P3 on day 36 where this subpopulation present 14.2% of its total cell population presenting an immunophenotype similar to P4 (CD20 and CD38 high expression). At day 100, P3 and P4 cells are growing as single cells and in clusters and P6 stats to present some few small cluster formations. This morphologic description is supported by flow cytometry analysis developed at day 100. At the same date, P3, P4 and P6 have 4.5%, 7.8% and 1.6% of the total cell for each subpopulation expressing CD20 and CD38 high. P5 does not show cluster formation during the date of visual analysis, a fact that also agrees with the flow cytometry analyses as this subpopulation only presenting values inferior to 1% its total cell expressing an immunophenotype resembling P4. Hence, CD20 high expression is related to cluster formation while CD20 low is related with singly cells growing.

5.6. Doxorubicin efficiency on U2932 cell line and subpopulations

Based on the distinct immunophenotype, growth pattern and morphological characteristic of the subpopulations and the unsorted cell line, drug screen analyses are performed to determine distinct sensibility to doxorubicin among the cell lines studied at day 36 and 100 after sorting. The cell lines are submitted to four distinct drug doses. The C2 corresponds to the GI50 value of doxorubicin to the U2932 cell line, C3 and C1 correspond to the double and half

concentration of C2, respectively. C0 corresponding only to isotonic saline. To determine static difference among the U2932 and its subpopulations submitted to the same doxorubicin dose, one-way ANOVA was developed for each C0, C1 C2 and C3 individually with Tukey multiple comparison as a post- test.

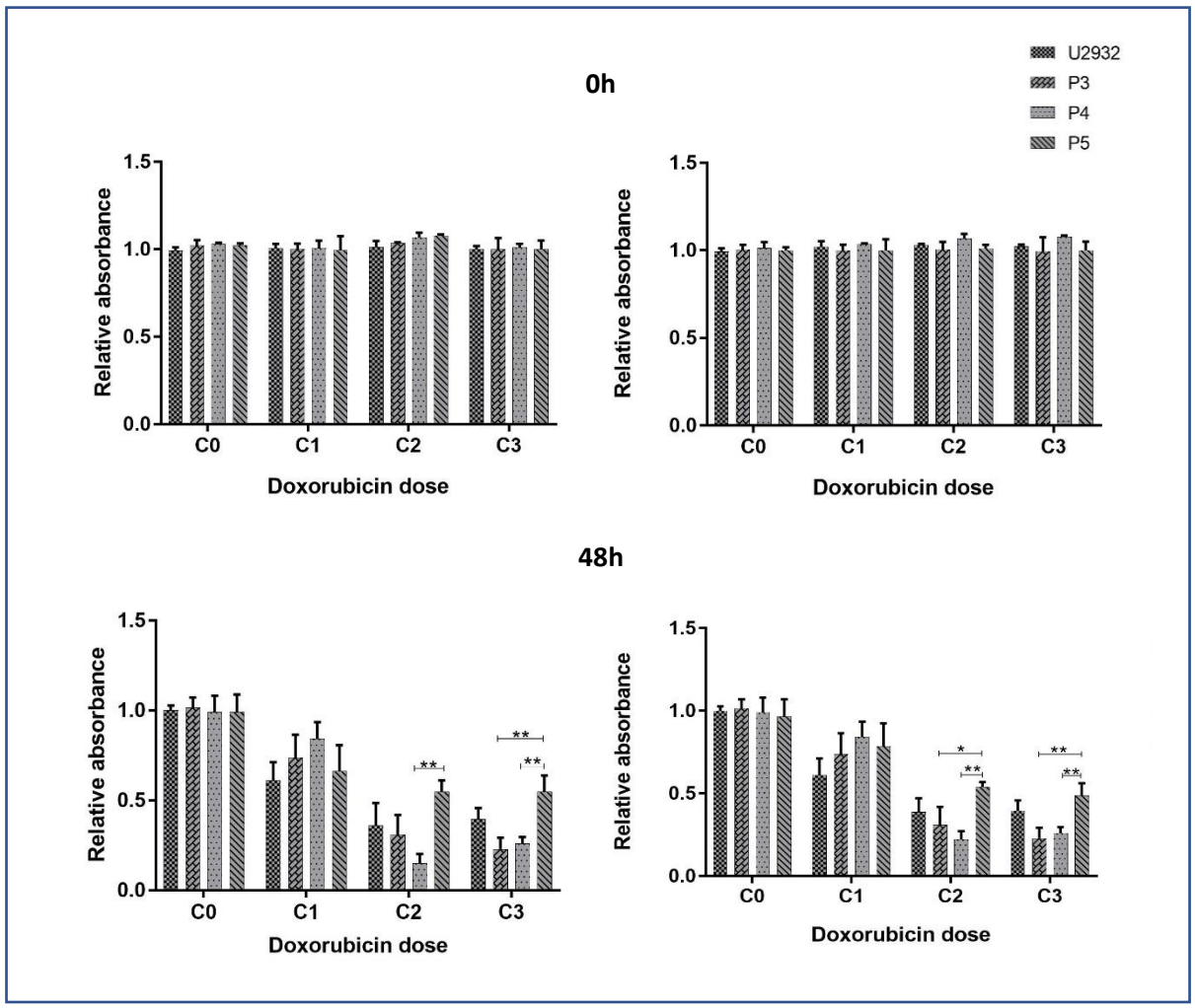

Figure 5.6.1*: Doxorubicin efficiency on U2932, P3, P4 and P5 on day 36 after sorting.* The cell lines are submitted to four distinct doses, C0 (isotonic saline only), C1 (GI25), C2 (GI50) and C3 (GI75). One-way ANOVA test indicates significant difference among the cell line results only on C2 (P value= 0,0046, 0,0053) and C3 (P value= 0,0026, 0,0012) doses after 48h exposition. Tukey multiple comparison test indicates significant difference at C2 for P4 vs P5 (P value=0,0035, 0,0034), P3 vs P5 (P value=0,0219, 0,0532). Tukey multiple comparison test indicates significant difference at C3 for P4 vs P5 (P value=0,0079, 0,0030) and between P3 vs P5 (P value=0,0036, 0,0015). (*P value= 0,032, **P value= 0,021; Significant P value <0.05, GraphPad Prisma 7.0).

After zero hour of exposition to the drug, none of the cell lines at 36 days after sorting demonstrate distinct sensibility to doxorubicin in any of the doses utilized. However, one-way ANOVA test presents significant values for both duplicates after 48h of exposition to the drug on C2 (P value= 0,0046, 0,0053) and C3 (P value= 0,0026, 0,0012). On multiple comparison test for C2, P4 demonstrates significant difference related to P5 in both duplicates (P value=

0,0035, 0,0034), P3 also presents significant difference related to P5 in one of the duplicates (P value= 0,0219, 0,0532). Multiple comparison test for C3, P4 demonstrates significant difference related to P5 (P value= 0,0079, 0,0030) and P3 also presented significant difference related to P5 (P value=0,0036, 0,0015) in both duplicates (Figure 5.6.1). Complete statistic results in appendix 2, table 3 and 4.

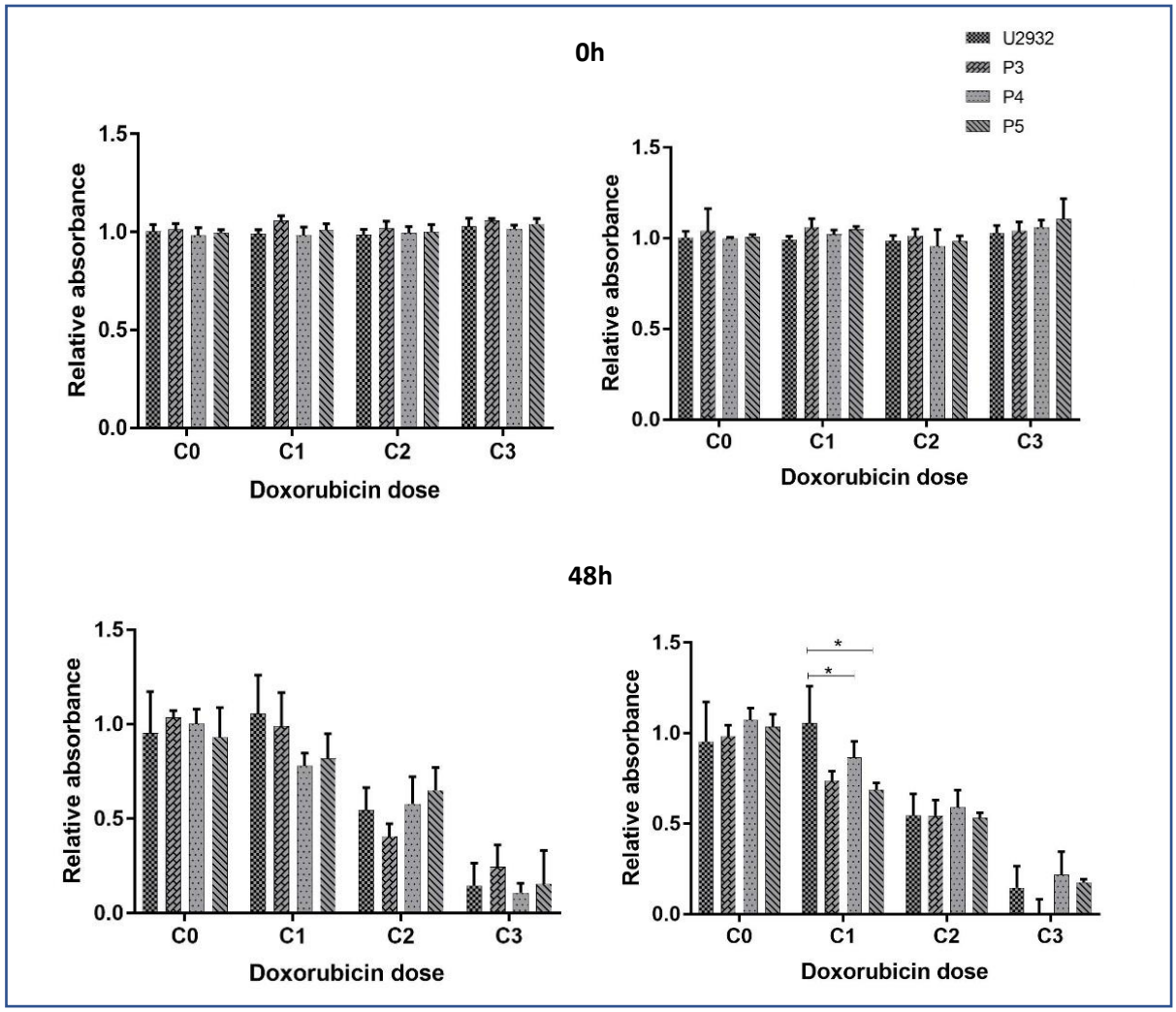

Figure 5.6.2: *Doxorubicin efficiency on U2932, P3, P4 and P5 after 100 days of sorting.* The cell lines were submitted to four distinct doses, C0 (isotonic saline only), C1 (GI25), C2 (GI50) and C3 (GI75). One-way ANOVA test indicates significant result among the cell line results only on C1 (P value= 0.0046, 0.0053) dose after 48h exposition. Tukey multiple comparison test indicates significant difference at C1 for U2932 vs P4 (P value= 0,2686, 0,2086) and U2932 vs P5 (P value= 0.0206, 0.3086). (*P value= 0.0332; Significant P value <0.05, GraphPad Prisma 7.0).

After zero hour of exposition to the drug, none of the cell lines at 100 days after sorting demonstrate distinct sensibility to doxorubicin in any of the doses utilized, and hence no significant difference is demonstrated. One-way ANOVA presented significant values for one of the duplicates after 48h of exposition to the drug on C1 (P value= 0,0203, 0,1638). On Tukey

multiple comparison test for C1, U2932 demonstrates significant difference related to P4 (P value= 0.02686, 0.2086) and P5 (P value= 0.0206, 0.3086) for one of the duplicate analyses (Figure 5.6.2). Complete statistic results in appendix 2, table 5 and 6.

The data presented indicates that U2932 and all its subpopulation after 36 days of sorting are not affected by doxorubicin on the same way. P3 and P4 present higher sensibility to C2 and C3 doses of doxorubicin than U2932 and P5. This data is in accordance with the distinct immunophenotype and morphologic features of U2932 subpopulations at this point. At 36 days after sorting, P5 and U2932 unsorted cell line present low CD20 expression and similar percentage of cells (54.9% for P5 and 47% for U2932) expressing CD38 low. Meanwhile P3 presents CD20 low, CD38 high expression and P4 presents CD20, CD38 high expression. U2932 and P5 are main morphologically characterized as small and round cells singly growing on media, while P3 and P4 present small and big cluster formation, respectively. However, the cells demonstrate the same sensibility to doxorubicin when exposed to the same drug dose after 100 days of sorting. This data is also in accordance with previous results as the subpopulations regained the U2932 immunophenotype and morphologic features, low CD20 and high CD38 expression, parallel to singly and cluster cell distribution on media. The only significant difference demonstrated by multiple comparison test were between U2932 compared to P4 and P5 after 100 days of sorting. Nevertheless, U2932 samples presents a high deviation which may have affected the final statistical analyses. To determine statistic difference among different drug doses for the same cell population, one-way ANOVA was developed for U2932, P3, P4 and P5 individually with Tukey multiple comparison as post- test for the samples after 36 and 100 days from sorting. In all cases, the cell lines demonstrate higher sensibility to C2 and C3 dose of doxorubicin.

5.7. Gene expression of U2932 cell line and subpopulations

Microarray gene expression analysis are established to substantiate the phenotypic characterization of U2932, P4 and P5. The different immunophenotype among U2932 subpopulations indicates the presence of cells in distinct stages of differentiation coexisting in this cell line. For this reason, *BCL6, MYC and BCL2* expression were further analysed considering that their aberration block or compromise the B cell differentiation process (Figure 5.7.1). Moreover, the gene expression of the transcription factors involved in the B cell differentiation (*PAX5, IRF4, PRMD1/Blimp-1, XBP1*) are evaluate for U2932, P4 and P5 (Figure 4.8.2). One- way ANOVA indicated significant difference (P value < 0.05) of

expression among the groups for all six genes (P value *PAX5* = 0.0176, P value *IRF4* = 0.0337, P value *BCL6*= 0.002, P value BCL2 = 0.0017, P value *PRDM1/Blimp-1* = 0.0015, P value *MYC* <0.0001, P value *XBP1*=0.0108). The P4 expression of *MYC* and *BCL6* is clearly distinct to the other groups. For *MYC*, the P4 expression is reduced compared to U2932 and P5. On the other hand, *BCL6* is highly expressed for P4 compared to U2932 and P5. P5 presented a reduced expression of *BCL6* compared to U2932 and P4. Tukey multiple comparison post-test (P value <0.05) indicated significant difference on P4 compared to U2932 (P value < 0.001) and P5 (P value < 0.001) for *MYC*, and it indicated significant difference on U2932 compared to P4 (P value = 0.011) and P5 (P value = 0.0024) for *BCL6.* For *BCL2* expression only demonstrated lower differences in the group comparisons, the U2932 have a higher expression of the gene compared to P4 and P5 (P value = 0.0016 and P value =0.008, respectively), P4 and P5 also demonstrated distinct expression for this gene among each other (P value= 0.0001).

For both *PAX5* and *XBP1*, only a slightly lower expression is observed for P4. For *IRF4*, only a slightly lower expression is observed for U2932. For *PRMD1/Blimp-1*, lower expression is observed for P4. Tukey multiple comparison post- test (P value <0.05) indicated significant difference on P4 compared to U2932 (P value = 0.0298) and P5 (P value = 0.0187) for *PAX5*, on U2932 compared to P4 (P value = 0.0415) and P5 (P value = 0.0467) for *IRF4.* The same test also indicated significant difference on U2932 compared to P4 (P value = 0.032) and P5 (P value = 0.032) and P4 compared to P5 (P value = 0.021) for *PRMD1/Blimp-1* and indicated significant difference on U2932 compared to P4 (P value = 0.032) and P4 compared to P5 (P value = 0.032) for *XBP1* (Figure 5.7.2). More than one probe set indicated the expression of *PAX5, PRMD1/Blimp-1* and *BCL6*. Detailed statistics analyses and probe set information on appendix 2, table 7 and 8.

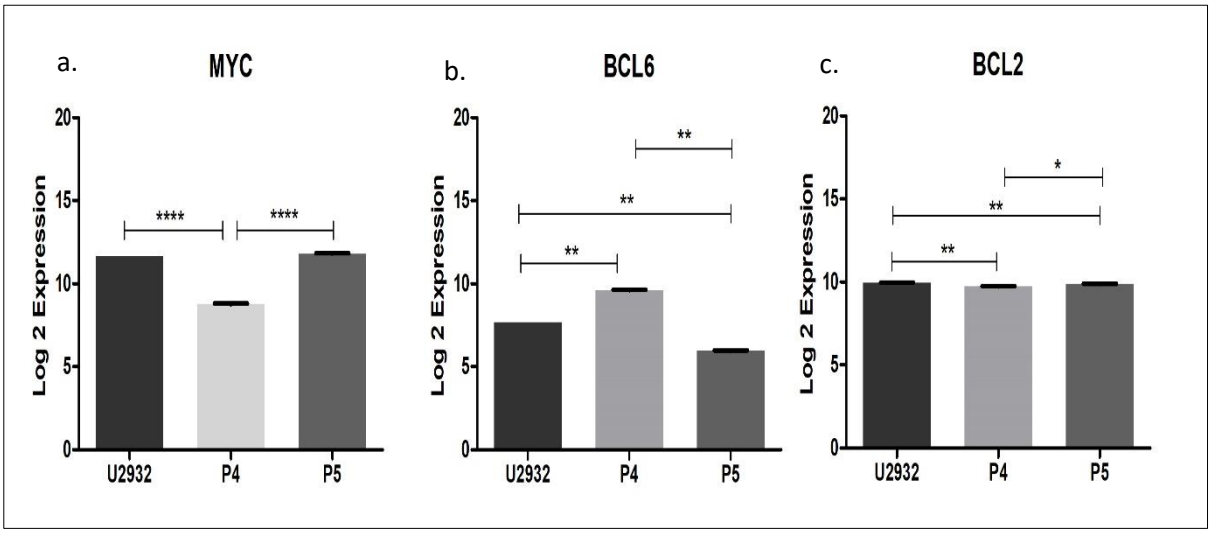

Figure 5.7.1: *Expression of MYC, BCL6and BCL2 among U2932 and subpopulations*. One-way ANOVA test indicates significant difference between the groups analysed for BCL6 (P value = 0.002), MYC (P value <0.0001). (a) Tukey multiple comparison test indicated significant difference between U2932 and P4 (P value < 0.001), and P4 and P5 (P value < 0.001) for *MYC* (b) Tukey multiple comparison test indicated significant difference between U2932 and P4 (P value = 0.011), and P4 and P5 (P value = 0.0024) for BCL6. (c) Tukey multiple comparison test indicates significant difference between U2932 and P4 (P value = 0.0016), and U2932 and P5 (P value =0.008) and P4 and P5 (P value= 0.0001) for BCL2. (*P value= 0.032, **P value= 0.021; ****P value< 0.0001, Significant P value <0.05, GraphPad Prisma 7.0).

