MASS SPECTROMETRY-BASED PROTEOMICS ANALYSIS OF WHOLE CELL AND SECRETED EXTRACELLULAR VESICLES OF FOUR DIFFUSE LARGE B-CELL LYMPHOMA CELL LINES

Henrique Silva Caio Simões Baeta

Supervisors: Rune Matthiesen, PhD, Principal Investigator of CEDOC

Ana Sofia Carvalho, PhD, of CEDOC

A dissertation submitted in partial fulfillment of the requirements for the Degree of Masters in Biomedical Research

Dissertação para obtenção do grau de Mestre em Investigação Biomédica

February, 2021

Contents

Figures Index

1. Fundamentation

In humans, many types of immune cells contribute to an extensive capacity to resolve environmental hostility. Lymphomas are routinely referred to as cancers of the immune system, mainly because lymphomagenesis occurs on lymphocytes. Most lymphomas arise from B cells and lymphomagenesis can occur at distinct stages of their maturation. Ultimately, this results in distinctive forms of disease that indicate different clinical outcomes to treatment approaches. The importance of B cell development and germinal center (GC) reaction in human immunity and how dysfunction can lead to lymphomagenesis are reviewed next.

1.1. B cells

An unmeasurable number of possible pathogenic infections requires a virtually infinite ability by the human immune system as to identify and successfully target pathogenic invaders for destruction and clearance. B cells play critical roles in innate and adaptive immunity, that act together in human immune responses. Unlike other lymphocytes, B cells can produce and secrete antibody proteins, encoded by the Immunoglobulin (Ig) genes. Antibodies bind foreign molecular targets and work by direct inactivation of viral or microbial activity, or by serving as molecular "tags" for their recognition and obliteration through other immune effectors, such as phagocytic cells. B cells are phenotypically characterized by the expression at the plasma membrane (PM) of the B cell antigen Receptor Complex (BCR). Immunoglobulin (Ig) chains form the antibody portion of the PM bound BCR and are non-covalently associated with CD79A-CD79B heterodimer and its co-activators[1] (Figure 1). BCR function is to bind antigen and thereby initiate signal transduction cascades that lead to B cell activation. Low-level "tonic" BCR signaling is enough to sustain B cell survival. But, it is the antigen encounter that drives BCR signaling to effectively engage its downstream effectors, such as NF-κB, and promote B cell survival, proliferation (clonal expansion) and differentiation[2]. Molecular mechanisms controlling B cell interactions in physiology are highjacked during lymphomagenesis. B cell lymphomas arise from virtually any stage of their development and/or differentiation and, to some extent, mirror the molecular phenotypes of their cell-of-origin[3]. This is evidenced by gene and protein expression profiling studies using patient biopsies and model systems. Likewise, many notions into B cell biology originated from studying their malignant counterparts. Importantly, these notions aid clinical practice by providing novel diagnostic and prognostic tools and exposing novel therapeutic targets to treat lymphomas.