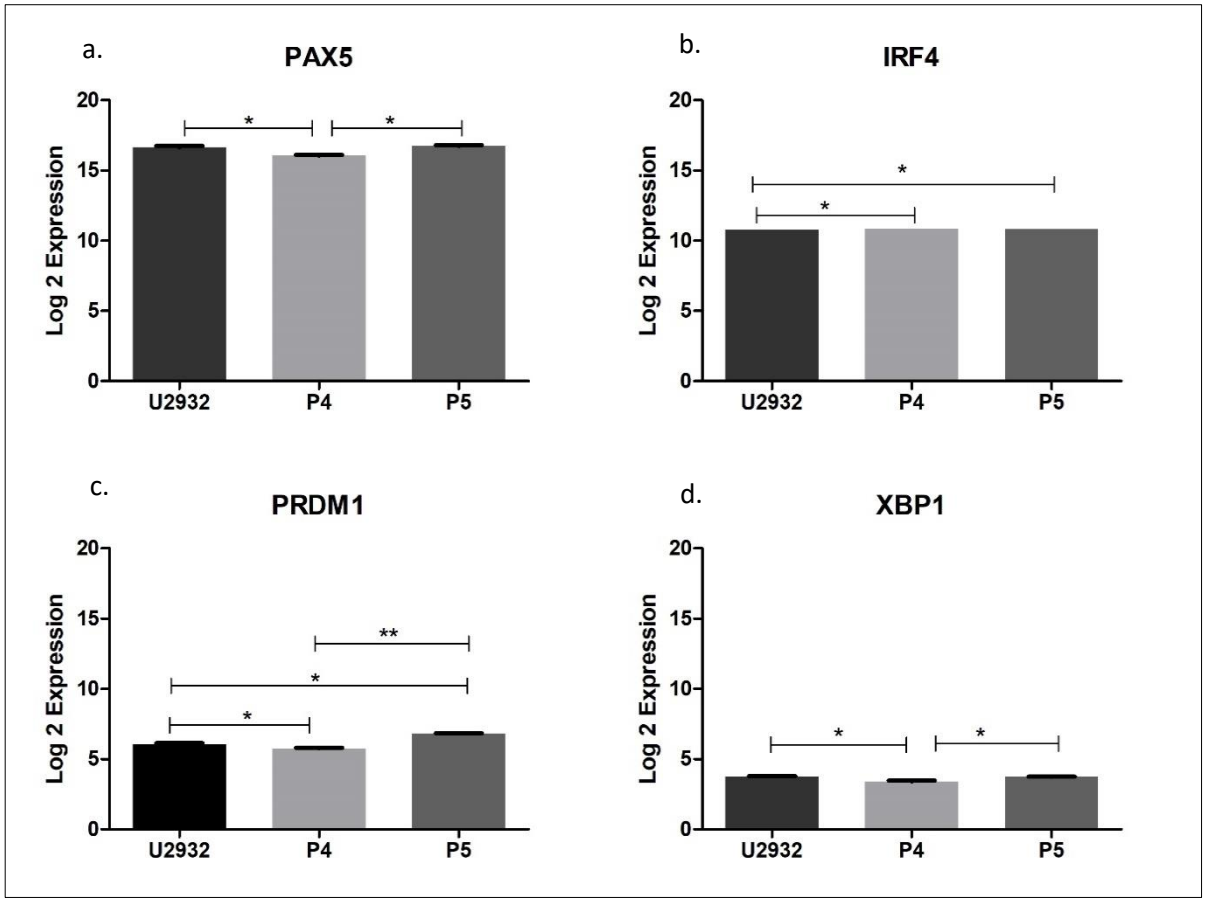

Figure 5.7.2: Expression of transcription factors involved in the B-cell differentiation among U2932 and subpopulations. One-way ANOVA test indicates significant difference between the groups analysed for *PAX5* (P value = 0.0176), *IRF4* (P value = 0.0337), P value *PRDM1/Blimp-1* = 0.0015, P value *XBP1*=0.0108. (a) Tukey multiple comparison test indicates significant difference between U2932 and P4 (P value = 0.0298), and P4 and P5 (P value = 0.0187) for *PAX5*. (b) Tukey multiple comparison test indicated significant difference between U2932 and P4 (P value = 0.0415), and U2932 and P5 (P value = 0.0467) for *IRF4*. (c) Tukey multiple comparison test indicated significant difference between U2932 and P4 (P value = 0.032), U2932 and P5 (P value = 0.032), P4 and P5 (P value = 0.021) for *PRMD1/Blimp-1*. (d) Tukey multiple comparison test indicated significant difference between U2932 and P4 (P value = 0.032), P4 and P5 (P value = 0.032) for *XBP1*(*P value= 0.032, **P value= 0.021; Significant P value <0.05, GraphPad Prisma 7.0).

5.8. Differentially expressed genes among U2932 cell line and its subpopulations and pathway determination.

The expression level analyses included 47,000 transcripts and variants, including 38,500 well-characterized human genes, consequently GEP included genes with low significance and fold change. Due to the high number of genes differently expressed and to guarantee an elevated significance and fold change of the genes composing the pathway enrichment analysis, a specified criteria of fold change > 2 OR change < -2 and p value with FDR < 0.04 were established. In this sense, the number of differently expressed genes between U2932 and P4 was reduced to 712 genes, and the number of differently expressed genes between P4 and P5 was reduced to 2.744 genes (Appendix 1, figure 4). No gene passed the specified criteria for the differently expressed genes analysis between U2932 and P5.

Table 3: Pathways enrichment for U2932 vs P4 by KEEGG

Pathway name	Genes in list/ Total no. of genes in pathway	Enrichment score
Rheumatoid arthritis	14/91	11.2548
Apoptosis	15/137	7.95095
MAPK signaling pathway	24/181	7.33434
Transcriptional misregulation in cancer	17/294	7.03758
NF-kappa B signaling pathway	11/93	6.78102

Utilizing the Fisher test, the pathway enrichments analysis is determined for the genes differently expressed among the original cell line and its subpopulations. For U2932 and P4, 10 pathways are identified containing more than one gene differently expressed also composing the pathway (Appendix 2, table 9). The five pathways presenting the genes with higher enrichment score are selected for further analysis (Table 3). 20 genes presenting differential expression are present in more than one pathway. TNF (fold of change -3.0, U2932 down vs P4) and *GADD45A* (fold of change 2.0, U2932 up vs P4) are presenting in four pathways. TNF compose the rheumatoid arthritis, apoptosis, MAPK signaling and NF-kappa B signaling pathway. *GADD45A* compose the apoptosis, MAPK signaling, Transcriptional misregulation in cancer and NF-kappa B signaling pathway. *MYC* (fold of change 7.5, U2932 up vs P4) is present in two pathways, MAPK signaling and Transcriptional misregulation. *BCL2A1* (fold of change -4, U2932 up vs P4) compose apoptosis, transcriptional misregulation in cancer and NF-kappa B signaling pathway. *BCL6* (fold of change -6.0 and -4.5, U2932 down vs P4) are

present only in transcriptional misregulation in cancer pathway. Complete list of genes composing the pathways on appendix 2, table 11- 15.

Table 4: Pathways enrichment for U2932 vs P5 by KEEGG

Pathway name	Genes in list/ Total no. of genes in pathway	Enrichment score
Epstein-Barr virus infection	42/199	9.77619
Rheumatoid arthritis	22/91	7.62471
HIF-1 signaling pathway	23/98	7.45854
B cell receptor signaling pathway	18/71	7.06229
Systemic lupus erythematosus	26/101	6.82448

For U2932 and P5, 10 pathways are identified containing more than one gene differently expressed also composing the pathway (Appendix 2, table10). The five pathways presenting the genes with higher enrichment score are selected for further analysis (Table 4). 21 genes presenting differential expression are present in more than one pathway. *IGH* (fold of change 3, P4 up vs P5) compose Epstein-Barr virus (EBV) infection, rheumatoid arthritis (RA), B cell receptor signaling pathway and systemic lupus erythematosus (SLE) pathway. *MYC* (fold of change -8.0, P4 down vs P5) were present only in Epstein-Barr virus infection. Complete list of genes composing the pathways on appendix 2, table 16- 20.

5.9. U2932 cell line and subpopulations classification.

For the classification of the samples into ABC/GCB, B-cell Associated Gene Signature (BAGS) and resistance gene signatures (REGS) for doxorubicin, it is utilized the Online One-By-One Microarray Normalization and Classification of Hematological Cancers for Precision Medicine (Hemaclass.org). U2932, P4 and P5 duplicates are normalized in a cohort- based RMA utilizing three distinct cell line groups, DLBCL cell line only, DLBCL and multiple myeloma (MM) cell lines and cancer cell lines (Table 5).

The BAGS classification varies according to the group utilized for normalization. U2932 and P5 are classified as Plasmablast and P4 is classified as plasmablast and centrocyte when normalized only with DLBCL cell lines. U2932 and P4 are classified as centrocyte when normalized with DLBCL + MM cell line and with cancer cell lines. P5 are classified as centroblast and centrocyte/centroblast when normalized with DLBCL + MM cell line and when

with cancer cell lines, respectively. In all cases, U2932 and P5 are classified as ABC presenting a probability value > 0.9. P4 is classified as no classified in all normalization cases, however, it always presented a probability value < 0.45 indicating a proximity to the GCB classification. For REGS, all samples are classified as doxorubicin resistant with a probability value > 0.9.

Table 5: U2932 and subpopulations classification by Hemaclass.

Cohort	U2932 samples	BAGS Probability	BAGS Classification	ABC/GCB Probability	ABC/GCB Classification	Doxorubicin Probability	Doxorubicin Classification
17 DLBCL cell lines	U2932	0.66420	Plasmablast	0.93737	ABC	0.96165	Resistant
	U2932	0.59596	Plasmablast	0.9371	ABC	0.95971	Resistant
6 U2932 samples	P4	0.78063	Plasmablast	0.31207	NC	0.9729	Resistant
	P4	0.55094	Centrocyte	0.37756	NC	0.97793	Resistant
	P5	0.80533	Plasmablast	0.9948	ABC	0.91526	Resistant
	P5	0.82175	Plasmablast	0.99267	ABC	0.91048	Resistant

Cohort	U2932 samples	BAGS Probability	BAGS Classification	ABC/GCB Probability	ABC/GCB Classification	Doxorubicin Probability	Doxorubicin Classification
17 DLBCL cell line	U2932	0.679247	Centrocyte	0.951282	ABC	0.992045	Resistant
	U2932	0.834284	Centrocyte	0.949958	ABC	0.99111	Resistant
17 MM cell line	P4	0.88271	Centrocyte	0.303703	NC	0.992915	Resistant
6 U2932 cell line	P4	0.976874	Centrocyte	0.437018	NC	0.993192	Resistant
	P5	0.425046	Centroblast	0.995992	ABC	0.975156	Resistant
	P5	0.634693	Centroblast	0.994776	ABC	0.977037	Resistant

Cohort	U2932 samples	BAGS Probability	BAGS Classification	ABC/GCB Probability	ABC/GCB Classification	Doxorubicin Probability	Doxorubicin Classification
41 cancer cell line	U2932	0.722347	Centrocyte	0.923478	ABC	0.994283	Resistant
	U2932	0.859973	Centrocyte	0.92305	ABC	0.99367	Resistant
6 U2932 cell line	P4	0.876138	Centrocyte	0.187754	NC	0.99582	Resistant
	P4	0.975445	Centrocyte	0.279426	NC	0.995902	Resistant
	P5	0.450009	Centrocyte	0.993216	ABC	0.982961	Resistant
	P5	0.368949	Centroblast	0.989726	ABC	0.982482	Resistant

5.10. Genetic correspondence among U2932 and its subpopulations.

All subpopulations present identical alleles also demonstrated on original cell line by barcoding analyses. It eliminates the possibility of U2932 contamination with other cell line which would explain the co- existence of subpopulation in the same tumor. In addition, it ensures the quality of the biological material utilized in this experiment.

6. Discussion

The variety of genetic mutations [1], the diversified patients' prognostic and overall survival [36, 47], and drug resistance [58] characterizing DLBCL subtypes are clear indications of this tumor heterogeneity. Elevated heterogeneity and plasticity is a hallmark of malignant cells composing the tumor environment. Carcinogenic cells can develop or acquire genetic alterations that allow the cells transition into different biologic states favouring the tumor growth, establishment of phenotypic equilibrium and drug resistance [101, 102]. Additionally, it has been recently identified the presence of different subpopulations in U2932, a DLBCL cell line [108]. Hence, this study hypophyses that different subpopulation composing U2932 correspond to B cells at different states of differentiation according to their phenotype and transcriptional profile, consequently demonstrating distinct drug sensibility. P4 and P5 demonstrate distinct CD markers expression for CD38/CD20, different gene expression for *BCL6*, *MYC* and *PRDM1*/Blimp-1 and distinct pathways between U2932 vs P4 and P4 vs P5 enrichment in microarray analysis at 36 days after sorting. P4 also shows higher sensibility to doxorubicin compared to the other subpopulations and hemaclass classification of U2932, P4 and P5 into BAGS demonstrated variation according to the normalization cohort. However, all U2932 subpopulations demonstrated the same drug sensibility and phenotype for CD10, CD19, CD38, CD45, IgM, lambda, CD5, CD20, CD23, CD27, kappa and CD43 at 100 days after sorting.

6.1. U2932 subpopulations present distinct stage of differentiation at day 36.

Firstly, it is necessary emphasizes that U2932 cell line does not correspond to cells in a homeostatic state. In this sense, this DLBCL cell line might not demonstrate markers surface expression similar to non-malignant B cells or other cancer types. In health adults, CD38 positive expression is a characteristic of immature, and plasma cells [76,77]; CD20 is high expressed throughout immature and mature B cell phases, extinguished at the terminal plasma cells differentiation [72,73]. In this sense, combined high CD30/CD20 expression indicates that P4 is on immature B cell state while P5 is composed of naïve or memory B cells [66, 112, 113]. These assumptions support the hemaclass BAGs classification for P4 and P5 as centroblast or centroblast/centrocyte respectively when normalized with 17DLBCL + 17MM+ 6U2932 cell line and 41Cancer cell line + 6 U2932. Moreover, MM plasma cells have a distinct immunophenotype compared to NHL based on CD markers expression and surface

immunoglobulin expression [114, 115]. In this sense, the normalization of U2932 and P5 samples with MM or other cell lines is inaccurate compared to the analyses only utilizing DLBCL. Thus, we assume that the hemaclass BAGS classification of U2932, P4 and P5 as plasmablast is the most appropriate considering that it utilized only DLBCL cell lines for normalization.

The combined low/absent and intermediate/high expression of CD38/CD20 has been demonstrated in aggressive plasma B cell lymphoma with plasma B cell differentiation cases, encompassing P4 and P5 CD markers phenotype. 35 cases of DLBCL demonstrated a plasma differentiation state and were analysed through immunostained tissue microarray. The exact plasmablastic phenotype was defined by *PRDM1*/*Blimp-1*-positive, *XBP1*s-positive and CD20/*PAX5*-negative or intermediate and commonly associated with positive cases of Epstein-Barr virus [116]. EBV is related with *PAX5* downregulation by promoting the hypermethylation of *PAX5* promoter and *XBP1* activation through stress mediated activation process [117]. In addition, a marked feature of the patient's phenotypes was the negative *BCL6* expression. A similar phenotype was observed on P5 presenting *PRDM1*/*Blimp-1*- positive, *XBP1*s-positive, CD38/CD20 -negative, *PAX5*-high, *BCL6*- intermediate, *IRF4*- positive and EBV as its main pathway enrichment [116]. This speculation about P5 cells state of differentiation agrees with the hemaclass BAGS classification for this subpopulation as plasmablast, when only normalized with 17DLBCL + 6U2932, and the ABC classification for all cohorts utilized for normalization.

The identification of plasma cells phenotype heterogeneity is recurrent among rodents and humans, indicating distinct function among plasma cells originated from GC or germinal follicular plasma cells [118]. In our analyses, P4's phenotype, *PRDM1*/*Blimp-1*-low, *XBP1*s-low, CD38/CD20 -positive, *PAX5*-high, *BCL6*-high and *IRF4*-positive, diverge from the most common plasmablast B lymphoma. Low blimp-1 expression contrasts the plasmablast characteristics due to its function as fundamental regulator and hallmark of terminal plasma cell differentiation. Analyses utilizing green florescence protein as a reporter of blimp-1 locus on mouse strain for MM and plasmacytoma cell lines (plasma cell tumours in mice) demonstrated blimp-1 high or intermediate (5 fold lower) level of expression as fundamental indicator of plasma cell specificities and progeny. The plasma cells in the bone marrow and the scarce plasma cell in the blood demonstrated blimp-1 high and intermediate expression, respectively [14].

Flow cytometry analyses showed the positive expression of early stage of development markers, such as B220, CD19, CD22 and major histocompatibility complex II (MHCII), on blimp-1 intermediate plasma cells whereas blimp-1 high plasma cells presented negative expression for the same markers. In addition, proliferation analyses in vivo after bromodeoxyuridine (BrdU) incorporation in mouse were done utilizing anti-BrdU mAb on flow cytometry. The proliferation capacity of blimp-1 high plasma cell was lower than blimp-1 intermediate plasma cell which rapidly turned over. The higher incidence of blimp-1 intermediate plasma cells on blood and its elevated proliferation capacity after immunisation demonstrates that blimp-1 intermediate phenotype is a consistent indicator of extra follicular plasma cells. Besides, the early stage phenotype of blimp-1 intermediate cell suggest that these cells are on distinct stage of differentiation compared to blimp-1 high cells [14].

Moreover, *PRDM1* has demonstrated tumor suppressor function on DLBCL, especially on ABC subtype. The OCI-Ly3, an ABC DLBCL cell line, demonstrated clonal deleterious mutation of *PRDM1* by copy number determination using real- time quantitative PCR (RT Q-PCR). Supporting this data, 8 clinical DLBCL cases phenotypically characterizing as non- GC (*BCL6+*, CD10- *IRF4+*) demonstrated deleted or transcriptionally silence paired alleles for the mutant *PRDM1* alleles by sequencing and agarose gel analysis of cDNA obtained from RT-PCR. It consequently promotes the reduction of *PRDM1* transcription factors and protein products [119]. P4 and P5 presented blimp-1 positive but low expression possibly as a consequence of *PRDM1* tumor suppressor function on ABC DLBCL, however, blimp-1 expression was even more diminished for P4. In this sense, we speculate if P4 actually could correspond to extrafollicular plasma cell. This speculation agrees with the hemaclass BAGS classification for this subpopulation as plasmablast, when only normalized with 17DLBCL + 6U2932. In addition, the positive expression of early B cell markers for extrafollicular plasma cells combined with its BAGS classification as plasmablast might justify the ABC/GCB classification of P4 as non- classified (NC) for all cohorts utilized for normalization.

6.2. Subpopulations growth rate reflects genetic mutation of enrichment pathways.

In this study, we observed P4 to have lower proliferation capacity compared to the U2932 cell line, P3 and P5 for the first 72h of growth curve experiments. In parallel, pathway enrichment of U2932 vs P4 demonstrated altered expression on 24 and 11 genes composing the MAPK signaling and NF- kappa B signaling pathway, respectively. Both pathways are commonly activated by growth factors related to cancerogenic plasma cell proliferation. IFN- alfa and IL-6 independently trigger MAP kinase signaling by promoting STAT3 phosphorylation [120]. IL-6 cytokine is highly produced in bone marrow of MM patients and its blockage has demonstrated to effectively reduces the MM tumours proliferation [121]. However, IFN-alfa have presented inhibitor capacity on ANBL-6, a MM cell line. IFN-alfa failed to induce this cell line progression to S/G2 and M cell cycles phase. In parallel, IFN-alfa induced p19 expression on RNA and protein level 2h after stimulation indicating this gene function as early signal of G1 arrest, finally conducing the cells to apoptosis, which is also a pathway established by the U2932 vs P4 enrichment [122].

U2932 vs P4 enrichment demonstrated alteration on insulin receptor gene (*INSR)* expression for the MAP kinase pathway presenting a fold change of 7.0 (U2932 down vs P4 up). Insulin receptors substrate 1 is rapidly phosphorylated by IFN-alfa for the interleukin 4 activation [123]. Interleukin 4, a protein that regulates immunologic process, promoted inhibition of bone marrow macrophages maturation into osteoclasts by avoiding the receptor activator of nuclear factor kappa-B ligand (RANKL) signaling consequently inhibiting NF- kappa B pathway and MAP kinase pathway. RANKL compose a signaling super family presenting TNF- alfa as archetypical component [124]. For P4, TNF presented a fold change of -3.0 and -2.5 (U2932 down vs P4, two different probe set) for NF- kappa B and MAP kinase pathway, respectively. In this sense, we speculate that the elevated expression on *INSR* is a consequence of high IFN-alfa activity promoting interleukin 4 high expression, culminating into a diminished proliferation capacity of P4 compared to the other subpopulations and the original cell line through the inhibition of NF- kappa B and MAP kinase pathway. However, the described gene alterations only initially reduce P4 proliferation due to other growth factors, such as IL-6, capacity to activate proliferation pathways.

Moreover, P4 vs P5 enrichment demonstrated altered expression on 23 of 93 genes composing the hypoxia-Inducible Factor (HIF-1 pathway). Under oxygen restricted conditions, as

observed on solid tumor formation, HIF pathway is commonly activated consequently promoting the transcriptional activation of a variety of genes participating in mechanisms that favours the cells survival on that specific environment, including the cell proliferation [125]. HIF- alfa and *c-MYC* demonstrated synergic proliferation capacity promoting a significant increase of Burkitt's lymphoma (P493-6) tumor volume in a xenograft model compared to the control *MYC* inhibited. HIF-1 interaction with *c-MYC* promotes glycolysis inhibiting pyruvate dehydrogenase and reducing mitochondrial respiration [126]. P5 demonstrated high *MYC* expression which combined with HIF-1 pathway up-regulation justify this subpopulation initial higher proliferation compared to P4. High insulin receptor expression with approximately -13, -10 and -7.25 fold change, corresponding to 3 different probe sets, (P4 down vs P5 where) substantiate the HIF-1 pathway occurrence. Insulin- like growth factor have triggered hypoxia-inducible factor 1 through MAP kinase mediation on human colon carcinoma cell line (HCT16) [127]. This data reinforces the importance of MAP kinase pathway for malignant cells proliferation supporting the initial higher proliferation of P5 compared to P4.

6.3. P4 and P5 genetic mutation relation with autoimmune disturbances and/or infectious conditions, and phenotype similarities.

U2932 vs P4 and P4 vs P5 pathway enrichment demonstrated rheumatoid arthritis with 11.25 and 7.6 as enrichment score, respectively. Rheumatoid arthritis and NHL is a common combined diagnostic. From 42 rheumatoid arthritis patients 33 demonstrated NHL after paraffin- embedded tissue analyses and 67% of NHL corresponded to DLBCL [128]. Further characterization of DLBCL diagnostic correspondence demonstrated elevated rheumatoid incidence among non-GCB DLBCL. Among the 139 DLBCL combined with rheumatoid arthritis 70% of cases demonstrated ABC phenotype with *BCL6* negative and *IRF4* positive in 78% and 59% of RA+ DLBCL cases, emphasizing the correlation between those conditions and pointing to a potential participation of active peripheral B cells on this correlation [129]. Epstein- Bar virus infection and systematic lupus erythematosus pathway were also presented for the P4 vs P5 pathway enrichment with 9.7 and 6.8 as enrichment score, respectively. From 114 DLBCL cases, 11.4% presented EBV encoded RNA in a *in situ* hybridization analyses. Among the EBV+ DLBCL patients, 92%, 31% and 23% of cases presented CD30, *BCL6* and IgG expression, respectively. The elevated incidence of CD30 compared to the lower *BCL6* expression indicate that the majority of EBV+ DLBCL are non- GCB DLBCL also indicating participation of activated B cell for the combined diagnosis [130]. Systematic lupus erythematosus has demonstrated phenotype correlation with NHL as well. In a cohort of 40

NHL patients, 16 cases showed positive diagnostic for SLE. For the 10 cases presenting SLE+DLBCL, CD10 and BCL-6 expression was predominant indicating ABC DLBCL [131]. Combined SLE, EBV in DLBCL patients has been characterized with elevated aggressiveness and high lymphocytic proliferation similar features to P5 and ABC DLBCL [132]. Moreover, EBV have demonstrated down regulator and activator function over *PAX5* and *IRF4* which clearly contribute for B cell progression to terminal GC differentiation [117].

6.4. CD19 heterogenous phenotype is a consequence of U2932 cell differentiation and CD19 internalization process.

Quentmeier *et al.* 2013 demonstrated temporary simultaneous CD19 +/- expression in the U2932 cell line, which was considered a consequence of the transitional state of B cell differentiation cells composing this cell line [108]. Initial trial developed by our group presented CD19 ambiguous phenotype for 29 days on the same cell line, which was partially reproduced on this study. In this study, all subpopulation presented an heterogenous expression for CD19 with the cells distribution in between 0 to 10^4 in SSC-A plot, but the samples did not present distinct peaks of expression for CD19+/- in the histogram plot as observed in our initial trials. B cells in a T cell dependent or independent differentiation process have already demonstrated CD19 heterogenous expression. Plasma cell originated from bone marrow demonstrated increased CD19- phenotype with or without toll like receptors stimulation, efficiently mimicking the T cell activation, *in vitro* throughout PB to PC differentiation process. In addition, CD19- expression was proportional to the time of differentiation process and the long lasting plasma cell formation Although the increase of CD19- phenotype after samples immune sensitization to influenza on peripheral blood plasmablast cells, it only represented a minor fraction of the total cells [133].

Moreover, the partial CD19 antigen internalization process was observed on six different B cell lymphoma cell lines after exposition to immunotoxins FMC63(Fv)-PE38–targeting CD19 [134]. Based on the presented data, we speculate that the CD19 heterogenous phenotype demonstrated on our group initial trial and this current work is a consequence of P4 and P5 differentiation into plasma cells combined with the partial CD19 internalization after anti-CD19 antigens exposition observed on B cell lymphomas. Considering P4 characteristcs as an extra follicular plasma cell, it is expected that this subpopulation presents a more predominant CD19+ phenotype than P5. The predominant CD19 negative phenotype demonstrated by

Quentmeier compared to our result is possibly due to distinct time on the differentiation process where both experiments were initiated, indicating that our experiment were initiated in an earlier stage of the differentiation process.

6.5. U2932 subpopulations demonstrate a stochastic state of transition from day 36 to 100.

All subpopulations in the U2932 cell line demonstrated CD10 high, CD19 high, CD38 high, CD45 high, IgM high, lambda high, CD5 low, CD20 low, CD23 low, CD27 low, kappa low and CD43 predominantly high expression at 100 days after sorting. This major phenotype has demonstrated efficiency for the characterization of B cell states of differentiation, by utilizing CD10, CD38, CD27, CD19, CD5 and CD45 as a "backbone", whereas IgM/IgD expression classify specifically the cells into hematogones, immature B cell or memory cells subtypes. Due to the marked CD19 positive expression throughout the entire B cell development and differentiation process, this CD marker positive or high expression indicates B cell identity. However, the heterogenous expression of CD38 presenting high and low expression excludes the plasma cells classification. For all subpopulations, CD38 high/low expression were combined with CD10 and IgM high expression, CD5 and CD27 low expression suggesting that U2932 cells correspond to immature B cell at day 100 after sorting [135]. This classification is supported by other studies which had also performed immunophenotype classification of B cells originated from bone marrow utilizing surface marker and flow cytometry in a large cohort [136, 137]. However, it is necessary to highlight that our phenotype characterization does not include the IgD which presented important function for the cells distinction into hematogones, immature B cell or memory cells. Moreover, our samples correspond to malignant cells which might not necessarily reflect normal B cells biological function.

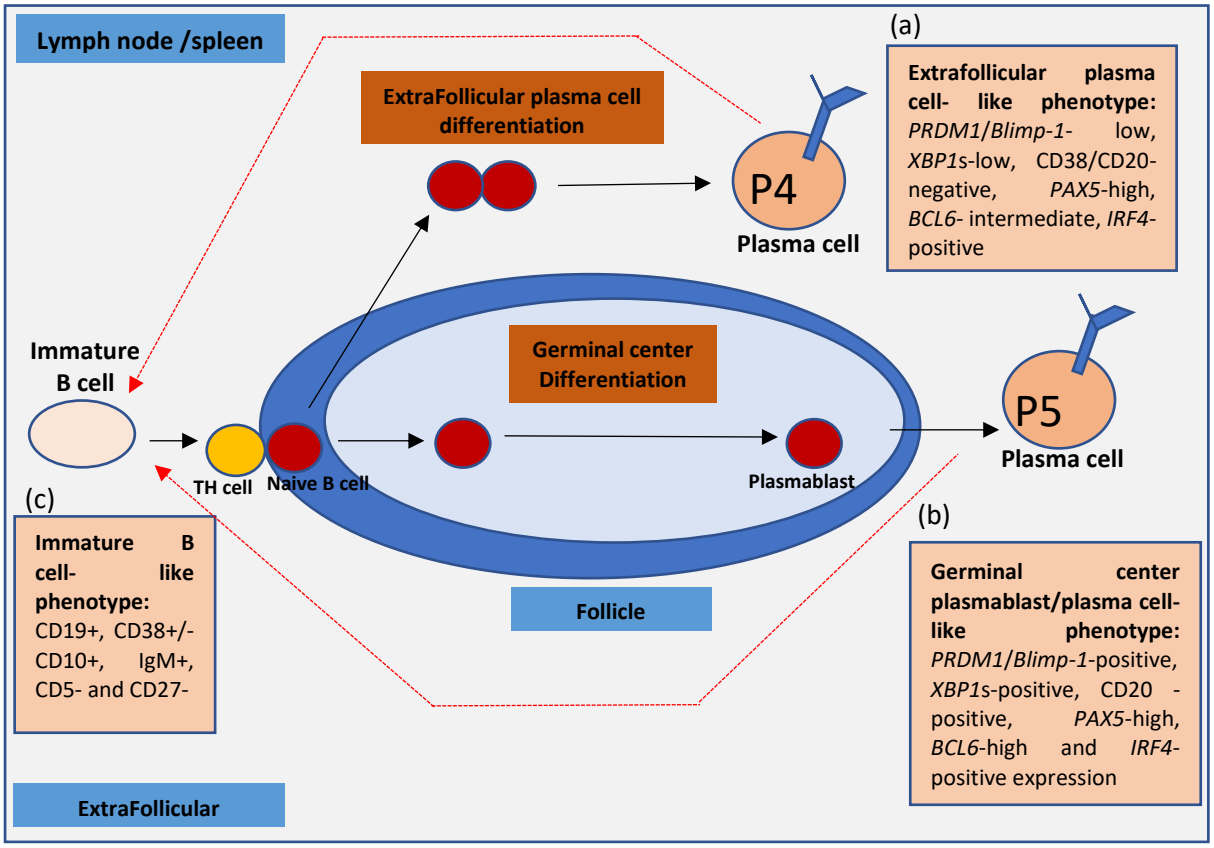

***Figure 6.5**: U2932 subpopulations phenotype and its correspondence to different states of B cell differentiation.* Naive antigen-activated B cells receive signals from 'T-cell help' and migrate to primary B-cell follicles in secondary lymphoid or to lymph node or spleen where the germinal centres differentiation or the extrafollicular differentiation are established, respectively. The red arrows indicate the stochastic state of transition admitted by both subpopulations day 107. (a). P4 phenotype resembles to an extrafollicular plasma cell, admitting *PRDM1*/*Blimp-1*-low, *XBP1*s-low, CD38/CD20 -positive, *PAX5*-high, *BCL6*-high and *IRF4*-positive, at 36 days. (b). P5 phenotype resembles to germinal center plasmablast/plasma cell, admitting *PRDM1*/*Blimp-1*- positive low, *XBP1*s-positive positive, CD38/CD20 -negative, *PAX5*-high, *BCL6*- intermediate, *IRF4*- positive at 36 days. (c). P4 and P5 phenotype resembles to an immature B cell, suggesting the occurrence of a B cell dedifferentiation. Both subpopulation express CD19+, CD38+/-, CD10+, IgM+, CD5- and CD27-.

This data suggests the regression of P4 and P5 on the differentiation process, acquiring earlier states of differentiation phenotype (Figure 6.5). P4 previously demonstrated extrafollicular plasma cell phenotype whereas P5 were classified and presented plasmablast/plasma cell phenotype. In this sense, we assume that the subpopulations acquired a stochastic state transitions between different GC B cells differentiation states to ensure phenotypic equilibrium for U2932 overall subpopulation. The phenotype proportional equilibrium among subpopulation composing a tumour were already proved by the Markov model, and verified in vivo and in vitro cell line. The mechanism of phenotype stabilization for U2932 might occur through establishment of equal proliferation rates among all subpopulations [102]. Thus, we

speculate that V(D)J recombination occurring on immature B cell is prone to mutations favouring the tumor survival, in this case, the DLBCL [101].

This data points to the U2932 cell line hierarchically organization based on the tumorigenic potential. In order words, this cell line apparently follow the CSCM of development with subpopulations transiting back into states mediated by genetic alterations [138]. Interestingly, mature B cells dedifferentiation into uncommitted progenitor cells was already observed in rodents presenting *PAX5* deletion, resulting into aggressive lymphoma with similar progenitor phenotype [138]. These results demonstrate an elevated plasticity of mature differentiated cells. In addition, it shows that the lack of a transcription factor that regulates the B cell differentiation result on cellular interconversion into previous states [138, 139]. In this sense, the lack of genetic profile data for the subpopulations at 100 days limited our analysis. However, we also speculate that a considerable fraction of the cell composing P4 died between day 36 and 100 after sorting, assuming the similarities of the P4 phenotype to the extrafollicular plasma cell and their short- lived period at the extrafollicular *foci*. After immune sensitization of B cell on spleen, the *foci* were gradually decomposed and lost from the day 10, whereas GC B lived at least 16 days [140]. This assumption is supported by the apoptosis pathway 8.0 enrichment score as result of the U2932 vs P4 enrichment analysis.

6.6. CD43 heterogeneous expression suggest U2932 is still in a B cell subset transition.

There is no specific phenotype classification determining the B cells differentiation state for malignant DLBCL based on surface markers expression. However, the CD43 positive expression is recurrent in approximately 25% of DLBCL cases in Western countries, indicating the occurrence of this antigen expression in NHL. This antigen expression has also been associated with lower prognostic of DLBCL patients with a predominant non- GC phenotype [98]. Flow cytometry analysis of B cells from peripheral blood samples of health donors utilized CD27 and CD43 expression to determine phenotypically the position of CD43 positive B cells. Flow cytometry analyses identified CD43 B cell subsets based on an isotype control phenotype, CD27 and CD43 negative expression was correspondent to naïve B cells on the isotype control. Further analyses utilized CD43 positive and negative B cells, evaluating its transition into a plasmablast- like cell thorough its final phenotype expression after the cell exposition to plasmablast inducers (CpG-ODN, CD40L, IL-2, IL-10 and IL-15) in an *in vitro* model. After 4 days of incubation, CD43 positive B cells demonstrated CD20 negative, CD27

negative and CD43 positive, demonstrating a plasmablast-like cell phenotype [141]. At 100 days after sorting, U2932 and all subpopulations predominantly demonstrated CD20 negative, CD27 negative and CD43 positive. In this sense, U2932 again demonstrated distinct phenotypes characterizing different states of differentiation of the subpopulations composing this cell line. We speculate that the U2932 cell line is still in a stage of transition presenting a fraction of its cell in the B cell subsets state between plasmablast/plasma- and immature B cell.

6.7. Distinct immunophenotype and genetic mutations among subpopulations results into a variable level of resistance to doxorubicin

HIF-1 pathway, demonstrated on P4 vs P5 enrichment, presents clear association with drug resistance of chemotherapeutic drugs, including anthracyclines, which might have caused P5 elevated drug resistance. Microarray and RT- PCR analyses demonstrated time relation between hypoxia condition and *multidrug resistance 1* (*MDR1*) gene mRNA expression on epithelial cell lines. In addition, oligonucleotides presenting HIF-1 site deletion and submitted to hypoxia demonstrated low *MDR1* activity, demonstrating HIF-1 importance for *MDR1* induction. ELISA analyses demonstrated that *MDR1* promoted surface expression of membrane-resident P-glycoprotein, an ATP- bidding cassette transporter that nonselectively transport myriad of amphipathic molecules causing drug efflux. Finally, doxorubicin presented diminished induction of cell death on samples previously submitted to hypoxia for 24 and 48h, presenting population approximately 3 times lower for the controls not submitted to restricted oxygen conditions [142]. Hence, we speculate that in our samples HIF-1 pathway induced *MRD1* expression which promoted P-gp expression on cell surface, culminating into doxorubicin efflux and U2932 subpopulation variable doxorubicin resistance.

Moreover, a marked difference on the P4 phenotype compared to the other subpopulations and original cell line is the constant high CD20 expression for 36 days after sorting. After human B cell activation, CD20 has demonstrated regulation of the cells progression in the cell cycle by modulating the calcium flow [73]. Whole-cell patch- clamp experiments demonstrated elevated Ca^{2+} flow through plasma membrane of T lymphoblasts, erythroleukemia and fibroblasts which had demonstrated CD20 positive phenotype on flow cytometry analysis. The same activity was not observed on the calcium channel of cells presenting CD20 negative expression. Consequently, further florescence analysis demonstrated a higher cytosolic free

Ca^{2+} on lymphoblasts expressing CD20 by cDNA transfection compared to the negative control [143].

Calcium influx promotes CD20 positive cells progression to S/G2 phases of the cell cycle when not inhibited by specific anti- CD20 mAb [73]. Doxorubicin main mechanism of action is the disruption of topoisomerase-II-mediated DNA repair [40]. Topoisomerase II activity is fundamental for the appropriate development of G2. Due to this DNA gyrase inhibition at G2 phase mammalian cells demonstrated delayed progression to mitosis, whereas once topoisomerase II is removed the cells rapidly progressed to the mitotic stage [144]. In this study, all cell lines exposed to the C2 and C3 doxorubicin dose demonstrated a reduction in of 50% to 75% of their original population after 48h. By promoting topoisomerase II disruption, doxorubicin maintains the cells in an erroneous pro- mitotic phase which might culminate into cellular apoptosis. In this sense, the CD20 positive phenotype facilitates the P4 progression to S/G2, resulting into the lower resistance of this subpopulation to doxorubicin due to the relevance of the topoisomerase II function at this cell cycle phase. Finally, we speculate that the common CD20 lower expression observed in all U2932 subpopulations reflects on the similar doxorubicin resistance level after 48h of all four doses of drug exposition at 100 days after sorting. The elevated standard deviation on U2932 samples is possibly the cause of significant difference at C1 for U2932 vs P4 by Tukey multiple comparison test.

Although P4 does not demonstrate HIF-1 as a pathway enrichment and has CD20 elevated expression, P4 presents transcriptional misregulation pathway with elevated BCL6 expression which have already demonstrated inhibition of p53 pathway, suppressing its pro- apoptotic function [145]. Consequently, BCL6 high expression promote drug resistance on apoptotic p53 dependent therapies, such as doxorubicin. In this sense, the apoptotic induction from doxorubicin becomes only effective through reactive oxygen species activity [146]. This assumption is consistent with REGS classification of U2932, P4, P5 as doxorubicin resistant.

7. Conclusion

The procedures developed in this study revealed that U2932 cell line is initially organized in subpopulations phenotypically and genetically characterizing distinct B cell states of differentiation. The subpopulations later assumed a stochastic state of transiting ensuring the tumor phenotypic equilibrium. P4 and P5 were the main subpopulations characterized is this study due to their elevated phenotype consistency during the first month of analyses. At day 36, P4 demonstrated an extrafollicular plasma cell- like phenotype, *PRDM1*/*Blimp-1*- low, *XBP1*s-low, CD38/CD20 -negative, *PAX5*-high, *BCL6*- intermediate, *IRF4*- positive and EBV expression, whereas P5 demonstrated a germinal center plasmablast/plasma cell- like phenotype, *PRDM1*/*Blimp-1*-positive, *XBP1*s-positive, CD20 -positive, *PAX5*-high, *BCL6*-high and *IRF4*-positive expression. The distinct genetic alterations potentially driven each subpopulation to the acquisition of specific B cell subset phenotype favouring the tumor heterogeneity and chemotherapeutic varied sensibility, which was demonstrated in this study by the P4 and P5 differentiated resistance to doxorubicin.

At day 107, both subpopulations presented an immature B cell- like phenotype, CD19+, CD38+/-, CD10+, IgM+, CD5- and CD27-, demonstrating the malignant cell dedifferentiation into a state prone to genetic mutation. This data partially agrees with the Quentmeier *et al. 2013*, both studies demonstrate a steady phenotype for U2932 subpopulations for 35 days. The two studies demonstrated a *BCL2*- high for both subpopulation, *MYC*- high for the subpopulation expressing CD30/CD20 low and *BCL6*- high for the subpopulation expressing CD30/CD20 high. However, the analysis developed at day 100 indicates that P4 and P5 regained of original cell line phenotype excluding the possibility of sub- clone formation pointed by the previous study. The result divergence of the two distinct studies might be caused by the original cell line sorting at distinct moment after initiation of U2932 culture. In this study, the original cell line was only submitted to FACS after 22 days of culture. Consequently, the subpopulations reanalysed at day 100 in Quentmeier *et al. 2013* corresponds to the samples in an earlier time point in our experiment. Hence, U2932 cell line represents an appropriate tool for the understanding of heterogeneity for ABC DLBCL or NHL, but not a model for clone development in tumours.

Although U2932 correspond to an ABC-like DLBCL, this cell line presents particularities not commonly observed in this tumor subtype. U2932 cell line was isolated from a 29 years old patient with Waldeyer's tonsillar ring as the tumor original location. Nevertheless, DLBCL is

more commonly recurrent among elderly patients with this tumour formation at Waldeyer's tonsillar ring representing approximately 10% of DLBCL cases. In this sense, U2932 cell line correspond to an exceptional DLBCL case. This cell line specificities should be taken into consideration when utilizing U2932 as a tool for cancer heterogeneity studies. Moreover, U2932 characterization is still requesting improvement for an enhanced comprehension of this cell line. In this study, data about genetic alteration was not developed for the day 100, as well as, the initial immunophenotype characterization was only substantiated by three CD markers expression, restricting the subpopulations appropriate classification into B cell subset. Additionally, the limited data about B cells differentiation state for malignant DLBCL based on surface markers expression turned necessary the comparison of the subpopulation phenotype to healthy B cells or another NHL, which not necessarily correspond to the ABC DLBCL phenotype.