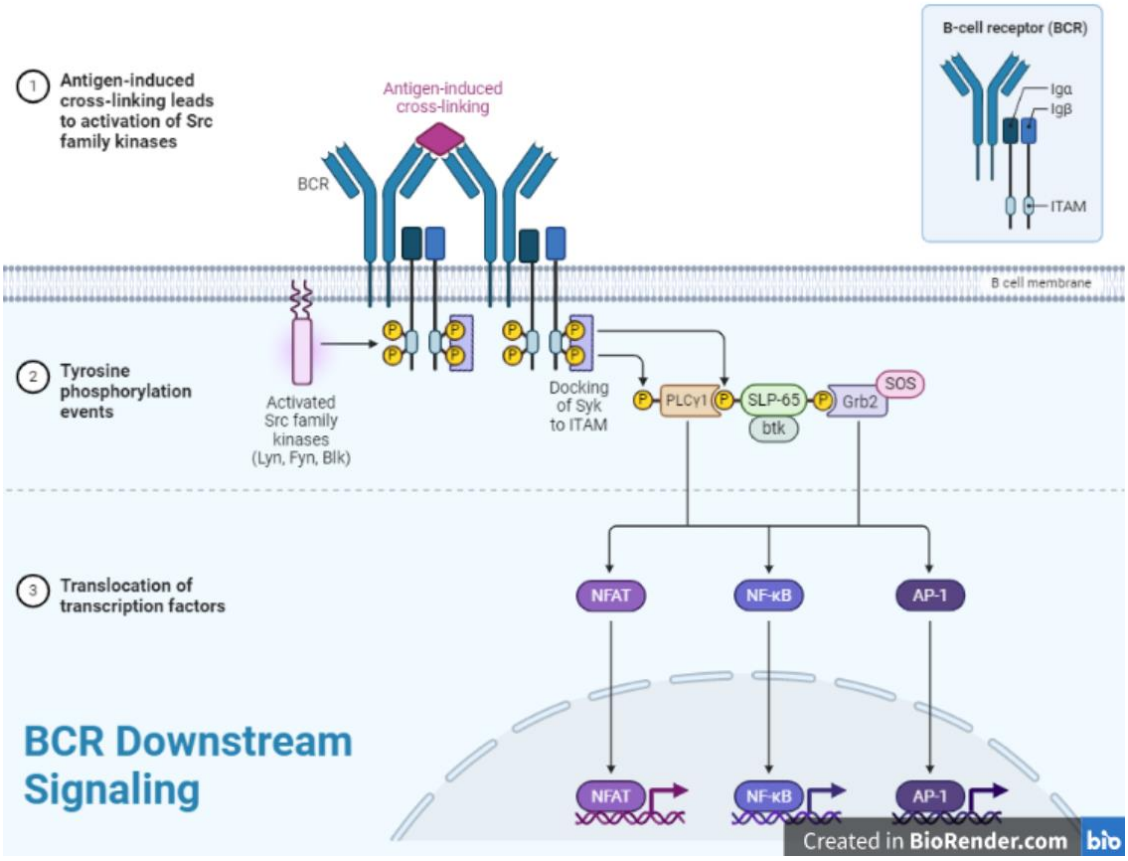

Figure 1. A simplified schematic of the B cell receptor (BCR) complex (binding to antigen) and downstream signaling effectors. *The B cell receptor complex (BCR) at the plasma membrane. Several proteins are involved in the regulation of signal transduction from the BCR. Upon antigen encounter, intracellular signaling drive B cell activation and/or differentiation by inducing well defined transcriptional programs. (created using BioRender.com)*

1.1.2. Early B cell development

In adult humans, B lymphocytes differentiate from hematopoietic stem cells in the bone marrow. There, functional rearrangement of immunoglobulin (Ig) gene segments take place in the RAG1/RAG2-mediated[4] somatic recombination process[1]. Many inherited V, J, and D gene segments - encoding antibody chains - vastly contribute for antibody diversity. In addition, the combinatorial joining of these gene segments (combinatorial diversification) and the joining mechanism itself allow the generation of an antibody repertoire able to recognize an estimated 5×10^{13} different antigen. Upon reliable Ig (heavy chain) V(D)J recombination, rearrangements on other chromosomes are suppressed to ensure B cell mono specific antibody commitment in the process of allelic exclusion[5]. A proper recombination is then followed by the pre-BCR complex assembly and expression at the PM[6]. These lymphocytes are considered immature B cells at this stage.

1.1.3. Bone marrow egress and systemic circulation

After recombination of Ig gene light chains, immature B cells acquire functional BCRs and eventually egress the BM in a process controlled by the interplay between G protein-coupled receptors, adhesion molecules and chemokine gradients. S1P1 receptors (S1P1R) expressed by immune cells act as key regulators of lymphocyte in-and-out movements from the BM[7] and other lymphoid tissues[8]. S1P1Rs bind extracellular lipid ligand sphingosine-1-phosphate (S1P)

present in systemic circulation and promote B cell exit from lymphoid or BM tissues. These cells, not yet presented with antigen, are considered naïve B cells and mainly express surface IgM or IgD antibody isotypes. They recirculate the human body through the bloodstream and lymph and regularly visit secondary lymphoid organs like the spleen, lymph nodes, and others.