8. References

[1]. Küppers, R. 2005. Mechanisms of B-cell lymphoma pathogenesis. Nature Reviews Cancer, 5(4), pp. 251-262.

[2]. Nogai, H. Dörken, H and Lenz, G. 2011. Pathogenesis of Non-Hodgkin's Lymphoma. Journal of Clinical Oncology, 29(14), pp. 1803-1811.

[3]. Seifert, M., Scholtysik, R. and Küppers, R. 2013. Origin and pathogenesis of B cell lymphomas. United States.

[4]. Pelanda, R. and Torres, R.M. 2012. Central B-Cell Tolerance: Where Selection Begins. Cold Spring Harbor Perspectives in Biology, 4(4), pp. a007146.

[5]. Basso, K., and Dalla-Favera, R. 2015. Germinal centres and B cell lymphomagenesis. Nature Reviews. Immunology, 15(3), pp. 172-184.

[6]. Suan, D., Sundling, C., and Brink, R. 2017. Plasma cell and memory B cell differentiation from the germinal center. Current Opinion in Immunology, 45, pp. 97-102.

[7]. Maclennan, I., Toellner, K., Cunningham, A.F., Serre, K., Sze, D., Zuniga, E., Cook, M.C. and Vinuesa, C.G. 2003. Extrafollicular antibody responses. Immunological Reviews, 194(1), pp. 8-18.

[8]. Paus, D., Phan, T.G., Chan, T.D., Gardam, S., Basten, A. and Brink, R. 2006. Antigen recognition strength regulates the choice between extrafollicular plasma cell and germinal center B cell differentiation. Journal of Experimental Medicine, 203(4), pp. 1081-1091

[9]. Basso, K., and Dalla-Favera, R. 2012. Roles of BCL6 in normal and transformed germinal center B cells. Immunological Reviews, 247(1), pp. 172-183

[10]. Busslinger, M., Delogu, A., Schebesta, A. and Cobaleda, C. 2007. Pax5: the guardian of B cell identity and function. Nature Immunology, 8(5), pp. 463-470.

[11]. Recaldin, T., And Fear, D.J. 2016. Transcription factors regulating B cell fate in the germinal centre. Clinical & Experimental Immunology, 183(1), pp. 65-75.

[12]. Conacci-sorrell, M., Mcferrin, L. and Eisenman, R.N. 2014. An overview of MYC and its interactome. Cold Spring Harbor perspectives in medicine, 4(1), pp. a014357.

[13]. Korać, P., Dotlić, S., Matulić, M., Petranović, M. Z., and Dominis, M. 2017. Role of MYC in B Cell Lymphomagenesis. Genes, 8(4), pp. 115.

[14]. Kallies, A., Hasbold, J., Tarlinton, D.M., Dietrich, W., Corcoran, L.M., Hodgkin, P.D. and Nutt, S.L., 2004. Plasma Cell Ontogeny Defined by Quantitative Changes in Blimp-1 Expression. Journal of Experimental Medicine, 200(8), pp. 967-977.

[15]. Tellier, J., Shi, W., Minnich, M., Liao, Y., Crawford, S., Smyth, G.K., Kallies, A., Busslinger, M. and NUTT, S.L. 2016. Blimp-1 controls plasma cell function through the regulation of immunoglobulin secretion and the unfolded protein response. Nature Immunology, 17(3), pp. 323-330.

[16]. The Non-Hodgkin's Lymphoma Classification Project. 1997. A clinical evaluation of the International Lymphoma Study Group classification of non-Hodgkin's lymphoma. Blood, 89(11), pp. 3909.

[17]. Hanahan, D., And Weinberg, R. 2011. Hallmarks of Cancer: The Next Generation. Cell, 144(5), pp. 646-674.

[18]. Swerdlow, S.H., Campo, E., Pileri, S.A., Harris, N.L., Stein, H., Siebert, R., Advani, R., Ghielmini, M., Salles, G.A., Zelenetz, A.D. And Jaffe, E.S. 2016. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. Blood, 127(20), pp. 2375-2390.

[19]. Shankland, K.R., Armitage, J.O., and Hancock, B.W. 2012. Non-Hodgkin lymphoma. Lancet, The, 380(9844), pp. 848-857.

[20]. Schneider, C., Pasqualucci, L., And Dalla-Favera, R. 2011. Molecular pathogenesis of diffuse large B-cell lymphoma. Seminars in Diagnostic Pathology, 28(2), pp. 167-177.

[21]. Martelli, M., Ferreri, A.J., Agostinelli, C., Di Rocco, A., Pfreundschuh, M., Pileri, S.A. 2013. Diffuse large B-cell lymphoma. Critical Reviews in Oncology/Hematology; 87(2):146-71.

[22]. Swerdlow, S.H., Campo, E., Harris, N.L., Jaffe, E.S., Pileri, S.A., Stein, H. WHO classification of tumors of haematopoietic and lymphoidtissues. 4th ed. Lyon: IARC Press; 2008

[23]. Lymphoma.dk

[24]. Tilly, H., Gomes da Silva, M., Vitolo, U., Jack, A., Meignan, M., Lopez-Guillermo, A., Walewski, J., André, M., Johnson, P.W., Pfreundschuh, M. and Ladetto, M., 2015. ESMO Guidelines Committee. Diffuse large B-cell lymphoma (DLBCL): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology; 26 Suppl 5:v116-25

[25]. Maartense, E., Kluin-Nelemans, H.C., Le Cessie, S., Kluin, P.M., Snijder, S. And Noordijk, E.M. 2000. Different age limits for elderly patients with indolent and aggressive non-Hodgkin lymphoma and the role of relative survival with increasing age. Cancer, 89(12), pp. 2667-2676.

[26]. López-Guillermo, A., Colomo, L., Jiménez, M., Bosch, F., Villamor, N., Arenillas, L., Muntañola, A., Montoto, S., Giné, E., Colomer, D., Beà, S., Campo, S. and Montserrat, E. 2005. Diffuse Large B-Cell Lymphoma: Clinical and Biological Characterization and Outcome According to the Nodal or Extranodal Primary Origin. Journal of Clinical Oncology, 23(12), pp. 2797-2804.

[27]. Moller, M.B., Pedersen, N.T. and Christensen, B.E. 2004. Diffuse large B-cell lymphoma: clinical implications of extranodal versus nodal presentation a population-based study of 1575 cases. British Journal of Haematology, 124(2), pp. 151-159.

[28]. Montoto, S. and Fitzgibbon, J. 2011. Transformation of Indolent B-Cell Lymphomas. Journal of Clinical Oncology, 29(14), pp. 1827-1834.

[29]. Jamroziak, K., Tadmor, T., Robak, T. and Polliack, A. 2015. Richter syndrome in chronic lymphocytic leukemia: updates on biology, clinical features and therapy. Leukemia & Lymphoma, 56(7):1949-58.

[30]. Lenz, G., Wright, G. W., Tolga Emre, N. C., Kohlhammer, H., Dave, S.S., Eric Davis, R., Carty, S., Lam, L.T., Shaffer, A.L., Xiao, W., Powell, J., Rosenwald, A., Ott, G., Muller-Hermelink, H.K., Gascoyne, R.D., Connors, J.M., Campo, E., Jaffe, E.S., Delabie, J., Smeland, E.B., Rimsza, L.M., Fisher, R.I., Weisenburger, D.D., Chan, W.C. and Staudt, L.M. 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. 13520-13525.

[31]. Alizadeh, A.A., Hudson, J., Weisenburger, D.D., Byrd, J.C., Moore, T., Boldrick, J.C., Greiner, T.C., Eisen, M.B., Rosenwald, A., Tran, T., Ma, C., Yang, L., Armitage, J.O., Warnke, R., Marti, G.E., Grever, M.R., Lu, L., Sherlock, G., Staudt, L.M., Levy, R., Chan, W.C., Davis, R.E., Tibshirani, R., Botstein, D., Brown, P.O., Lossos, I.S., Powell, J.I., Sabet, H., Wilson, W., Yu, X. and Lewis, D.B. 2000. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature, 403(6769), pp. 503-511.

[32]. Muppidi, J.R., Schmitz, R., Green, J.A., Xiao, W., Larsen, A.B., Braun, S.E., AN, J., XU, Y., Rosenwald, A., Ott, G., Gascoyne, R.D., Rimsza, L.M., Campo, E., Jaffe, E.S., Delabie, J., Smeland, E.B., Braziel, R.M., Tubbs, R.R., Cook, J.R., Weisenburger, D.D., Chan, W.C., Vaidehi, N., Staudt, L.M. and Cyster, J.G. 2014. Loss of signalling via Gα13 in germinal centre B-cell-derived lymphoma. Nature, 516(7530), pp. 254-258.

[33]. Hans, C.P., Weisenburger, D.D., Greiner, T.C., Gascoyne, R.D., Delabie, J., Ott, G., Müller-Hermelink, H.K., Campo, E., Braziel, R.M., Jaffe, E.S., Pan, Z., Farinha, P., Smith, L.M., Falini, B., Banham, A.H., Rosenwald, A., Staudt, L.M., Connors, J.M., Armitage, J.O. and Chan, W.C. 2004. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. Blood, 103(1), pp. 275-282.

[34]. Davis, R. E., Brown, K. D., Siebenlist, U. and Staudt, L. M. 2001. Constitutive Nuclear Factor kappaB Activity Is Required for Survival of Activated B Cell – like Diffuse Large B Cell Lymphoma Cells, J Exp Med., vol. 194, no. 12, pp. 1861–1874, 2001.

[35]. Dybkær, K., Bøgsted, M., Falgreen, S., Bødker, J.S., Kjeldsen, M.K., Schmitz, A., Bilgrau, A.E., Xu-Monette, Z.Y., Li, L., Bergkvist, K.S., Laursen, M.B., Rodrigo-Domingo, M., Marques, S.C., Rasmussen, S.B., Nyegaard, M., Gaihede, M., Møller, M.B., Samworth, R.J., Shah, R.D., Johansen, P., El-Galaly, T.C., Young, K.H. And Johnsen, H.E. 2015. Diffuse large B-cell lymphoma classification system that associates normal B-cell subset phenotypes with prognosis. Journal of clinical oncology: official journal of the American Society of Clinical Oncology, 33(12), pp. 1379-1388.

[36]. Falgreen, S., Dybkær, K., Young, K.H., Xu-Monette, Z.Y., El-Galaly, T.C., Laursen, M.B., Bødker, J.S., Kjeldsen, M.K., Schmitz, A., Nyegaard, M., Johnsen, H.E., Bøgsted, M. 2015. Predicting response to multidrug regimens in cancer patients using cell line experiments and regularised regression models. BMC Cancer,15:235.

[37]. Held, G., Pöschel, V. and Pfreundschuh, M. 2006. Rituximab for the treatment of diffuse large B-cell lymphomas. Expert Review of Anticancer Therapy, 6(8):1175-86.

[38]. Dotan, E., Aggarwal, C., Smith, M. R. 2010. Impact of Rituximab (Rituxan) on the Treatment of B-Cell Non-Hodgkin's Lymphoma. Physical therapy science, 35(3):148-57.

[39]. Fleming, R. A. 1997. An overview of cyclophosphamide and ifosfamide pharmacology. Pharmacotherapy, 17(5 Pt 2):146S-154S. US.

[40]. Thorn, C., Oshiro, C., Marsh, S., Hernandez-Boussard, T., McLeod, H., Klein, T. and Altman, R. 2011. Doxorubicin pathways: pharmacodynamics and adverse effects. Pharmacogenetics and Genomics, 21(7), pp. 440-446.

[41]. Jordan, M.A. 2002. Mechanism of action of antitumor drugs that interact with microtubules and tubulin. Current medicinal chemistry. Anti-cancer agents, 2(1), pp. 1-17.

[42]. Sept, D. 2007. Microtubule Polymerization: One Step at a Time. Current Biology, 17(17), pp. R766.

[43]. Lamar, Z. S. 2016. The Role of Glucocorticoids in the Treatment of Non-Hodgkin Lymphoma. Annals of Hematology & Oncology, 3(7): 1103.

[44]. Wilson, W.H. 2013 Treatment strategies for aggressive lymphomas: what works? Hematology. American Society of Hematology. Education Program; 2013:584-90.

[45]. Hagemeister, F. 2010. Rituximab for the Treatment of Non-Hodgkins Lymphoma and Chronic Lymphocytic Leukaemia. Drugs, 70(3), pp. 261-272.

[46]. Friedberg, J.W. 2011. Relapsed/refractory diffuse large B---cell lymphoma. Hematology Am Soc Hematol Educ Program, 498–505.

[47]. Huang, J. Z., Sanger, W. G., Greiner, T. C., Staudt, L. M., Weisenburger, D.D, Pickering, D. L., Lynch, J. C., Armitage, J. O., Warnke, R. A, Alizadeh, A., Lossos, I. S., Levy, R. and Chan, W. C. 2002. The t(14;18) defines a unique subset of diffuse large B-cell lymphoma with a germinal center B-cell gene expression profile," Blood, vol. 99, no. 7, pp. 2285–2291.

[48]. Bhatia, S., Frangioni, J.V., Hoffman, R.M., Iafrate, A.J. and Polyak, K. 2012. The challenges posed by cancer heterogeneity. Nature biotechnology, 30(7), pp. 604.

[49]. Swanton, C. 2012. Intratumor heterogeneity: evolution through space and time. Cancer research, 72(19), pp. 4875.

[50]. Rang, H. P., Dale, M.M., Ritter, J. M., G. 2011, "Anticancer drugs" in Rang, H. P., Dale, M.M., Ritter, J. M., Flower, R.J., Henderseon, G (7 ed.), Pharmacology, Elsevier, pp.681- 682.

[51]. Arcamone, F., Cassinelli, G., Fantini, G., Grein, A., Orezzi, P., Pol, C., and Spalla, C. 1969. Adriamycin, 14- hydroxydaunomycin: a new antitumor antibiotic from S. peucetius var. caesius. Biotechnol Bioeng; 11:1101–1110.

[52]. Cortes-Funes, H. and Coronado, C. 2007. Role of anthracyclines in the era of targeted therapy. Cardiovascular Toxicology, 7:56–60.

[53]. Tewey, K.M., Rowe, T.C., Yang, L., Halligan, B.D. and Liu, L.F. 1984. Adriamycin-induced DNA damage mediated by mammalian DNA topoisomerase-II. Science, 226:466–468.

[54]. Weiss, R. B. 1992. The anthracyclines: will we ever find a better doxorubicin? Seminars in Oncology, 19:670–686.

[55]. Doroshow, J.H. 1986. Role of hydrogen peroxide and hydroxyl radical formation in the killing of Ehrlich tumor cells by anticancer quinones. Proceedings of the National Academy of Sciences, 83(12):4514-8.

[56]. Carvalho, C., Santos, R.X., Cardoso, S., Correia, S, Oliveira, P.J. and Santos, M.S. 2009. Doxorubicin: the good, the bad and the ugly effect. Current Medicinal Chemistry, 16:3267–3285.

[57]. Swain, S.M., Whaley, F.S. and Ewer, M.S. 2003. Congestive heart failure in patients treated with doxorubicin: a retrospective analysis of three trials. Cancer,97:2869–2879.

[58]. Vaidya, R. and Witzig and T.E. 2014. Prognostic factors for diffuse large B-cell lymphoma in the R(X)CHOP era. Annals of oncology, 25(11):2124–33.

[59]. Gottesman, M.M. 2002. Mechanisms of Cancer Drug Resistance. Annual Review of Medicine, 53:615–27.

[60]. Zheng, T., Wang, J., Chen, X. and Liu L. 2010. Role of microRNA in anticancer drug resistance. International Journal of Cancer 126(1):2–10.

[61]. Allen, K.E. and Weiss, G.J. 2010. Resistance may not be futile: microRNA biomarkers for chemoresistance and potential therapeutics. Molecular Cancer Therapy, 9(12):3126–36.

[62]. Biomarkers Definitions Working Group. 2001. "Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework," Clinical Pharmacology and Therapeutics, 69 (3):89–95.

[63]. Pepe, M. S., Janes, H., Longton, G., Leisenring, H.W., and Newcomb, P. 2004. Limitations of the Odds Ratio in Gauging the Performance of a Diagnostic, Prognostic, or Screening Marker, American Journal of Epidemiology, 159 (9): 882–890.

[64]. Pepe, M. S., Etzioni, R., Feng, Z., Potter, J. D., Lou, M., Thornquist, M., Winget, M. and Yasui, Y. 2001 Phases of Biomarker Development for Early Detection of Cancer, Journal of the National Cancer Institute, 93(14): 054–1061.

[65]. Lossos, I.S., and Morgensztern, D. 2006. Prognostic Biomarkers in Diffuse Large B-Cell Lymphoma. Journal of Clinical Oncology, 24(6): 995-1007.

[66]. Perez-Andres, M., Paiva, B., Nieto, W.G., Caraux, A., Schmitz, A., Almeida, J., Vogt, R.F. Jr, Marti, G.E., Rawstron, A.C., Van Zelm, M.C., Van Dongen, J.J., Johnsen, H.E., Klein, B. and Orfao, A. 2010. Primary Health Care Group of Salamanca for the Study of MBL. Human peripheral blood B-cell compartments: a crossroad in B-cell traffic. Cytometry. Part B, Clinical cytometry,78.

[67]. Wang, K., Wei, G. and Liu, D. 2012. CD19: a biomarker for B cell development, lymphoma diagnosis and therapy. Experimental Hematology & Oncology, 29;1(1):36.

[68]. Otero, D.C., Anzelon, A.N. and Rickert, R.C. 2003. CD19 function in early and late B cell development: I. Maintenance of follicular and marginal zone B cells requires CD19-dependent survival signals. The Journal of Immunology, 170(1):73-83.

[69]. Yang, W., Agrawal, N., Patel, J., Edinger, A., Osei, E., Thut, D., Powers J, Meyerson, H. 2005. Diminished expression of CD19 in B-cell lymphomas. Cytometry. Part B, Clinical cytometry, 63(1):28-35.

[70]. Del Nagro, C.J., Otero, D.C., Anzelon, A.N., Omori, S.A., Kolla, R.V. and Rickert, R.C. 2005. CD19 function in central and peripheral B-cell development. Immunologic Research, 31(2):119-31.

[71]. Anolik, J., Looney, R.J., Bottaro, A., Sanz, I. and Young, F. 2003. Down-regulation of CD20 on B cells upon CD40 activation. European Journal of Immunology, 33(9), pp. 2398-2409.

[72]. Uchida, J., Lee, Y., Hasegawa, M., Liang, Y., Bradney, A., Oliver, J.A., Bowen, K., Steeber, D.A., Haas, K.M., Poe, J.C. and Tedder, T.F. 2004. Mouse CD20 expression and function. International Immunology, 16(1), pp. 119-129.

[73]. O'Keefe, T.L., Williams G.T., Davies S.L. and Neuberger M.S. 1998. Mice carrying a CD20 gene disruption. Immunogenetics, 48(2):125-32.

[74]. Tedder, T.F. and Engel, P.1994. CD20: a regulator of cell-cycle progression of B lymphocytes. Immunology Today, 15(9):450-4.

[75]. Deans, J.P., Li, H. and Polyak, M.J. 2002. CD20-mediated apoptosis: signalling through lipid rafts. Immunology, 107(2), pp. 176-182.

[76]. Romero-Ramírez H., Morales-Guadarrama, M.T., Pelayo, R., López-Santiago, R., Santos-Argumedo, L. 2015. CD38 expression in early B-cell precursors contributes to extracellular signal-regulated kinase-mediated apoptosis. Immunology,144(2):271-81.

[77]. Funaro, A., Morra, M., Calosso, L., Zini, M.G., Ausiello, C.M. and Malavas, I. F.1997. Role of the human CD38 molecule in B cell activation and proliferation. Tissue Antigens, 49(1): 7-15.

[78]. Vences-Catalán, F. and Santos-Argumedo, L., 2011. CD38 through the life of a murine B lymphocyte. IUBMB Life, 63(10): 840-846.

[79]. Xu, D. 2006. Dual surface immunoglobulin light-chain expression in B-cell lymphoproliferative disorders. Archives of Pathology & Laboratory Medicine,130(6):853-6.

[80]. Koziner, B., Stavnezer, J., Al-Katib, A., Gebhard, D., Mittelman, A., Andreeff, M. and Clarkson, B.D.1986. Surface immunoglobulin light chain expression in pre-B cell leukemias. Annals of the New York Academy of Sciences, 468(1 Clinical Cyto): 211-226.

[81]. Shusfik, C., Bergsagel, D. E and Pruzanski, W. 1976. K and A Light Chain Disease: Survival Rates and Clinical Manifestations.

[82]. Hermiston, M. L., Xu, Z. and Weiss, A. 2003. CD45: A Critical Regulator of Signaling. Annual Review of Immunology, 21:107-37.

[83]. Huntington, N.D. and Tarlinton, D.M. 2004. CD45: direct and indirect government of immune regulation. Immunology Letters, 94(3):167-74.

[84]. Uherova, P., Ross, C.W., Schnitzer, B., Singleton, T.P. and Finn and W.G. 2001. The clinical significance of CD10 antigen expression in diffuse large B-cell lymphoma. American journal of clinical pathology, 115(4): 582-588

[85]. Mishra, D., Singh, S. and Narayan, G. 2016. Role of B Cell Development Marker CD10 in Cancer Progression and Prognosis. Molecular Biology International, 2016:4328697, pp. 1-9.

[86]. Capolunghi, F., Rosado, M.M., Sinibaldi, M., Aranburu, A. and Carsetti, R. 2013. Why do we need IgM memory B cells? Immunology Letters, 152(2):114-120.

[87]. Schmidlin, H., Diehl, S.A. and Blom, B. 2009. New insights into the regulation of human B-cell differentiation. Trends in Immunology, 30(6): 277-285.

[88]. Agematsu, K. 2000. Memory B cells and CD27. Histology and Histopathology, 15(2):573-6.

[89]. Wu, Y.C., Kipling, D. and Dunn-Walters, D.K. 2011. The relationship between CD27 negative and positive B cell populations in human peripheral blood. Frontiers in Immunology, 2:81.

[90]. Agematsu, K., Hokibara, S., Nagumo, H., Komiyama, A. 2000. CD27: a memory B-cell marker. Immunology Today, 21(5):204-6.

[91]. Kipps, T.J. 1989. The CD5 B Cell. Advances in Immunology,47: 117-187.

[92]. Youinou, P., Jamin, C. and Lydyard, P.M. 1999. CD5 expression in human B-cell populations. Immunology Today, 20(7):312-6.

[93]. Bonnefoy, J., Lecoanet-Henchoz, S., Aubry, J., Gauchat, J. and Graber, P. 1995. CD23 and B-cell activation. Current Opinion in Immunology, 7(3): 355-359.

[94]. Liu, C., Richard, K., Wiggins, M., Zhu, X., Conrad, D.H. and Song, W., 2016. CD23 can negatively regulate B-cell receptor signaling. Scientific reports, 6, pp. 25629.

[95]. Pedraza-Alva, G., Rosenstein, Y. 2007. CD43 – One molecule, many tales to recount. Signal Transduction, 7:372- 385

[96]. Bjoerck, P., Axelsson, B. and Paulie, S. 1991. Expression of CD40 and CD43 during activation of human B lymphocytes. Scandinavian Journal of Immunology, 33(2): 211-218.

[97]. Arlindo, E.M., Marcondes, N.A., Fernandes, F. B. and Faulhabera, G.A.M. 2017. Quantitative flow cytometric evaluation of CD200, CD123, CD43 and CD52 as a tool for the differential diagnosis of mature B-cell neoplasms. Revista Brasileira De Hematologia & Hemoterapia, 39(3): 252–258.

[98]. Xiao-Bo, M., Zheng, Y., Yuan, H., Jing, J. and Wang, Y.P. 2015. CD43 expression in diffuse large B-cell lymphoma, not otherwise specified: CD43 is a marker of adverse prognosis. Human Pathology, 46(4): 593-599.

[99]. Nowell, P.C. 1976 The clonal evolution of tumor cell populations. Science.; 194:23–28.

[100]. Greaves, M. and Maley, C. C. 2012. Clonal evolution in cancer. Nature, 481:306–313.

[101]. Meacham, C.E. and Morrison, S.J. 2013. Tumour heterogeneity and cancer cell plasticity. Nature, 501(7467):328-37.

[102]. Gupta, P.B., Fillmore, C.M., Jiang, G., Shapira, S.D., Tao, K., Kuperwasser, C. and Lander, E. S. 2011. Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. Cell, 146(4):633-44.

[103]. Creighton, C.J., Li, X., Landis, M., Dixon, J.M., Neumeister, V.M., Sjolund, A., Rimm, D.L., Wong, H., Rodriguez, A., Herschkowitz, J.I. 2009. Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. Proceedings of the National Academy of Sciences,106(33):13820-5

[104]. Gupta, P.B., Onder, T.T., Jiang, G., Tao, K., Kuperwasser, C., Weinberg, R.A., and Lander, E.S. 2009. Identification of selective inhibitors of cancer stem cells by high-throughput screening. Cell 138(4):645-659

[105]. Li, X., Lewis, M.T., Huang, J., Gutierrez, C., Osborne, C.K., Wu, M.F., Hilsenbeck, S.G., Pavlick, A., Zhang, X., Chamness, G.C., 2008. Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. Journal of the National Cancer Institute, 100(9):672-9

[106]. Woodward, W.A., Chen, M.S., Behbod, F., Alfaro, M.P., Buchholz, T.A., and Rosen, J.M. 2007. WNT/beta-catenin mediates radiation resistance of mouse mammary progenitor cells. Proceedings of the National Academy of Sciences 104(2):618-23

[107]. Sharma S.V., Lee, D.Y., Li, B., Quinlan, M.P., Takahashi, F., Maheswaran, S., McDermott, U., Azizian, N., Zou, L., Fischbach, M.A., Wong, K.K., Brandstetter, K., Wittner, B., Ramaswamy, S., Classon, M. and Settleman, J. 2010. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. Cell, 141(1):69-80

[108]. Quentmeier, H., Amini, R.M., Berglund, M., Dirks, W.G., Ehrentraut, S., Geffers, R., Macleod, R.A., Nagel, S., Romani, J., Scherr, M., Zaborski, M. and Drexler, H. G. 2013. U-2932: two clones in one cell line, a tool for the study of clonal evolution. Leukemia, 27(5):1155-64.

[109]. Amini, R.M., Berglund, M., Rosenquist, R., Von Heideman, A., Lagercrantz, S., Thunberg, U., Bergh, J., Sundström, C., Glimelius, B. and Enblad, G. 2002. A novel B-cell line (U-2932) established from a patient with diffuse large B-cell lymphoma following Hodgkin lymphoma. Leukemia & Lymphoma, 43(11):2179-89.

[110]. Bergkvist, K.S., Nyegaard, M., Bøgsted, M., Schmitz, A., Bødker, J.S., Rasmussen, S.M., Perez-Andres, M., Falgreen, S., Bilgrau, A.E., Kjeldsen, M.K., Gaihede, M., Nørgaard, M.A., Bæch, J., Grønholdt, M.L., Jensen, F.S., Johansen, P., Dybkær, K. and Johnsen, H.E. 2014. Validation and implementation of a method for microarray gene expression profiling of minor B-cell subpopulations in man. BMC Immunology, 15:3.

[111]. Hebert, P.D., Cywinska, A., Ball, S.L. and de Waard, J.R. 2003. Biological identifications through DNA barcodes. Proceedings. Biological Sciences, 270(1512):313-21.

[112]. Caraux, A., Klein, B., Paiva, B., Bret, C., Schmitz, A., Fuhler, G., Bos, N., Johnsen, H., Orfao, A. and Perez-andres, M. 2010. Circulating human B and plasma cells. Age-associated changes in counts and detailed characterization of circulating normal CD138- and CD138+ plasma cells. Haematologica, 95(6), pp. 1016-20.

[113]. Bemark, M., Holmqvist, J., Abrahamsson, J. and Mellgren, K. 2012. Translational Mini-Review Series on B cell subsets in disease. Reconstitution after haematopoietic stem cell transplantation - revelation of B cell developmental pathways and lineage phenotypes. Clinical & experimental immunology,167(1):15-25.

[114]. Jackson, N., Ling, N.R., Ball, J., Bromidge, E., Nathan, P.D. and Franklin, I.M. 1989.An analysis of myeloma plasma cell phenotype using antibodies defined at the IIIrd International Workshop on Human Leucocyte Differentiation Antigens. Clinical & Experimental Immunology, 72(3): 351–356.

[115]. Seegmiller, A.C., Xu, Y., McKenna, R.W. and Karandikar, N.J. 2007. Immunophenotypic differentiation between neoplastic plasma cells in mature B-cell lymphoma vs plasma cell myeloma. American Journal of Clinical Pathology, 127(2):176-81.

[116]. Montes-moreno, S., Gonzalez- Medina, A., Rodriguez- Pinilla, S., Maestre, L., Sanchez- Verde, L., Roncador, G., Mollejo, M., García, J. F., Menarguez, J., Montalbán, C., Ruiz- Marcellan, M.N., Conde, E. and Piris, M. A. 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):1342-1349.

[117]. Tierney, R. J., Kirby, H. E., Nagra, J. K., Desmond, J., Bell, A. I. and Rickinson, A. B. 2000. Methylation of Transcription Factor Binding Sites in the Epstein-Barr Virus Latent Cycle Promoter Wp Coincides with Promoter Down-Regulation during Virus-Induced B-Cell Transformation. Journal of Virology, 74(22):10468-10479.

[118] Underhill, G.H., Kolli, K.P. and Kansas, G.S. 2003. Complexity within the plasma cell compartment of mice deficient in both E- and P-selectin: implications for plasma cell differentiation. Blood, 102(12): 4076-4083.

[119] Tam, W., Gomez, M., Chadburn, A., Lee, J.W., Chan, W.C. and Knowles, D.M. 2006. Mutational analysis of PRDM1 indicates a tumor-suppressor role in diffuse large B-cell lymphomas. Blood, 107(10): 4090-4100.

[120] Klein, B., Tarte, K., Jourdan, M., Mathouk, K., Moreaux, J., Jourdan, E., Legouffe, E., Vos, J. and Rossic, J. 2003. Survival and Proliferation Factors of Normal and Malignant Plasma Cells. International Journal of Hematology, 78(2): 106-113.

[121] Klein, B., Zhang, X.G., Jourdan, M., Content, J., Houssiau, F., Aarden, L., Piechaczyk, M. and Bataille, R. 1989. Paracrine rather than autocrine regulation of myeloma-cell growth and differentiation by interleukin-6. Blood, 73(2): 517.

[122] Arora, T. 1998. Differential Myeloma Cell Responsiveness to Interferon-alpha Correlates with Differential Induction of p19INK4d and Cyclin D2 Expression. Journal of Biological Chemistry, 273(19):11799-11805

[123] Uddin, S., Yenush, L., Sun, X.J., Sweet, M.E., White, M.F., and Platanias, L.C. 1995. Interferon-alpha engages the insulin receptor substrate-1 to associate with the phosphatidylinositol 3'-kinase. The Journal of Biological Chemistry, 270(27):15938-41.

[124] Wei, S., Wang, M.W., Teitelbaum, S.L. and Ross, F.P., 2002. Interleukin-4 Reversibly Inhibits Osteoclastogenesis via Inhibition of NF- B and Mitogen-activated Protein Kinase Signaling. Journal of Biological Chemistry, 277(8): 6622-6630.

[125] Masoud, G.N. and Li, W., 2015. HIF-1α pathway: role, regulation and intervention for cancer therapy. Acta pharmaceutica Sinica. B, 5(5): 378.

[126] Kim, J. K., Gao, P., Liu, Y., Semenza, G. L. and Dang, C. V. 2007. Hypoxia-Inducible Factor 1 and Dysregulated c-Myc Cooperatively Induce Vascular Endothelial Growth Factor and Metabolic Switches Hexokinase 2 and Pyruvate Dehydrogenase Kinase 1. Molecular and Cellular Biology, 27(21): 7381-7393.

[127] Fukuda, R., Hirota, K., Fan, F., Jung, Y., Ellis, L.M. and Semenza, G. L. 2002. Insulin-like Growth Factor 1 Induces Hypoxia-inducible Factor 1-mediated Vascular Endothelial Growth Factor Expression, which is Dependent on MAP Kinase and Phosphatidylinositol 3-Kinase Signaling in Colon Cancer Cells. Journal of Biological Chemistry, 277(41): 38205-38211.

[128] Baecklund, E., Sundström, C., Ekbom, A., Catrina, A.I., Biberfeld, P., Feltelius, N. and Klareskog, L. 2003. Lymphoma subtypes in patients with rheumatoid arthritis: Increased proportion of diffuse large B cell lymphoma. Arthritis & Rheumatism, 48(6):1543-1550.

[129] Baecklund, E., Backlin, C., Iliadou, A., Granath, F., Ekbom, A., Amini, R., Feltelius, N., Enblad, G., Sundstrom, C., Klareskog, L., Askling, J. and Rosenquist, R. 2006. Characteristics of diffuse large B cell lymphomas in rheumatoid arthritis. Arthritis & Rheumatism, 54(12): 3774-3781.

[130] Kuze, T., Nakamura, N., Hashimoto, Y., Sasaki, Y. and Abe, M. 2000. The Characteristics of Epstein-Barr Virus (EBV)-positive Diffuse Large B-Cell Lymphoma: Comparison between EBV+ and EBV- Cases in Japanese Population. Cancer Science, 91(12): 1233-1240

[131] Löfström, B., Backlin, C., Sundström, C., Ekbom, A. and Lundberg I.E. 2007A closer look at non-Hodgkin's lymphoma cases in a national Swedish systemic lupus erythematosus cohort: a nested case-control study. Annals of the Rheumatic Diseases, 66(12):1627-32.

[132] Abenavoli, L., Pennacchia, I., Stigliano, E., Carbone, A., Vecchio, F.M. and Arena, V., 2011. Aggressive Large B-Cell Lymphoma in a Systemic Lupus Erythematosus Patient with Chronic Active Epstein-Barr Virus Infection: A Case Report. International Journal of Immunopathology and Pharmacology, 24(4):1083-1086.

[133] Arumugakani, G., Stephenson, S.J., Newton, D.J., Rawstron, A., Emery, P., Doody, G.M., Mcgonagle, D. and Tooze, R.M., 2017. Early Emergence of CD19-Negative Human Antibody-Secreting Cells at the Plasmablast to Plasma Cell Transition. Journal of Immunology, 198(12): 4618-4628.

[134]. Du, X., Beers, R., Fitzgerald, D.J. and Pastan, I., 2008. Differential Cellular Internalization of Anti-CD19 and -CD22 Immunotoxins Results in Different Cytotoxic Activity. Cancer Research, 68(15): 6300-6305.

[135]. Clavarino, G., Delouche, N., Vettier, C., Laurin, D., Pernollet, M., Raskovalova, T., Cesbron, J., Dumestre-Pérard, C. and Jacob, M.C. 2016. Novel Strategy for Phenotypic Characterization of Human B Lymphocytes from Precursors to Effector Cells by Flow Cytometry. PLoS One, 11(9): e0162209.

[136]. Mckenna, R.W., Washington, L.T., Aquino, D.B., Picker, L.J. and Kroft, S.H. 2001. Immunophenotypic analysis of hematogones (B-lymphocyte precursors) in 662 consecutive bone marrow specimens by 4-color flow cytometry. Blood, 98(8): 2498-2507.

[137]. Van Lochem, E.G., Van Der Velden, V.H.J., Wind, H.K., Te Marvelde, J.G., Westerdaal, N.A.C. and Van Dongen, J.J.M. 2004. Immunophenotypic differentiation patterns of normal hematopoiesis in human bone marrow: Reference patterns for age-related changes and disease-induced shifts. Cytometry Part B: Clinical Cytometry, 60B(1): 1-13.

[138]. Marjanovic, N.D., Weinberg, R.A. and Chaffer, C.L. 2013. Cell plasticity and heterogeneity in cancer. Clinical chemistry, 59(1): 168.

[139]. Cobaleda, C., Schebesta, A., Delogu, A. and Busslinger, M. 2007. Pax5: the guardian of B cell identity and function. Nature Immunology, 8(5):463-70.

[140]. Jacob, J. and Kelsoe, G.1992. In situ studies of the primary immune response to (4-hydroxy-3-nitrophenyl)acetyl. II. A common clonal origin for periarteriolar lymphoid sheath-associated foci and germinal centers. Journal of Experimental Medicine, 176(3): 679-687

[141]. Inui, M., Hirota, S., Hirano, K., Fujii, H., Sugahara-Tobinai, A., Ishii, T., Harigae, H. and Takai, T. 2015. Human CD43+ B cells are closely related not only to memory B cells phenotypically but also to plasmablasts developmentally in healthy individuals. International immunology, 27(7): 345-355.

[142]. Comerford, K.M., Wallace, T.J., Karhausen, J., Louis, N.A., Montalto, M.C. and Colgan, S.P. 2002. Hypoxia-inducible factor-1-dependent regulation of the multidrug resistance (MDR1) gene. Cancer Research, 62(12):3387-94.

[143]. Bubien, J.K., Zhou, L.J., Bell, P.D., Frizzell, R.A., Tedder, T.F.1993. Transfection of the CD20 cell surface molecule into ectopic cell types generates a Ca2+ conductance found constitutively in B lymphocytes. Journal of Cell Biology,121(5):1121-32.

[144]. Downes, C.S., Clarke, D.J., Mullinger, A.M., Giménez-Abián, J.F., Creighton, A.M. and Johnson, R.T. 1994. A topoisomerase II-dependent G2 cycle checkpoint in mammalian cells. Nature, 372(6505):467-70.

[145]. Dalla-Favera, R. and Phan, R.T., 2004. The BCL6 proto-oncogene suppresses p53 expression in germinal-centre B cells. Nature, 432(7017): 635-639.

[146]. Wang, S., Konorev, E.A., Kotamraju, S., Joseph, J., Kalivendi, S. and Kalyanaraman, B. 2004. Doxorubicin Induces Apoptosis in Normal and Tumor Cells via Distinctly Different Mechanisms: Intermediacy of H sub(2)O sub(2)- and p53-dependent pathways. Journal of Biological Chemistry, 279(24): 25535-25543

Appendix

Appendix 1- Supplementary figures

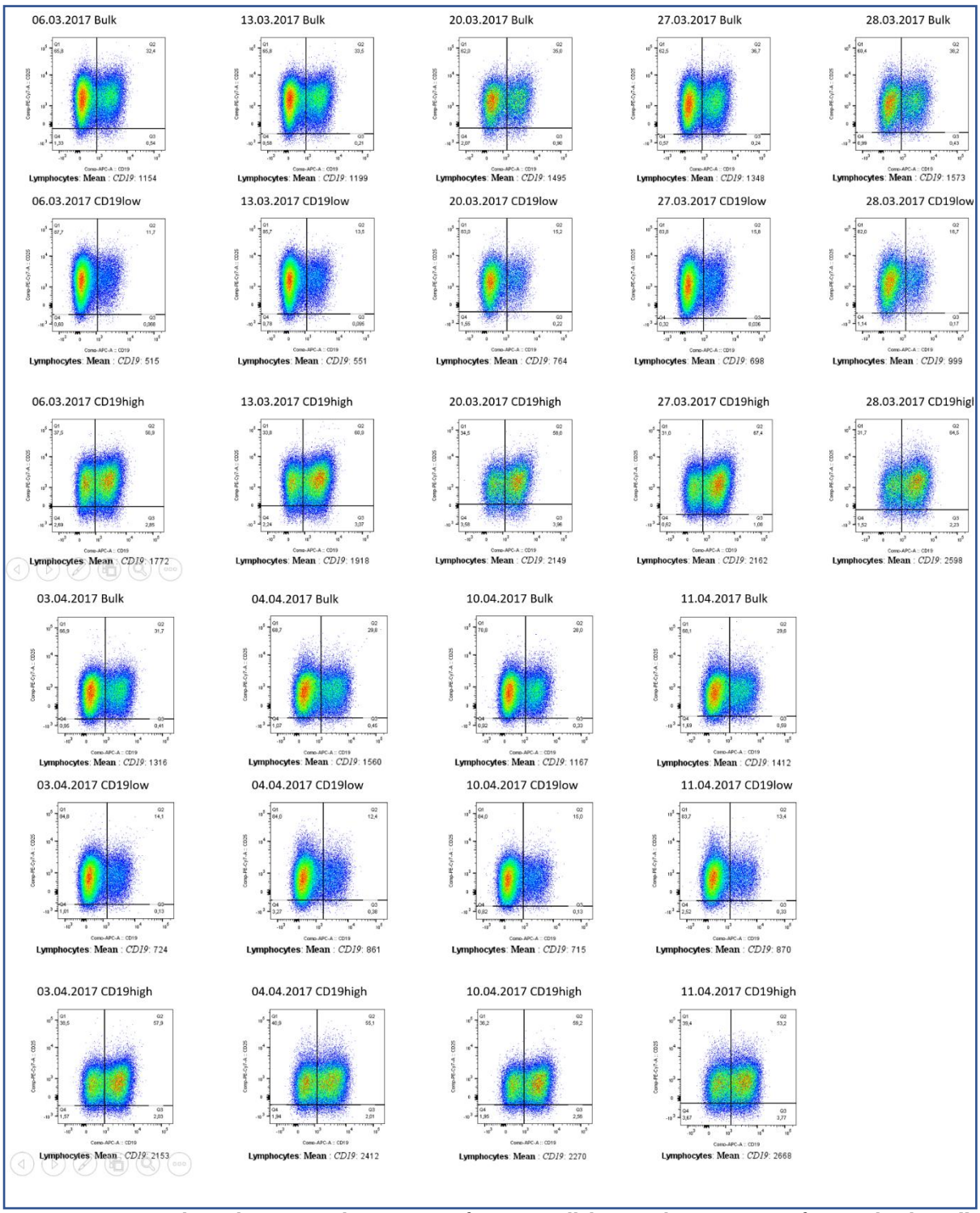

Figure 1: CD19 subpopulations resulting FACS of U2932 cell line at the experiment first trial. The cells maintained a CD19 heterogenous expression for 29 days, which characterizes a process of B cell differentiation or the internalization of the surface marker.