1.1.4. Secondary lymphoid organs in Immunity

In the spleen, differentiation into marginal zone (MZ) or follicular B cell subtypes can occur, although none of the cell types seem exclusive to a particular area[9]. Naïve follicular B cells maintain in blood and lymph circulation and home to the spleen, bone marrow and secondary lymph nodes[10]. One can think of naïve lymphocytes as innate immune components that, due to direct contact with circulation, form a first line defense against blood-borne pathogens. They can detect pathogen-associated molecular patterns (PAMPs) such as lipopolysaccharides (LPS) from gram-negative bacteria or viral dsRNA, among others. Activation of Toll-like receptor (TLR) signaling and other evolutionary conserved mechanisms of intruder recognition induces lymphocyte activation and prompt physiological responses against invaders. Although innate and adaptive immune responses are functionally distinct, innate system components can stimulate the antigen-specific adaptive system, which, in turn, may amplify its response by recruiting innate effectors[11]. Intercellular interactions between lymphocytes and other immune or stromal cells are essential for humoral immune responses. A major challenge is to bring together antigen-presenting cells (APCs) and rare antigen-specific naïve B and T cells as to produce more specialized forms of immunity. Adult humans contain around 400 lymph nodes (LN) - secondary lymphoid organs - that form efficient venues for lymph filtration and consequent antigen surveillance, lymphocyte clonal expansion and antibody secretion (Figure 2A). Inside the LN, circulating naïve B lymphocytes encounter antigen displayed on APCs and, with help by local T and follicular dendritic cells (DCs), form highly dynamic structures called germinal centers[12]. B cell activation, antibody affinity maturation and posterior selection in germinal centers provide basis for human antibody mediated (humoral) immune responses.

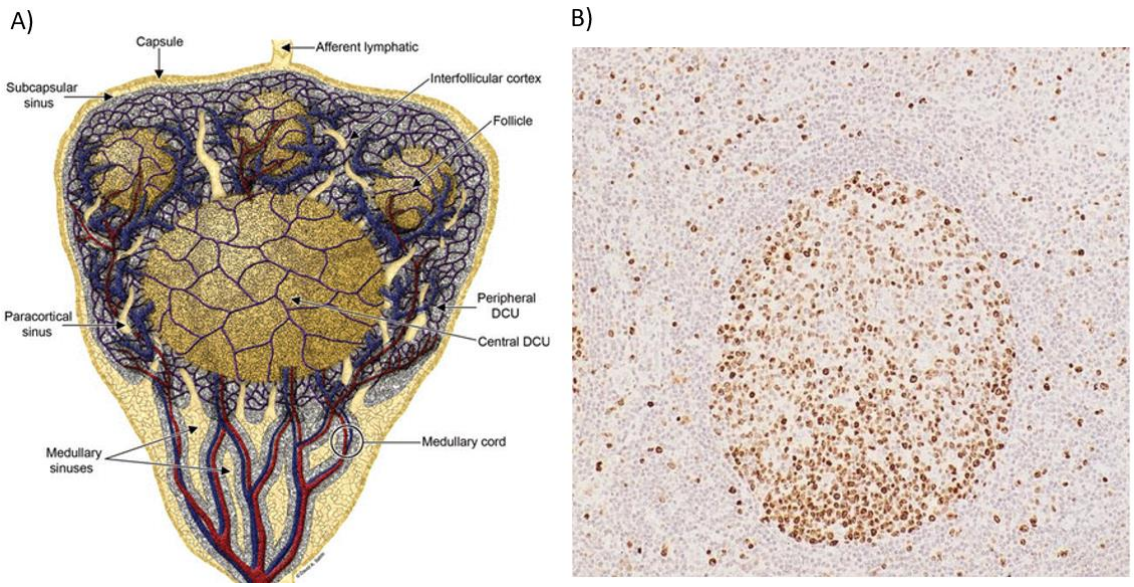

Figure 2. Representation of a human lymph node and histology of lymph node follicular region containing a germinal center (GC). *Lymph nodes are dispersed throughout the human body and convey the anatomical basis (A) for circulating lymphocytes to interact closely with other immune cells and produce high affinity antibodies. (B) "Dark zone" proliferating B cells in germinal centers, stained in brown*

(Ki-67 proliferative marker), displace mature B cells from the periphery to undergo antibody-antigen affinity selection in the "light zone". (from reference 12)