Cell line:	**U-2932**
DSMZ no.:	**ACC 633**
Species:	human (*Homo sapiens*)
Cell type:	B cell lymphoma
Origin:	established in 1996 from the ascites of a 29-year-old woman with diffuse large B cell lymphoma, who 16 years earlier was diagnosed with advanced stage Hodgkin lymphoma, and relapsed several times after multiple chemo- and radiotharapy regimens to complete remissions; cells were described in the literature to overexpress BCL2, BCL6 and p53; assigned to ABC-like lymphoma subtype (activated B-cell)
Reference(s):	14356 , 17355 , 18133 , 18134 , 18135
Biosafety level:	1
Permissions and restrictions:	A
	DSMZ Cell Culture Data:
Morphology:	small round cells growing singly and in clusters in suspension; image
Medium:	90% RPMI 1640 + 10% h.i. FBS
Subculture:	seed out at ca. 0.5 x 10^6 cells/ml; optimal split ratio 1:4 twice a week, maintain at 1-1.5 x 10^6 cells/ml
Incubation:	at 37 °C with 5% CO_2
Doubling time:	ca. 50 hours
Harvest:	cell harvest of ca. 2 x 10^6 cells/ml
Storage:	frozen with 70% medium, 20% FBS, 10% DMSO
	DSMZ Scientific Data:
Mycoplasma:	negative in microbiological culture, PCR assays
Immunology:	CD3 -, CD10 +, CD13 -, CD19 +, CD20 +, CD37 +, CD38 +, cyCD79a +, CD80 +, CD138 +, HLA-DR +, sm/cyIgG -, sm/cyIgM +, sm/cykappa +, sm/cylambda -; image
Fingerprint:	fluorescent nonaplex PCR of short tandem repeat markers revealed a unique DNA profile
Species:	confirmed as human by species PCR
Cytogenetics:	human polyclonal hypodiploid karyotype with 20% polyploidy; 45(43-46)<2n>, XX, add(X)(q22), add(1)(q24), der(3)ins(3;18)(q27;q2?q2?)hsr(18)(q21), add(5)(q32), der(6)t(6;18)(p24;?q23)del(6)(q13), der(10)t(10;14)(q24;q23), der(11)t(1;11)(q25;q2?), der(14)t(3;14)(q27;p11), der(18)t(1;18)(q21;q21), der(18)t(3;18)(q2?;q27), add(19)(p13), del(19)(q13); sdl with del(2)(q11), t(4;15)(q22;q14), add(7)(q21), del(13)(?q21), der(13)add(13)(p11)add(13)(q32); extensive genomic amplification of BCL2 region; resembles published karyotype
Viruses:	PCR: EBV -, HBV -, HCV -, HIV-1 -, HIV-2 -, HTLV-I/II -, MLV -, SMRV -

Supplied as:	Delivery form	Prices
	Frozen culture	360,- €
	Growing culture (please inquire for exact delivery time)	720,- €
	DNA isolated from cell line (25 µg)	440,- €
	DNA isolated from cell line (5 µg)	100,- €

Figure 2: Scientific data about U2932 cell line provided by DSMZ biobank

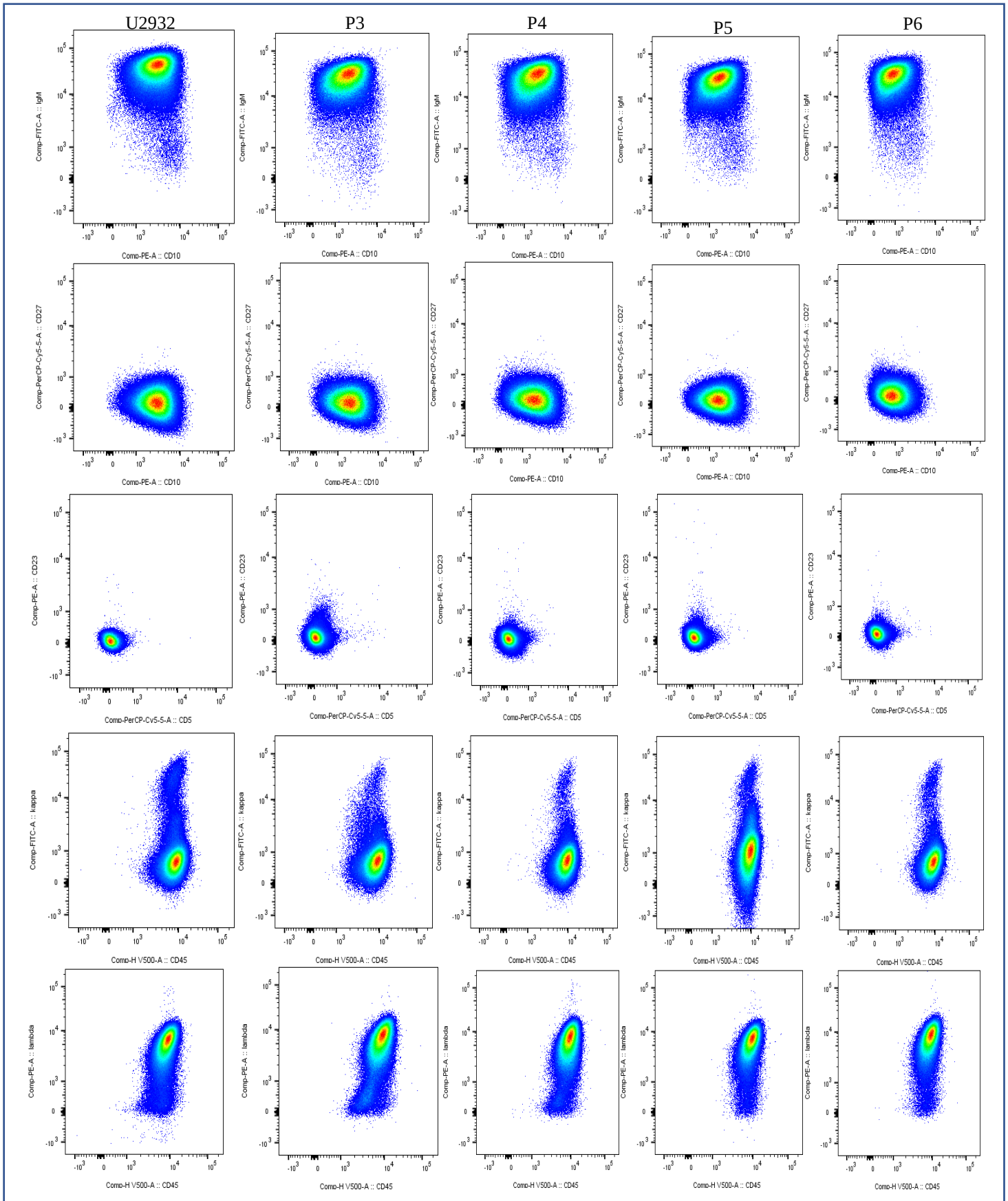

Figure 3: Major immunophenotype characterization of U2932 and subpopulations at day 107 after sorting. IgM, CD10, CD27, kappa, lambda, CD45, CD5, CD23 were used as markers.

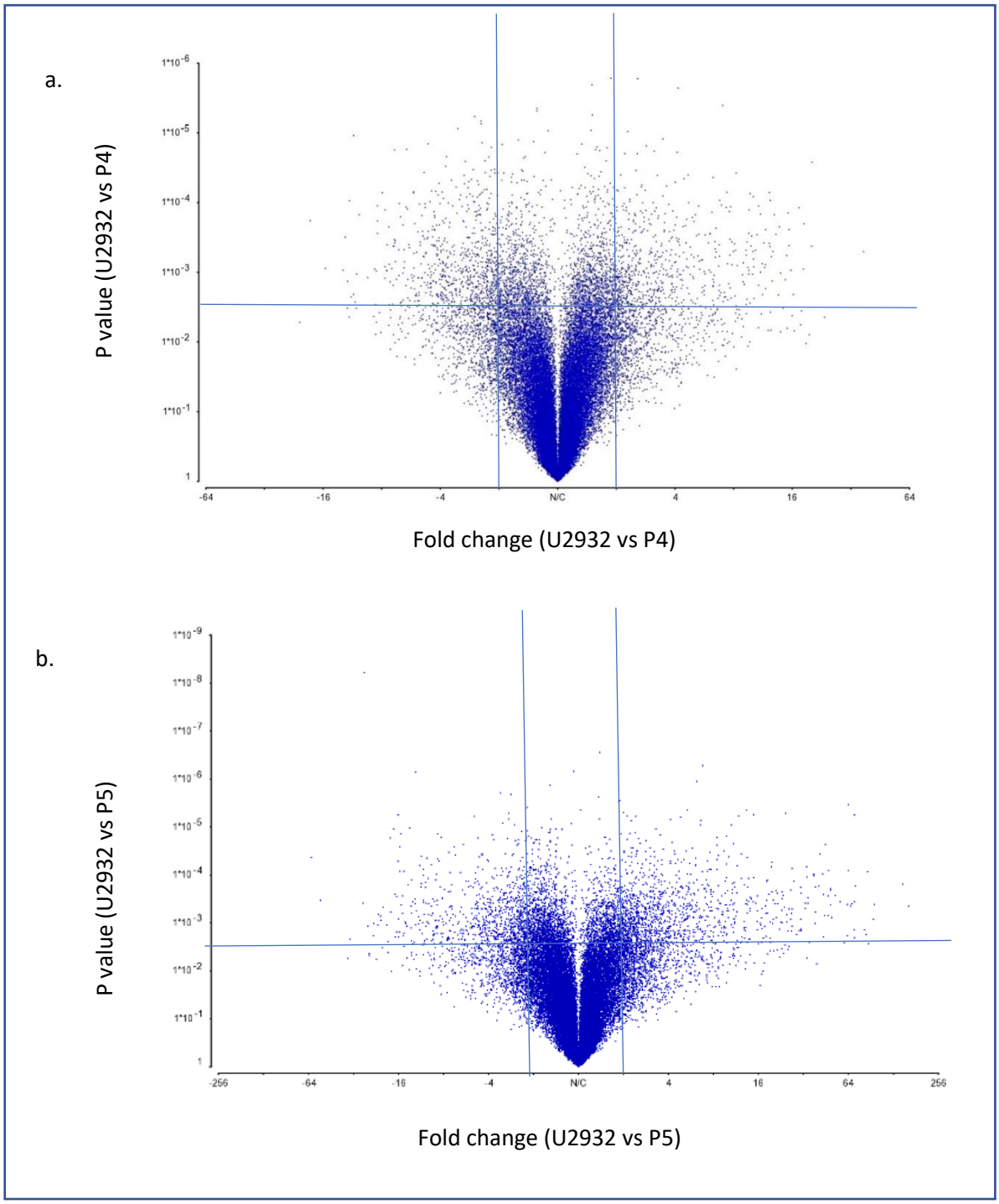

Figure 4: Genes differently expressed between U2932 cell line and its subpopulation. The genes with increased expression localizes on the right side of the N/C (no change) line; genes with reduced expression localizes on the left. The genes that have larger and more significant changes between U2932 and subpopulations are on the upper right and upper left corner. The blue lines represent the specified criteria of fold change > 2 OR change < -2 and p value with FDR < 0.04. (a) Genes differently expressed between U2932 and P4, with 712 genes passing the specified criteria. (b) Genes differently expressed between P4 and P5, with 2.744 genes passing the specified criteria.

Appendix 2- Supplementary Tables

Table 1: Percentage of cells presenting CD markers high and low expression for each subpopulation and U2932 cell line per day of flow cytometry analyses. The sum of each CD marker high and low expression corresponds approximately 100% of the cells composing each population. Data obtained from SSC-A/ CD marker analysis on FLOWJO. P3 had not reached the number of 12 e10^6 cells at day 12, consequently not being submitted to flow cytometry analyses at that date.

Day 0	CD19 high	CD19 low	CD20 high	CD20 low	CD38 high	CD38 low
U2932	**78.0**	22	18.3	**81.3**	47.8	**51.7**
P3	**64.5**	33.8	0.61	**99.4**	**99.9**	0.076
P4	**85.7**	14.2	**96.6**	2.05	**98.9**	1.12
P5	**51.8**	47.4	1.35	**98.6**	0.97	**99**
P6	**70.2**	28	1.28	**98.4**	**93.5**	6.0
Day 12	**CD19 high**	**CD19 low**	**CD20 high**	**CD20 low**	**CD38 high**	**CD38 low**
U2932	**72.4**	26.7	3.56	**96.4**	38.8	**61**
P4	**85.3**	14.3	**78.8**	23.2	**96.4**	3.56
P5	**69.8**	28.8	0.21	**99.7**	13.8	**85.9**
P6	**77.1**	22.2	0.26	**99.7**	**92.1**	7.67
Day 26	**CD19 high**	**CD19 low**	**CD20 high**	**CD20 low**	**CD38 high**	**CD38 low**
U2932	**75**	24.1	2.27	**97.6**	46.2	**53.7**
P3	**82.3**	17.4	1.39	**98.6**	**93.3**	6.62
P4	**79.7**	20.3	**77.7**	21.7	**92.3**	7.63
P5	**79.7**	19.5	0.13	**99.8**	34.7	**65.1**
P6	**81.0**	19	0.19	**99.8**	**95**	5.06
Day 36	**CD19 high**	**CD19 low**	**CD20 high**	**CD20 low**	**CD38 high**	**CD38 low**
U2932	**73.9**	26.1	6.4	**93.5**	**52**	47
P3	**81.5**	18.5	14.5	**85.4**	**82.9**	17.1
P4	**73.2**	26.8	**75.2**	24.6	**93.6**	6.36
P5	**79**	21	0.7	**99.3**	45	**54.9**
P6	**77.1**	22.9	1.02	**98.9**	**92.7**	7.25
Day 93	**CD19 high**	**CD19 low**	**CD20 high**	**CD20 low**	**CD38 high**	**CD38 low**
U2932	**72.2**	27.4	0.27	**99.7**	**79.7**	19.9
P3	**63.7**	35.8	8.21	**91.7**	**50.6**	49.4
P4	**62.5**	37	11.4	**88.5**	43.8	**55.6**
P5	**59.1**	40.4	0.47	**99.5**	49	**50.7**
P6	**69**	31	3.4	**96.9**	45.6	**54**
Day 100	**CD19 high**	**CD19 low**	**CD20 high**	**CD20 low**	**CD38 high**	**CD38 low**
U2932	**87.5**	11.9	0.57	**99.3**	**83.1**	16
P3	**73.1**	26	5.22	**94.7**	**54.2**	45.6
P4	**74**	25.1	9.66	**90.2**	**60.5**	39.3
P5	**72.1**	27	0.42	**99.4**	**51.7**	48
P6	**73.5**	25.6	2.14	**97.8**	**62**	37.6

Table 2: One- way ANOVA statistics analysis of U2932 and subpopulations on exponential phase of growth curve

Tukey's multiple comparisons test DAY1	**Adjusted P Value**
U2932 vs. P3	0,9993
U2932 vs. P4	0,0032
U2932 vs. P5	0,6557
U2932 vs. P6	0,9707
P3 vs. P4	0,0044
P3 vs. P5	0,5301
P3 vs. P6	0,9149
P4 vs. P5	0,0005
P4 vs. P6	0,0014
P5 vs. P6	0,9345
One- way ANOVA	0,0005
Tukey's multiple comparisons test DAY2	**Adjusted P Value**
U2932 vs. P3	0,2555
U2932 vs. P4	<0,0001
U2932 vs. P5	0,0805
U2932 vs. P6	0,0001
P3 vs. P4	<0,0001
P3 vs. P5	0,9290
P3 vs. P6	0,0016
P4 vs. P5	0,0002
P4 vs. P6	0,1418
P5 vs. P6	0,0047
Tukey's multiple comparisons test DAY3	**Adjusted P Value**
U2932 vs. P3	0,9026
U2932 vs. P4	0,0004
U2932 vs. P5	0,5598
U2932 vs. P6	0,0220
P3 vs. P4	0,0012
P3 vs. P5	0,9570
P3 vs. P6	0,0843
P4 vs. P5	0,0030
P4 vs. P6	0,0942
P5 vs. P6	0,2293
Tukey's multiple comparisons test DAY4	**Adjusted P Value**
U2932 vs. P3	0,8362
U2932 vs. P4	0,0788
U2932 vs. P5	>0,9999
U2932 vs. P6	0,4972
P3 vs. P4	0,3457
P3 vs. P5	0,8523
P3 vs. P6	0,9669
P4 vs. P5	0,0832
P4 vs. P6	0,6755
P5 vs. P6	0,5167
One-way ANOVA	0,0656
Tukey's multiple comparisons test DAY5	**Adjusted P Value**
U2932 vs. P3	0,8868
U2932 vs. P4	0,3968
U2932 vs. P5	0,9996
U2932 vs. P6	0,9584
P3 vs. P4	0,8726
P3 vs. P5	0,7987
P3 vs. P6	0,9991
P4 vs. P5	0,3091
P4 vs. P6	0,7599
P5 vs. P6	0,9016
One-way ANOVA	0,3235

Table 3: Statistics analysis of drug screen data per dose at 36 days after sorting.

Tukey's multiple comparisons test C0	Adjusted P Value - 0h		Adjusted P Value - 0h	
U2932 vs. P3	0,9105	>0,9999	0,9947	0,9942
U2932 vs. P4	0,7520	>0,9999	0,9970	0,9967
U2932 vs. P5	0,9666	0,9988	0,9393	0,9976
P3 vs. P4	0,9846	0,9998	0,9699	0,9673
P3 vs. P5	0,9969	0,9977	0,8495	0,9715
P4 vs. P5	0,9446	0,9996	0,9818	>0,9999
One-way ANOVA	0, 7898	0,9980	0,8709	0,9647
Tukey's multiple comparisons test C1	**Adjusted P Value - 0h**		**Adjusted P Value - 48h**	
U2932 vs. P3	0,9131	0,9996	0,5674	0,5761
U2932 vs. P4	0,9665	0,9772	0,1491	0,1551
U2932 vs. P5	0,9253	>0,9999	0,3240	0,9498
P3 vs. P4	0,6966	0,9902	0,7058	0,7128
P3 vs. P5	>0,9999	0,9996	0,9564	0,8559
P4 vs. P5	0,7168	0,9770	0,9334	0,3122
One- way ANOVA	0,6554	0,9738	0,1736	0,1690
Tukey's multiple comparisons test C2	**Adjusted P Value - 0h**		**Adjusted P Value - 48h**	
U2932 vs. P3	0,7209	0,9946	0,5990	0,9045
U2932 vs. P4	0,3957	0,9615	0,0990	0,0906
U2932 vs. P5	0,8033	0,5261	0,1314	0,1372
P3 vs. P4	0,0969	0,9949	0,5122	0,2290
P3 vs. P5	0,9985	0,4038	0,0219	0,0532
P4 vs. P5	0,1208	0,3029	0,0035	0,0034
One- way ANOVA	0,0866	0,3008	0,0046	0,0053
Tukey's multiple comparisons test C3	**Adjusted P Value - 0h**		**Adjusted P Value - 48h**	
U2932 vs. P3	0,8860	0,9888	0,0399	0,0561
U2932 vs. P4	0,4998	0,9801	0,0993	0,1306
U2932 vs. P5	0,9458	0,8318	0,3170	0,0830
P3 vs. P4	0,2074	0,9999	0,9138	0,9292
P3 vs. P5	0,9978	0,6680	0,0036	0,0015
P4 vs. P5	0,2622	0,6297	0,0079	0,0030
One- way ANOVA	0,1983	0,6142	0,0026	0,0012

Table 4: Statistics analysis of drug screen data per population at 36 days after sorting.

Tukey's multiple comparisons test U2932	Adjusted P Value- 0h		Adjusted P Value- 48h	
C0 vs. C1	0,3553	0,9998	0,0009	0,0026
C0 vs. C2	0,1652	0,9995	<0,0001	<0,0001
C0 vs. C3	0,3075	0,9048	<0,0001	0,0001
C1 vs. C2	0,9330	>0,9999	0,0235	0,0308
C1 vs. C3	0,9994	0,9301	0,0283	0,0616
C2 vs. C3	0,9636	0,9412	0,9991	0,9583
One- way ANOVA	0,1753	0,9001	<0,0001	<0,0001
Tukey's multiple comparisons test P3	**Adjusted P Value- 0h**		**Adjusted P Value- 48h**	
C0 vs. C1	0,9992	>0,9999	0,0277	0,0287
C0 vs. C2	>0,9999	0,9898	<0,0001	<0,0001
C0 vs. C3	0,9926	0,9923	<0,0001	<0,0001
C1 vs. C2	0,9986	0,9898	0,0023	0,0032
C1 vs. C3	0,9989	0,9923	0,0007	0,0005
C2 vs. C3	0,9905	>0,9999	0,7205	0,7505
One- way ANOVA	0,9899	0,9817	<0,0001	<0,0001
Tukey's multiple comparisons test P4	**Adjusted P Value- 0h**		**Adjusted P Value- 48h**	
C0 vs. C1	>0,9999	0,9844	0,1324	0,1369
C0 vs. C2	0,9938	0,9595	<0,0001	<0,0001
C0 vs. C3	>0,9999	0,9980	<0,0001	<0,0001
C1 vs. C2	0,9938	0,8357	<0,0001	<0,0001
C1 vs. C3	>0,9999	0,9511	<0,0001	<0,0001
C2 vs. C3	0,9938	0,9885	0,9121	0,3153
One- way ANOVA	0,0240	08634	<0,0001	<0,0001
Tukey's multiple comparisons test P5	**Adjusted P Value- 0h**		**Adjusted P Value- 48h**	
C0 vs. C1	0,6371	0,9999	0,1761	0,0187
C0 vs. C2	0,0680	0,1796	0,0027	0,0033
C0 vs. C3	0,0274	0,1025	0,0013	0,0033
C1 vs. C2	0,3510	0,1968	0,0539	0,5523
C1 vs. C3	0,1495	0,0932	0,0212	0,5489
C2 vs. C3	0,9126	0,0047	0,9048	>0,9999
One-way ANOVA	0,9916	0,0077	0,0009	0,0022

Table 5: Statistics analysis of drug screen data per dose at 100 days after sorting.

Tukey's multiple comparisons test C0	Adjusted P Value - 0h		Adjusted P Value - 48h	
U2932 vs. P3	0,9798	0,8903	0,9906	0,8875
U2932 vs. P4	0,8405	0,9996	0,6433	0,9698
U2932 vs. P5	0,9849	>0,9999	0,8470	0,9973
P3 vs. P4	0,6391	0,8486	0,8016	0,9919
P3 vs. P5	0,8840	0,9077	0,9520	0,8014
P4 vs. P5	0,9610	0,9988	0,9788	0,9193
One- way ANOVA	0,6756	0,8439	0,6491	0,7969
Tukey's multiple comparisons test C1	**Adjusted P Value - 0h**		**Adjusted P Value - 48h**	
U2932 vs. P3	0,0799	0,5761	0,0424	0,9551
U2932 vs. P4	0,6192	0,1551	0,2686	0,2086
U2932 vs. P5	0,1532	0,9498	0,0206	0,3086
P3 vs. P4	0,4172	0,7128	0,5694	0,3959
P3 vs. P5	0,9648	0,8559	0,9510	0,5489
P4 vs. P5	0,6648	0,3122	0,3168	0,9898
One- way ANOVA	0,0601	0,0787	0,0203	0,1638
Tukey's multiple comparisons test C2	**Adjusted P Value - 0h**		**Adjusted P Value - 48h**	
U2932 vs. P3	0,6844	0,9213	>0,9999	0,5068
U2932 vs. P4	0,9900	0,8949	0,9186	0,9831
U2932 vs. P5	0,9663	0,9995	0,9980	0,6868
P3 vs. P4	0,8398	0,5751	0,9115	0,3367
P3 vs. P5	0,9056	0,8825	0,9986	0,1216
P4 vs. P5	0,9985	0,9315	0,8503	0,8682
One- way ANOVA	0,7212	0,6403	0,8559	0,1512
Tukey's multiple comparisons test C3	**Adjusted P Value - 0h**		**Adjusted P Value - 48h**	
U2932 vs. P3	0,4189	0,9951	0,2452	0,7661
U2932 vs. P4	0,7953	0,9305	0,7748	0,9827
U2932 vs. P5	0,9712	0,5093	0,9787	0,9997
P3 vs. P4	0,1085	0,9826	0,0649	0,5667
P3 vs. P5	0,6638	0,6380	0,1450	0,8141
P4 vs. P5	0,5486	0,8305	0,9378	0,9666
One- way ANOVA	0,1441	0,5271	0,0715	0,6075

Table 6: Statistics analysis of drug screen data per population at 100 days after sorting.

Tukey's multiple comparisons test U2932	Adjusted P Value- 0h		Adjusted P Value- 48h	
C0 vs. C1	0,2344	0,9603	0,8842	0,0009
C0 vs. C2	0,8775	0,9120	0,0773	<0,0001
C0 vs. C3	0,9985	0,8082	0,0020	<0,0001
C1 vs. C2	0,5605	0,9982	0,0279	0,0235
C1 vs. C3	0,2874	0,5441	0,0009	0,0283
C2 vs. C3	0,9339	0,4566	0,0844	0,9991
One- way ANOVA	0,2235	0,4558	0,0007	<0,0001
Tukey's multiple comparisons test P3	**Adjusted P Value- 0h**		**Adjusted P Value- 48h**	
C0 vs. C1	0,2601	0,9869	0,0084	0,9599
C0 vs. C2	0,9969	0,9652	0,0002	0,0006
C0 vs. C3	0,2601	>0,9999	<0,0001	0,0001
C1 vs. C2	0,3353	0,8570	0,0297	0,0010
C1 vs. C3	>0,9999	0,9859	<0,0001	0,0002
C2 vs. C3	0,3353	0,9670	<0,0001	0,3504
One-way ANOVA	0,1384	0,8861	<0,0001	<0,0001
Tukey's multiple comparisons test P4	**Adjusted P Value- 0h**		**Adjusted P Value- 48h**	
C0 vs. C1	>0,9999	0,9472	0,1094	0,0670
C0 vs. C2	0,9696	0,7443	0,0013	0,0020
C0 vs. C3	0,6416	0,4873	<0,0001	<0,0001
C1 vs. C2	0,9728	0,4515	0,0347	0,1015
C1 vs. C3	0,6511	0,7812	0,0002	<0,0001
C2 vs. C3	0,8688	0,1337	0,0066	0,0011
One- way ANOVA	0,6128	0,1673	<0,0001	<0,0001
Tukey's multiple comparisons test P5	**Adjusted P Value- 0h**		**Adjusted P Value- 48h**	
C0 vs. C1	0,9153	0,8079	<0,0001	0,7900
C0 vs. C2	0,9978	0,9553	<0,0001	0,1738
C0 vs. C3	0,3094	0,2293	<0,0001	0,0009
C1 vs. C2	0,9652	0,5322	0,0110	0,5420
C1 vs. C3	0,6250	0,6364	<0,0001	0,0026
C2 vs. C3	0,3850	0,1146	<0,0001	0,0145
One- way ANOVA	0,3045	0,1221	<0,0001	0,0010

Table 7: Probe set ID detailed information for transcription factors involved in the B-cell differentiation.

Probe set ID	Gene	P4 Expression	P4 Expression	P5 Expression	P5 Expression	U2932 Expression	U2932 Expression
202431_s_at	MYC	8.79058	8.70003	11.7432	11.824	11.6466	11.6342
215990_s_at	BCL6	10.3175	10.1194	5.95171	5.46212	7.69441	7.71468
228758_at	BCL6	10.7015	10.6375	6.25535	6.0192	8.05139	8.02525
239249_at	BCL6	4.98997	4.96105	4.21759	4.68644	4.40283	4.54493
203140_at	BCL6	12.5025	12.4725	7.52697	7.44884	10.3884	10.258
203684_s_at	BCL2	9.29813	9.43599	9.06405	9.05562	9.21522	9.17532
203685_at	BCL2	12.7339	12.8472	12.6589	12.7251	12.7436	12.6783
207004_at	BCL2	8.52491	8.35923	8.92314	9.01385	9.104	9.13919
207005_s_at	BCL2	8.42926	8.23835	8.7364	8.7514	8.68078	8.73765
221969_at	PAX5	10.4296	10.4446	11.1556	11.111	10.9682	11.0445
206802_at	PAX5	5.58185	5.66883	5.65974	5.57178	5.78553	5.4853
204562_at	IRF4	10.8446	10.825	10.8485	10.8145	10.7538	10.7656
217192_s_at	PRDM1	6.7084	6.67705	6.20272	6.22807	6.38201	6.24137
228964_at	PRDM1	5.37991	5.5299	8.31358	8.34277	6.35981	6.26218
235668_at	PRDM1	5.34262	4.9002	5.93119	5.97625	5.69462	5.51823
206398_s_at	CD19	8.66.75975	8.68091	9.75605	9.66079	9.6686	9.5816
205692_s_at	CD38	9.85441	9.78082	7.62598	7.70586	8.253	8.30786
207792_s_at	XBP1	3.4553	3.32953	3.74683	3.70056	3.79102	3.77461

Table 8: One- way ANOVA and Turkey multiple comparison test of expression of genes that regulates the B-cell differentiation.

Tukey's multiple comparisons test MYC	Adjusted P value
U2932 vs. P4	< 0.0001
U2932 vs. P5	0.1242
P4 vs. P5	< 0.0001
One-way ANOVA	< 0.0001
Tukey's multiple comparisons test BCL6	**Adjusted P value**
U2932 vs. P4	0.0011
U2932 vs. P5	0.0024
P4 vs. P5	0.0001
One- way ANOVA	< 0.0001
Tukey's multiple comparisons test IRF4	**Adjusted P value**
U2932 vs. P4	0.0415
U2932 vs. P5	0.0467
P4 vs. P5	0.9789
One- way ANOVA	0.0337
Tukey's multiple comparisons PAX5	**Adjusted P value**
U2932 vs. P4	0.0298
U2932 vs. P5	0.6562
P4 vs. P5	0.0187
One- way ANOVA	0.0173
Tukey's multiple comparisons PRDM1	**Adjusted P value**
U2932 vs. P4	0.032
U2932 vs. P5	0.032
P4 vs. P5	0.021
One- way ANOVA	0.0015
Tukey's multiple comparisons XBP1	**Adjusted P value**
U2932 vs. P4	0,032
U2932 vs. P5	0,032
P4 vs. P5	>0.05
One- way ANOVA	0.0108
Tukey's multiple comparisons BCL2	**Adjusted P value**
U2932 vs. P4	0.0016
U2932 vs. P5	0.008
P4 vs. P5	0.0174
One- way ANOVA	0.0017

Table 9: Pathways resulting from U2932 vs P4 enrichment

Pathway name	Enrichment score	%genes in pathway that are on list	Genes in list in pathway	Genes not in list in pathway
Rheumatoid arthritis	11.2548	15.3846	14	77
Apoptosis	7.95095	10.9489	15	122
MAPK signaling pathway	7.33434	8.16327	24	270
Transcriptional misregulation in cancer	7.03758	9.39227	17	164
NF-kappa B signaling pathway	6.78102	11.828	11	82
Osteoclast differentiation	6.42471	10.2362	13	114
B cell receptor signaling pathway	6.24132	12.6761	9	62
Chagas disease (American trypanosomiasis)	6.02631	10.7843	11	91
Systemic lupus erythematosus	5.7557	9.91736	12	109
Cellular senescence	5.35762	8.75	14	146

Table 10: Pathways resulting from P4 vs P5 enrichment

Pathway name	Enrichment score	%genes in pathway that are on list	Genes in list in pathway	Genes not in list in pathway
Epstein-Barr virus infection	9.77619	21.1055	42	157
Rheumatoid arthritis	7.62471	24.1758	22	69
HIF-1 signaling pathway	7.45854	23.4694	23	75
B cell receptor signaling pathway	7.06229	25.3521	18	53
Systemic lupus erythematosus	6.82448	21.4876	26	95
Apoptosis	6.4361	20.438	28	109
Hepatocellular carcinoma	6.3616	19.5122	32	132
Intestinal immune network for IgA production	6.06685	27.0833	13	35
MAPK signaling pathway	5.94571	17.0068	50	244
Choline metabolism in cancer	5.72056	21.4286	21	77

Table 11: Gene list of rheumatoid arthritis as GO function for Pathway enrichment of U2932 vs P4.

Gene symbol	Gene title	Fold change	Fold change (Description)
TNF	tumor necrosis factor	-2.47342	U2932 down vs P4
CCL3 /// CCL3L1 /// CCL3L3	chemokine (C-C motif) ligand 3 /// chemokine (C-C motif) ligand 3-like 1 /// chemokine	-2.48489	U2932 down vs P4
HLA-DQA1	major histocompatibility complex, class II, DQ alpha 1	-4.28254	U2932 down vs P4
ATP6V1C1	ATPase, H+ transporting, lysosomal 42kDa, V1 subunit C1	-2.66935	U2932 down vs P4
TNFRSF11A	tumor necrosis factor receptor superfamily, member 11a, NFKB activator	-12.3155	U2932 down vs P4
ATP6V1D	ATPase, H+ transporting, lysosomal 34kDa, V1 subunit D	-2.29943	U2932 down vs P4
FLT1	fms-related tyrosine kinase 1	4.55925	U2932 up vs P4
TLR4	toll-like receptor 4	3.58539	U2932 up vs P4
ITGAL	integrin alpha L	-4.00114	U2932 down vs P4
JUN	jun proto-oncogene	-2.01923	U2932 down vs P4
FOS	FBJ murine osteosarcoma viral oncogene homolog	-7.09727	U2932 down vs P4
ATP6V1C1	ATPase, H+ transporting, lysosomal 42kDa, V1 subunit C1	-2.67936	U2932 down vs P4
TNF	tumor necrosis factor	-2.96175	U2932 down vs P4
CD86	CD86 molecule	-2.40063	U2932 down vs P4

Table 12: Gene list of apoptosis as GO function for Pathway enrichment of U2932 vs P4

Gene symbol	Gene title	Fold change	Fold change (Description)
TNF	tumor necrosis factor	-2.47342	U2932 down vs P4
BID	BH3 interacting domain death agonist	-2.2905	U2932 down vs P4
MCL1	myeloid cell leukemia 1	-2.31023	U2932 down vs P4
BID	BH3 interacting domain death agonist	-2.60276	U2932 down vs P4
BCL2A1	BCL2-related protein A1	-4.27459	U2932 down vs P4
ITPR2	inositol 1,4,5-trisphosphate receptor, type 2	2.26451	U2932 up vs P4
AKT3	v-akt murine thymoma viral oncogene homolog 3	6.74745	U2932 up vs P4
BIRC3	baculoviral IAP repeat containing 3	-5.06858	U2932 down vs P4
ERN1	endoplasmic reticulum to nucleus signaling 1	3.41988	U2932 up vs P4
GADD45A	growth arrest and DNA-damage-inducible, alpha	2.10758	U2932 up vs P4
CTSO	cathepsin O	-2.2001	U2932 down vs P4
JUN	jun proto-oncogene	-2.01923	U2932 down vs P4
PIDD1	p53-induced death domain protein 1	2.30202	U2932 up vs P4
GADD45B	growth arrest and DNA-damage-inducible, beta	-2.27203	U2932 down vs P4
BID	BH3 interacting domain death agonist	-2.43128	U2932 down vs P4
FOS	FBJ murine osteosarcoma viral oncogene homolog	-7.09727	U2932 down vs P4
CFLAR	CASP8 and FADD like apoptosis regulator	-2.0907	U2932 down vs P4
TNF	tumor necrosis factor	-2.96175	U2932 down vs P4

Table 13: Gene list of MAPK signaling pathway as GO function for Pathway enrichment of U2932 vs P4

Gene symbol	Gene title	Fold change	Fold change (Description)
INSR	insulin receptor	7.03561	U2932 down vs P4
TNF	tumor necrosis factor	-2.47342	U2932 up vs P4
PRKCB	protein kinase C, beta	3.61959	U2932 up vs P4
RRAS2	related RAS viral (r-ras) oncogene homolog 2	-2.93525	U2932 down vs P4
IGF1R	insulin-like growth factor 1 receptor	2.0117	U2932 up vs P4
PRKCB	protein kinase C, beta	3.03184	U2932 up vs P4
NF1	neurofibromin 1	2.49783	U2932 up vs P4
MEF2C	myocyte enhancer factor 2C	-2.58691	U2932 down vs P4
HSPA1A /// HSPA1B	heat shock 70kDa protein 1A /// heat shock 70kDa protein 1B	-4.2851	U2932 down vs P4
PPM1A	protein phosphatase, Mg2+/Mn2+ dependent, 1A	-2.62166	U2932 down vs P4
CDC25B	cell division cycle 25B	4.01424	U2932 up vs P4
FGFR1	fibroblast growth factor receptor 1	-2.14068	U2932 down vs P4
PDGFD	platelet derived growth factor D	-4.94142	U2932 down vs P4
RRAS2	related RAS viral (r-ras) oncogene homolog 2	-3.69022	U2932 down vs P4
FLT1	fms-related tyrosine kinase 1	4.55925	U2932 up vs P4
HSPA1A /// HSPA1B	heat shock 70kDa protein 1A /// heat shock 70kDa protein 1B	-6.86555	U2932 down vs P4

MAP4K4	mitogen-activated protein kinase kinase kinase kinase 4	8.08342	U2932 up vs P4
AKT3	v-akt murine thymoma viral oncogene homolog 3	6.74745	U2932 up vs P4
INSR	insulin receptor	8.17867	U2932 up vs P4
HSPA1A /// HSPA1B	heat shock 70kDa protein 1A /// heat shock 70kDa protein 1B	-5.28935	U2932 down vs P4
TGFBR1	transforming growth factor, beta receptor 1	2.48239	U2932 up vs P4
RRAS2	related RAS viral (r-ras) oncogene homolog 2	-3.53925	U2932 down vs P4
RASGRP3	RAS guanyl releasing protein 3 (calcium and DAG-regulated)	-2.92227	U2932 down vs P4
GADD45A	growth arrest and DNA-damage-inducible, alpha	2.10758	U2932 up vs P4
MYC	v-myc avian myelocytomatosis viral oncogene homolog	7.44272	U2932 up vs P4
JUN	jun proto-oncogene	-2.01923	U2932 down vs P4
GADD45B	growth arrest and DNA-damage-inducible, beta	-2.27203	U2932 down vs P4
FOS	FBJ murine osteosarcoma viral oncogene homolog	-7.09727	U2932 down vs P4
TNF	tumor necrosis factor	-2.96175	U2932 down vs P4
ERBB4	erb-b2 receptor tyrosine kinase 4	13.9331	U2932 up vs P4

Table 14: Gene list of transcriptional misregulation as GO function for Pathway enrichment of U2932 vs P4

Gene symbol	Gene title	Fold change	Fold change (Description)
BCL6	B-cell CLL/lymphoma 6	-5.96679	U2932 down vs P4
REL	v-rel avian reticuloendotheliosis viral oncogene homolog	-2.82305	U2932 down vs P4
IGF1R	insulin-like growth factor 1 receptor	2.0117	U2932 up vs P4
MEF2C	myocyte enhancer factor 2C	-2.58691	U2932 down vs P4
BCL2A1	BCL2-related protein A1	-4.27459	U2932 down vs P4
SS18	synovial sarcoma translocation, chromosome 18	-2.1507	U2932 down vs P4
FLT1	fms-related tyrosine kinase 1	4.55925	U2932 up vs P4
BCL6	B-cell CLL/lymphoma 6	-4.42932	U2932 down vs P4
BIRC3	baculoviral IAP repeat containing 3	-5.06858	U2932 down vs P4
HIST1H2AD /// HIST1H3D	histone cluster 1, H2ad /// histone cluster 1, H3d	-3.99069	U2932 down vs P4
HPGD	hydroxyprostaglandin dehydrogenase 15-(NAD)	10.0572	U2932 up vs P4
SMAD1	SMAD family member 1	37.1744	U2932 up vs P4
SS18	synovial sarcoma translocation, chromosome 18	-2.57624	U2932 down vs P4
ID2 /// ID2B	inhibitor of DNA binding 2, dominant negative helix-loop-helix protein /// inhibitor of	-2.95248	U2932 down vs P4
GADD45A	growth arrest and DNA-damage-inducible, alpha	2.10758	U2932 up vs P4
MYC	v-myc avian myelocytomatosis viral oncogene homolog	7.44272	U2932 up vs P4
JMJD1C	jumonji domain containing 1C	2.61767	U2932 up vs P4
GADD45B	growth arrest and DNA-damage-inducible, beta	-2.27203	U2932 down vs P4
HPGD	hydroxyprostaglandin dehydrogenase 15-(NAD)	9.60002	U2932 up vs P4
CD86	CD86 molecule	-2.40063	U2932 down vs P4

Table 15: Gene list of NF-kappa B signaling pathway GO function for Pathway enrichment of U2932 vs P4

Gene symbol	Gene title	Fold change	Fold change (Description)
TNF	tumor necrosis factor	-2.47342	U2932 down vs P4
PRKCB	protein kinase C, beta	3.61959	U2932 up vs P4
PRKCB	protein kinase C, beta	3.03184	U2932 up vs P4
BCL2A1	BCL2-related protein A1	-4.27459	U2932 down vs P4
SYK	spleen tyrosine kinase	2.63734	U2932 up vs P4
TNFRSF11A	tumor necrosis factor receptor superfamily, member 11a, NFKB activator	-12.3155	U2932 down vs P4
BIRC3	baculoviral IAP repeat containing 3	-5.06858	U2932 down vs P4
LCK	LCK proto-oncogene, Src family tyrosine kinase	-2.17494	U2932 down vs P4
TLR4	toll-like receptor 4	3.58539	U2932 up vs P4
PIDD1	p53-induced death domain protein 1	2.30202	U2932 up vs P4
GADD45B	growth arrest and DNA-damage-inducible, beta	-2.27203	U2932 down vs P4
CFLAR	CASP8 and FADD like apoptosis regulator	-2.0907	U2932 down vs P4
TNF	tumor necrosis factor	-2.96175	U2932 down vs P4

Table 16: Gene list of Epstein-Barr virus infection GO function for Pathway enrichment of P4 vs P5.