1.1.5. Germinal centers and antibody affinity maturation

Germinal centers (GCs) comprise unique lymphoid structures where B cell response to antigen is amplified and its specificity refined through antibody affinity maturation followed by clonal selection. GCs contain 2 histologically well-defined areas, the dark zone, and the light zone. Independent of T cell help, positive BCR engagement is followed by co-receptor activation and/or cytokine signaling and leads to cell division and differentiation[13]. A classical model of GC function states dark zone B cells undergo somatic hypermutation (SHM) on Immunoglobulin (Ig) genes, rapidly proliferate and later experience antigen affinity-based selection in the light zone (Figure 2B). Actually, two-photon microscopy of mouse lymph nodes showed B cells can transit back-and-forth between the dark and light zones, suggesting B cells experience rounds of mutation and selection within a GC[14]. Finally, activated B cells can differentiate into antibody secreting (plasmacytic) and/or memory B cells, relying on class switch recombination (CSR) for the generation of further antibody classes with distinct effector functions. High affinity antibody production relies on SHM and CSR tight regulation by the activity of activation-induced cytidine deaminase (AID) protein[15]. The Polymerase II stalling prominently at the Ig loci recruits AID to initiate antibody diversification at these sites[16], though off target mutations are well described[17].

1.1.6. Germinal center reaction double-edge sword

An amplified B cell proliferative capacity combined with physiologically induced DNA breaks account for increased risk of malignant transformation. Abrogation of AID in B cell lymphoma mouse models prevented development of GC and post-GC, but not of pre-GC derived lymphomas[18]. AID-induced DNA remodeling events (SHM and CSR) thus seem strongly involved in GC lymphomagenesis. A recent study suggests CSR mostly ceases upon the onset of SHM, which seems to work as a protective biological mechanism whereby DNA breaks likelihood is diminished[19]. Two major forms of genetic damage caused by errors in antibody diversification are frequently observed in GC-derived lymphomas. Namely, chromosomal translocations and aberrant somatic hypermutation (ASHM)[20]. Interestingly, these translocations infrequently generate gene fusions (and fusion proteins) typically reported in acute myeloid leukemia[21]. In contrast, they often juxtapose oncogenes with heterologous promoters and/or enhancers. Translocations involving the Ig loci are frequently observed in GC B cell lymphomas and generally result in ectopic expression of GC reaction master regulators, such as BCL2, BCL6 and/or MYC[22].

A common feature of GC B cell lymphomas is the retention of their cell-of-origin molecular phenotypes. Exception occurs in some classical Hodgkin's lymphomas, in which BCR-deficient cells end up losing most features of mature B cells[23]. That malignant GC B cell lymphomas express functional BCRs is puzzling given the frequent AID-induced mutagenesis and translocations involving the Ig loci detected in patient tumors[24] and model systems[18]. Actually, BCR signals appear of considerable importance for lymphoma growth and survival[2] and for their retention of GC B cell phenotypes. In accordance, Ig translocations were reported to mainly target non-productive Ig alleles in GC lymphomas[25,26]. Further, many genetic aberrations spotted in B cell lymphoma tumors target BCR constituents and downstream effectors. These considerations highlight Immunoglobulin genes as driver "lymphoma oncogenes" endorsing expansion of pre-malignant B cells, requiring additional cooperating "oncogenic hits" for tumor development.

1.2. B cell non-Hodgkin's lymphoma

Non-Hodgkin's lymphomas (NHL) comprise a diverse spectrum of mature lymphoid malignancies that typically arise in secondary lymphoid organs. Genome-wide association studies revealed particular gene loci associate with excessive risk for non-Hodgkin's lymphoma development[27,28]. Immune deficiencies, previous infection or therapy, lifestyle[29] and genetics are also being probed as risk factors[30]. Interestingly, none seem to reason why approximately 85-90% NHLs derive from B cells. Besides significant bias towards the B cell lineage, at least 65% NHL cases are germinal center (GC) derived B cell lymphomas as indicated by the presence of somatic mutations in their coding genomes[31]. During GC lymphomagenesis, B cells experience functional reprogramming of signaling pathways and transcriptional profiles in control of the GC reaction[22]. This results in the extreme heterogeneity observed within GC-derived lymphoma patients. Still, molecular profiling studies could discriminate dark zone B cells as normal counterparts to Burkitt's (BL) lymphoma, while Follicular (FL) and germinal center-like (GCB)-diffuse large B cell lymphomas appear more as transiting light zone B cells. Activated B cell like (ABC)-diffuse large B cell lymphomas (phenotypically) resemble post-GC B cells committed to plasmacytic differentiation (Figure 3).

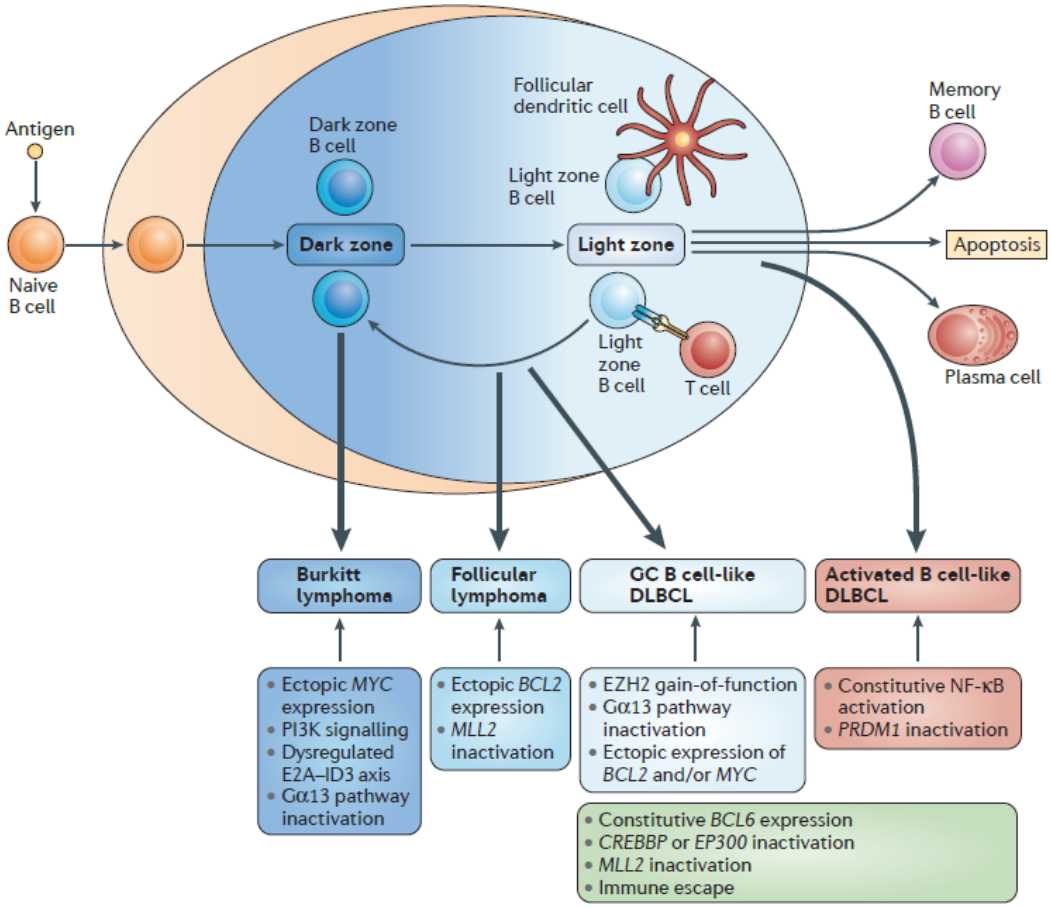

Figure 3. Origin of germinal center B cell lymphomas and main oncogenic features. *Most B cell lymphomas originate from germinal center B cells. Molecular profiling studies showed distinct GC B cell lymphomas arise from different stages of GC maturation. The aggressive diffuse large b cell lymphoma transcriptional phenotypes resemble those of late stage GC B cells. (from reference 22)*

Diffuse large B cell lymphomas (DLBCL) are aggressive NHLs that, aside "*de novo*" cases, can stem from clinical evolution of indolent follicular lymphoma or chronic lymphocytic leukemia

cases[22]. Transformation of canine marginal zone lymphoma (nMZL) into a deadly hepatic DLBCL tumor (shared clonal identity) is also reported[32]. Apparently, distinctive mechanisms drive DLBC lymphomagenesis and can strongly influence pathophysiology and clinical outcomes. The heterogeneous genetic, epigenetic and clinical presentation of DLBCL patients reported in landmark studies turned into a symbolic case study for precision medicine interventions[33].