Gene symbol	Gene title	Fold change	Fold change (Description)
RB1	retinoblastoma 1	70.1212	P4 up vs P5
MYC	v-myc avian myelocytomatosis viral oncogene homolog	-8.29189	P4 down vs P5
CD44	CD44 molecule (Indian blood group)	-2.03533	P4 down vs P5
EIF2AK4	eukaryotic translation initiation factor 2 alpha kinase 4	3.06462	P4 up vs P5
ITGAL	integrin alpha L	12.3257	P4 up vs P5
PSMC4	proteasome 26S subunit, ATPase 4	-2.90038	P4 down vs P5
TRAC /// TRAJ17 /// TRAV20 /// TRDV2	T-cell receptor alpha constant /// T cell receptor alpha joining 17 /// T cell receptor	4.79386	P4 up vs P5
CD58	CD58 molecule	5.48934	P4 up vs P5
HSPA1A /// HSPA1B	heat shock 70kDa protein 1A /// heat shock 70kDa protein 1B	11.4779	P4 up vs P5
HLA-DQA1 /// HLA-DQA2	major histocompatibility complex, class II, DQ alpha 1 /// major histocompatibility com	2.20304	P4 up vs P5
TP53	tumor protein p53	-2.34251	P4 down vs P5
CD44	CD44 molecule (Indian blood group)	-2.40957	P4 down vs P5
AKT3	v-akt murine thymoma viral oncogene homolog 3	-9.13951	P4 down vs P5
POLR2K	polymerase (RNA) II (DNA directed) polypeptide K, 7.0kDa	2.05861	P4 up vs P5
CD58	CD58 molecule	5.54243	P4 up vs P5
MAP2K4	mitogen-activated protein kinase kinase 4	2.10464	P4 up vs P5
CD40	CD40 molecule, TNF receptor superfamily member 5	3.54652	P4 up vs P5
MAPK9	mitogen-activated protein kinase 9	-2.81213	P4 down vs P5
CD58	CD58 molecule	5.61727	P4 up vs P5
CSNK2A1	casein kinase 2, alpha 1 polypeptide	-2.026	P4 down vs P5
CD38	CD38 molecule	4.48054	P4 up vs P5
ITGAL	integrin alpha L	6.5589	P4 up vs P5
IRAK1	interleukin 1 receptor associated kinase 1	-2.4307	P4 down vs P5
IL10RA	interleukin 10 receptor, alpha	6.71885	P4 up vs P5
HSPA1A /// HSPA1B	heat shock 70kDa protein 1A /// heat shock 70kDa protein 1B	7.08691	P4 up vs P5
HSPA1A /// HSPA1B	heat shock 70kDa protein 1A /// heat shock 70kDa protein 1B	5.10385	P4 up vs P5
IGH	immunoglobulin heavy locus	3.08466	P4 up vs P5
HLA-DQA1	major histocompatibility complex, class II, DQ alpha 1	23.1982	P4 up vs P5
VIM	Vimentin	-14.2627	P4 down vs P5
ICAM1	intercellular adhesion molecule 1	2.07973	P4 up vs P5
AKT3	v-akt murine thymoma viral oncogene homolog 3	-4.63676	P4 down vs P5
GTF2E2	general transcription factor IIE subunit 2	4.30891	P4 up vs P5
POLR3GL	polymerase (RNA) III (DNA directed) polypeptide G (32kD)-like	4.52783	P4 up vs P5
SYK	spleen tyrosine kinase	-2.95498	P4 down vs P5
POLR2I	polymerase (RNA) II (DNA directed) polypeptide I, 14.5kDa	-2.19236	P4 down vs P5
AKT3	v-akt murine thymoma viral oncogene homolog 3	-3.01814	P4 down vs P5
NFKBIE	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, epsilon	3.27866	P4 up vs P5
TP53	tumor protein p53	-2.27895	P4 down vs P5
ICAM1	intercellular adhesion molecule 1	4.39919	P4 up vs P5
HDAC4	histone deacetylase 4	-3.10613	P4 down vs P5
RB1	retinoblastoma 1	7.97378	P4 up vs P5
POLR3C	polymerase (RNA) III (DNA directed) polypeptide C (62kD)	2.78592	P4 up vs P5
CD44	CD44 molecule (Indian blood group)	-2.36198	P4 down vs P5
CD19	CD19 molecule	-2.02922	P4 down vs P5
IL10RB	interleukin 10 receptor, beta	-2.03687	P4 down vs P5
CD44	CD44 molecule (Indian blood group)	-2.04762	P4 down vs P5
TRAC	T-cell receptor alpha constant	6.00851	P4 up vs P5
SKP2	S-phase kinase-associated protein 2, E3 ubiquitin protein ligase	-3.14461	P4 down vs P5
CD40	CD40 molecule, TNF receptor superfamily member 5	2.15865	P4 up vs P5
USP7	ubiquitin specific peptidase 7 (herpes virus-associated)	2.07788	P4 up vs P5
PIK3R3	phosphoinositide-3-kinase, regulatory subunit 3 (gamma)	2.58693	P4 up vs P5
ICAM1	intercellular adhesion molecule 1	3.04389	P4 up vs P5
STAT3	signal transducer and activator of transcription 3 (acute-phase response factor)	2.43377	P4 up vs P5
IL10	interleukin 10	10.8833	P4 up vs P5
PLCG2	phospholipase C, gamma 2 (phosphatidylinositol-specific)	2.10483	P4 up vs P5
POLR3C	polymerase (RNA) III (DNA directed) polypeptide C (62kD)	2.84293	P4 up vs P5
TRAF3	TNF receptor-associated factor 3	-2.07643	P4 down vs P5
SYK	spleen tyrosine kinase	-2.51979	P4 down vs P5
JUN	jun proto-oncogene	2.22557	P4 up vs P5
SYK	spleen tyrosine kinase	-2.78945	P4 down vs P5

Table 17: Gene list of Rheumatoid arthritis virus infection GO function for Pathway enrichment of P4 vs P5.

Gene symbol	Gene title	Fold change	Fold change (Description)
CD86	CD86 molecule	2.63003	P4 up vs P5
ITGAL	integrin alpha L	12.3257	P4 up vs P5
HLA-DQA1 /// HLA-DQA2	major histocompatibility complex, class II, DQ alpha 1 /// major histocompatibility com	2.20304	P4 up vs P5
ATP6V0E1	ATPase, H+ transporting, lysosomal 9kDa, V0 subunit e1	2.45293	P4 up vs P5
ATP6V0E1	ATPase, H+ transporting, lysosomal 9kDa, V0 subunit e1	2.37085	P4 up vs P5
ATP6V1D	ATPase, H+ transporting, lysosomal 34kDa, V1 subunit D	3.06106	P4 up vs P5
ATP6V1G2	ATPase, H+ transporting, lysosomal 13kDa, V1 subunit G2	-3.32778	P4 down vs P5
ATP6V0E1	ATPase, H+ transporting, lysosomal 9kDa, V0 subunit e1	2.4917	P4 up vs P5
TNF	tumor necrosis factor	3.16861	P4 up vs P5
ITGAL	integrin alpha L	6.5589	P4 up vs P5
ATP6V1D	ATPase, H+ transporting, lysosomal 34kDa, V1 subunit D	3.20666	P4 up vs P5
TLR4	toll-like receptor 4	-4.99913	P4 down vs P5
ATP6V1C1	ATPase, H+ transporting, lysosomal 42kDa, V1 subunit C1	4.10795	P4 up vs P5
IGH	immunoglobulin heavy locus	3.08466	P4 up vs P5
CD86	CD86 molecule	2.08975	P4 up vs P5
IL18	interleukin 18	4.9278	P4 up vs P5
HLA-DQA1	major histocompatibility complex, class II, DQ alpha 1	23.1982	P4 up vs P5
ATP6V0E1	ATPase, H+ transporting, lysosomal 9kDa, V0 subunit e1	2.43363	P4 up vs P5
ICAM1	intercellular adhesion molecule 1	2.07973	P4 up vs P5
CCL3 /// CCL3L1 /// CCL3L3	chemokine (C-C motif) ligand 3 /// chemokine (C-C motif) ligand 3-like 1 /// chemokine	9.4619	P4 up vs P5
ATP6V1C1	ATPase, H+ transporting, lysosomal 42kDa, V1 subunit C1	3.92104	P4 up vs P5
TNFRSF11A	tumor necrosis factor receptor superfamily, member 11a, NFKB activator	29.9127	P4 up vs P5
ATP6V1H	ATPase, H+ transporting, lysosomal 50/57kDa, V1 subunit H	2.14838	P4 up vs P5
FOS	FBJ murine osteosarcoma viral oncogene homolog	6.01616	P4 up vs P5
ICAM1	intercellular adhesion molecule 1	4.39919	P4 up vs P5
FLT1	fms-related tyrosine kinase 1	-5.45163	P4 down vs P5
TNFRSF11A	tumor necrosis factor receptor superfamily, member 11a, NFKB activator	60.2175	P4 up vs P5
CD80	CD80 molecule	3.7323	P4 up vs P5
ICAM1	intercellular adhesion molecule 1	3.04389	P4 up vs P5
ATP6V1C1	ATPase, H+ transporting, lysosomal 42kDa, V1 subunit C1	2.36416	P4 up vs P5
JUN	jun proto-oncogene	2.22557	P4 up vs P5
CD86	CD86 molecule	3.09883	P4 up vs P5

Table 18: Gene list of HIF-1 signaling pathway GO function for Pathway enrichment of P4 vs P5.

Gene symbol	Gene title	Fold change	Fold change (Description)
PRKCB	protein kinase C, beta	-3.28371	P4 down vs P5
LDHA	lactate dehydrogenase A	3.7573	P4 up vs P5
AKT3	v-akt murine thymoma viral oncogene homolog 3	-9.13951	P4 down vs P5
MKNK2	MAP kinase interacting serine/threonine kinase 2	2.35927	P4 up vs P5
PRKCA	protein kinase C, alpha	4.36719	P4 up vs P5
TCEB1	transcription elongation factor B (SIII), polypeptide 1 (15kDa, elongin C)	2.28032	P4 up vs P5
PRKCB	protein kinase C, beta	-3.95232	P4 down vs P5
EGLN2 /// RAB4B-EGLN2	egl-9 family hypoxia-inducible factor 2 /// RAB4B-EGLN2 readthrough (NMD candidate)	-2.42346	P4 down vs P5
INSR	insulin receptor	-13.02	P4 down vs P5
INSR	insulin receptor	-10.0214	P4 down vs P5
TLR4	toll-like receptor 4	-4.99913	P4 down vs P5
PFKFB3	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3	6.6352	P4 up vs P5
PDHB	pyruvate dehydrogenase (lipoamide) beta	2.15527	P4 up vs P5
CYBB	cytochrome b-245, beta polypeptide	-10.0899	P4 down vs P5
IGF1	insulin-like growth factor 1 (somatomedin C)	4.35054	P4 up vs P5
AKT3	v-akt murine thymoma viral oncogene homolog 3	-4.63676	P4 down vs P5
PGK1	phosphoglycerate kinase 1	2.30704	P4 up vs P5
IGF1R	insulin-like growth factor 1 receptor	-2.35402	P4 down vs P5
CYBB	cytochrome b-245, beta polypeptide	-3.89518	P4 down vs P5
PGK1	phosphoglycerate kinase 1	2.00342	P4 up vs P5
PDHB	pyruvate dehydrogenase (lipoamide) beta	2.20281	P4 up vs P5
PGK1	phosphoglycerate kinase 1	2.60844	P4 up vs P5
AKT3	v-akt murine thymoma viral oncogene homolog 3	-3.01814	P4 down vs P5
FLT1	fms-related tyrosine kinase 1	-5.45163	P4 down vs P5
ENO2	enolase 2 (gamma, neuronal)	2.24257	P4 up vs P5
EGLN1	egl-9 family hypoxia-inducible factor 1	2.06073	P4 up vs P5

PIK3R3	phosphoinositide-3-kinase, regulatory subunit 3 (gamma)	2.58693	P4 up vs P5
EIF4E	eukaryotic translation initiation factor 4E	-2.08981	P4 down vs P5
STAT3	signal transducer and activator of transcription 3 (acute-phase response factor)	2.43377	P4 up vs P5
PLCG2	phospholipase C, gamma 2 (phosphatidylinositol-specific)	2.10483	P4 up vs P5
ARNT	aryl hydrocarbon receptor nuclear translocator	-2.51093	P4 down vs P5
IGF1	insulin-like growth factor 1 (somatomedin C)	2.56482	P4 up vs P5
INSR	insulin receptor	-7.25905	P4 down vs P5
PRKCA	protein kinase C, alpha	2.9003	P4 up vs P5

Table 19: Gene list of B cell receptor signaling pathway GO function for Pathway enrichment of P4 vs P5.

Gene symbol	Gene title	Fold change	Fold change (Description)
PRKCB	protein kinase C, beta	-3.28371	P4 down vs P5
RASGRP3	RAS guanyl releasing protein 3 (calcium and DAG-regulated)	4.99178	P4 up vs P5
INPP5D	inositol polyphosphate-5-phosphatase D	-2.73285	P4 down vs P5
MALT1	MALT1 paracaspase	2.06648	P4 up vs P5
AKT3	v-akt murine thymoma viral oncogene homolog 3	-9.13951	P4 down vs P5
PRKCB	protein kinase C, beta	-3.95232	P4 down vs P5
IGH	immunoglobulin heavy locus	3.08466	P4 up vs P5
SOS1	SOS Ras/Rac guanine nucleotide exchange factor 1	-2.22269	P4 down vs P5
SOS2	SOS Ras/Rho guanine nucleotide exchange factor 2	-2.75236	P4 down vs P5
AKT3	v-akt murine thymoma viral oncogene homolog 3	-4.63676	P4 down vs P5
SYK	spleen tyrosine kinase	-2.95498	P4 down vs P5
INPP5D	inositol polyphosphate-5-phosphatase D	-2.5195	P4 down vs P5
AKT3	v-akt murine thymoma viral oncogene homolog 3	-3.01814	P4 down vs P5
FOS	FBJ murine osteosarcoma viral oncogene homolog	6.01616	P4 up vs P5
NFKBIE	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, epsilon	3.27866	P4 up vs P5
FCGR2B	Fc fragment of IgG, low affinity IIb, receptor (CD32)	-5.93192	P4 down vs P5
CD22	CD22 molecule	5.51793	P4 up vs P5
CD19	CD19 molecule	-2.02922	P4 down vs P5
CD22	CD22 molecule	2.71482	P4 up vs P5
MALT1	MALT1 paracaspase	2.03822	P4 up vs P5
PIK3R3	phosphoinositide-3-kinase, regulatory subunit 3 (gamma)	2.58693	P4 up vs P5
IFITM1	interferon induced transmembrane protein 1	2.13966	P4 up vs P5
PLCG2	phospholipase C, gamma 2 (phosphatidylinositol-specific)	2.10483	P4 up vs P5
SYK	spleen tyrosine kinase	-2.51979	P4 down vs P5
JUN	jun proto-oncogene	2.22557	P4 up vs P5
SYK	spleen tyrosine kinase	-2.78945	P4 down vs P5

Table 20: Gene list of Systemic lupus erythematosus GO function for Pathway enrichment of P4 vs P5.

Gene symbol	Gene title	Fold change	Fold change (Description)
HIST1H2BK	histone cluster 1, H2bk	10.1634	P4 up vs P5
H2BFS /// HIST1H2BK /// LOC102724334	H2B histone family, member S (pseudogene) /// histone cluster 1, H2bk /// histone H2B t	13.384	P4 up vs P5
HIST1H4H	histone cluster 1, H4h	6.11459	P4 up vs P5
CD86	CD86 molecule	2.63003	P4 up vs P5
HIST2H2AA3 /// HIST2H2AA4	histone cluster 2, H2aa3 /// histone cluster 2, H2aa4	4.23268	P4 up vs P5
HIST1H2BG	histone cluster 1, H2bg	7.72212	P4 up vs P5
HLA-DQA1 /// HLA-DQA2	major histocompatibility complex, class II, DQ alpha 1 /// major histocompatibility com	2.20304	P4 up vs P5
HIST1H2AC	histone cluster 1, H2ac	10.4596	P4 up vs P5
H2AFJ	H2A histone family, member J	0.388068	P4 down vs P5
TNF	tumor necrosis factor	3.16861	P4 up vs P5
CD40	CD40 molecule, TNF receptor superfamily member 5	3.54652	P4 up vs P5
HIST1H2BF	histone cluster 1, H2bf	9.93283	P4 up vs P5
HIST1H2AE	histone cluster 1, H2ae	8.63859	P4 up vs P5
IGH	immunoglobulin heavy locus	3.08466	P4 up vs P5
HIST1H2BE	histone cluster 1, H2be	9.59668	P4 up vs P5
CD86	CD86 molecule	2.08975	P4 up vs P5
HLA-DQA1	major histocompatibility complex, class II, DQ alpha 1	23.1982	P4 up vs P5
HIST1H2BC	histone cluster 1, H2bc	16.3696	P4 up vs P5
HIST1H2BD	histone cluster 1, H2bd	8.79576	P4 up vs P5
HIST2H2AA3 /// HIST2H2AA4	histone cluster 2, H2aa3 /// histone cluster 2, H2aa4	4.80389	P4 up vs P5
HIST1H2AJ	histone cluster 1, H2aj	2.52574	P4 up vs P5
HIST2H2BE	histone cluster 2, H2be	3.11853	P4 up vs P5

HIST1H2BJ /// LOC105374995	histone cluster 1, H2bj /// uncharacterized LOC105374995	6.54354	P4 up vs P5
HIST1H2BI	histone cluster 1, H2bi	6.71909	P4 up vs P5
HIST1H2AD /// HIST1H3D	histone cluster 1, H2ad /// histone cluster 1, H3d	3.91175	P4 up vs P5
CD40	CD40 molecule, TNF receptor superfamily member 5	2.15865	P4 up vs P5
HIST1H2BC	histone cluster 1, H2bc	9.88139	P4 up vs P5
HIST1H2BG	histone cluster 1, H2bg	8.34053	P4 up vs P5
HIST1H4H	histone cluster 1, H4h	4.34216	P4 up vs P5
CD80	CD80 molecule	3.7323	P4 up vs P5
IL10	interleukin 10	10.8833	P4 up vs P5
CD86	CD86 molecule	3.09883	P4 up vs P5