1.3. Diffuse large B cell lymphoma (DLBCL)

Diffuse large B cell lymphomas (DLBCL) are the most frequently diagnosed subset of non-Hodgkin's lymphoma (~40% of cases) and represent about 1/3 of all B cell lymphomas. As the name hints, DLBCL cells show enlarged morphology and are present in a diffuse pattern along homing tissues[34] (Figure 4). DLBCL tumors usually grow within lymph nodes, which leads to lymphadenopathy. The European society of molecular oncology (ESMO) recommends full lymphadenectomy as optimal method to achieve final diagnosis of nodal DLBCL tumors[35]. This approach allows examination of nodal architecture and molecular phenotyping but carries the need for surgical intervention, causing unwanted burden to patients and clinicians. Roughly 40% of cases present extra-nodal primary tumors, being considered distinct disease entities that require specific diagnostic and therapeutic modalities[36]. Although apparent preference for the gastrointestinal tract, extra nodal DLBCLs can home virtually any human tissue, typically accompanied by fibrosis[37]. DLBCLs patients are mainly divided into 2 cell-of-origin biological subtypes that differ in their clinical course to frontline treatment. (see below)

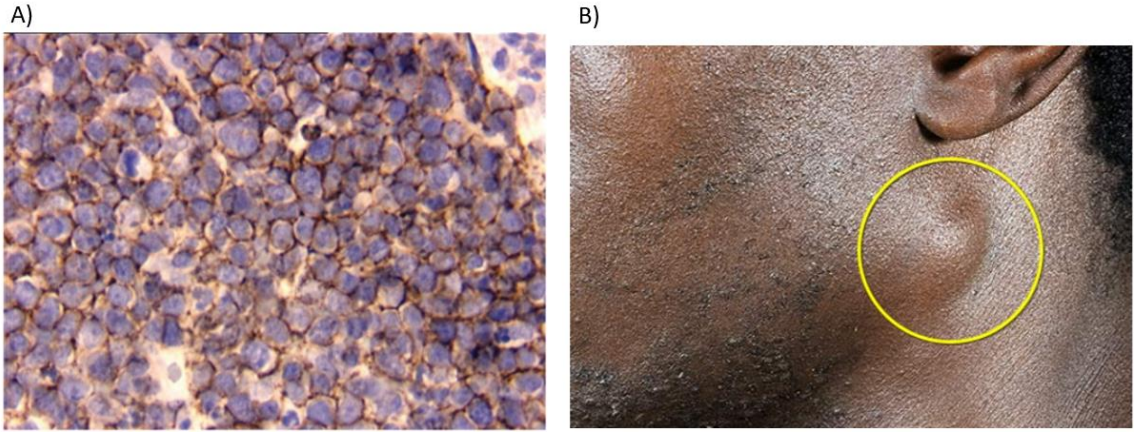

Figure 4. Diffuse large b cell lymphoma histopathology and pathophysiology. *(A) Diffuse large B cell lymphoma (DLBCL) tumor cells (in blue) usually have enlarged morphology and are "diffused" in the homing tissue. These cells stain positive for the B cell marker CD20 protein (in brown). (B) Most DLBCL tumors grow inside lymph nodes (LNs), leading to excessive swelling and systemic symptoms. To achieve definitive diagnosis a full excision of LNs is recommended which requires invasive surgery. (from reference 34)*

1.3.2. Cell-of-Origin classification and R-CHOP (rituximab – cyclophosphamide, doxorubicin, vincristine, and prednisone) as standard of care

By the turn of the millennium gene expression profiling (GEP) of tumor biopsies prompted DLBCL dissection into a cell of origin (COO) classification. Tumor GEPs matched those of B cells at late stage germinal center (GC) maturation and postulated DLBCLs as GC-derived NHLs[38]. Besides 10-15% unclassified cases, DLBCL patients were annotated as germinal center B cell like (GCB) if a "normal" transiting light zone B cell transcription program is expressed; or as activated B cell like (ABC) if expressing similar programs to activated B cells in late GC differentiation (Figure 5A). Functional studies keep informing into how BCR signaling participates in these two DLBCL tumor

types[39,40], though heterogeneity remain within each disease entity[41]. DLBCL COO subtypes correlated distinctly to combination "CHOP"[42] (cyclophosphamide, doxorubicin, vincristine, and prednisone) chemotherapy prognosis. Worst overall response and survival rates reported for ABC cases[43].

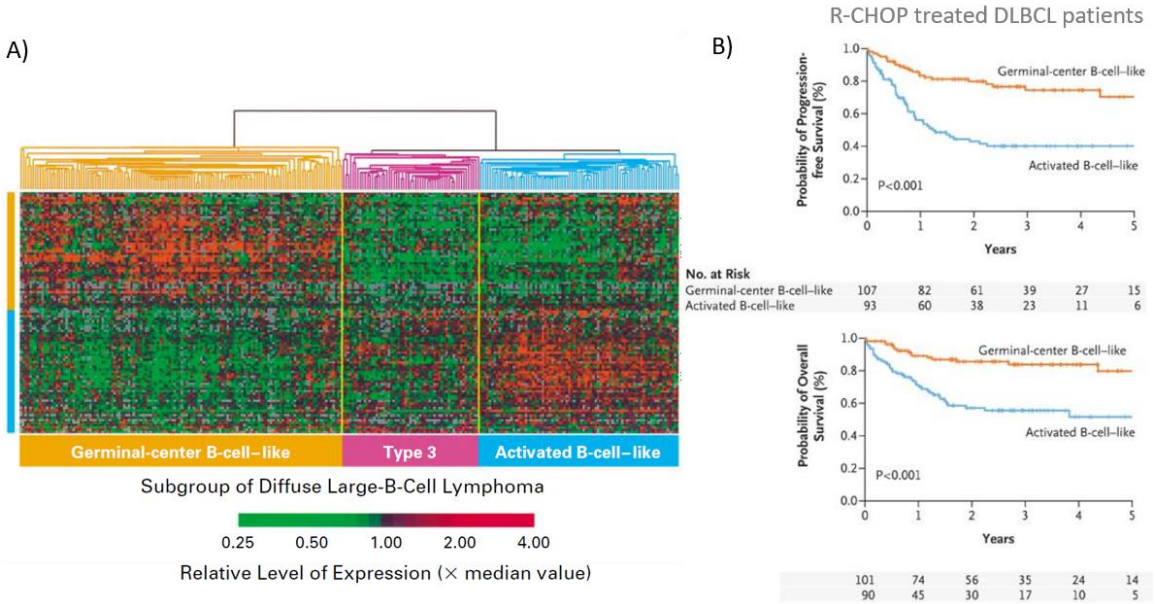

Figure 5. Diffuse large b cell lymphoma tumor transcriptional phenotypes and treatment prognosis. *(A) Gene expression profiling of diffuse large b cell lymphoma tumors allowed distinction of two molecular subtypes that reflect the transcription profiles of their cell of origin. (B) The germinal center b cell like (GCB) and the activated b cell like (ABC) cell of origin subtypes correlate distinctly to treatment prognosis. Frontline treatment consists on cytotoxic based chemotherapy (CHOP) added with monoclonal antibody Rituximab (anti CD20) and remains standard of care for these patients (from references 38 and 45).*

The same year, a pivotal trial showed anti-CD20 monoclonal antibody Rituximab (R) addition to standard cytotoxic-based combination (CHOP) therapy results in significant benefit for DLBCL patients. Rituximab (R-CHOP) treated patients exhibited superior complete response (76% vs 63%) and prolonged event-free and overall survival rates compared to CHOP alone, with no significant increase in toxicity[44]. GEP-based COO stratification was later (2008) probed for its prognostic value upon R-CHOP treatment. Rituximab addition benefited both subtypes but predictive disadvantage kept in ABC cases[45] (FIGURE 5B). Rituximab is shown to cross-link membrane CD20 proteins and favor complement recruitment *in vitro*[46]. But, also, to increase complement-independent macrophage mediated tumor clearance *in vivo*[47]. Its combination with cytotoxic chemotherapy remains as standard of care for newly diagnosed DLBCL patients[48] and usage in follicular lymphoma patients reduced the risk of histological transformation into aggressive lymphomas[49] as DLBCLs. Loss-of-function RNA interference[50] and CRISPR-Cas9 based[41] screenings allowed better understanding of key cellular pathways driving DLBCL biology. The latter experiments revealed the underlying biology of exceptional responses to added targeted therapy (using ibrutinib – BCR signaling inhibitor) reported in patients within the ABC DLBCL subtype[51]. Reported strategies beyond R-CHOP try to combine novel therapeutic agents targeting DLBCL subtype specific molecular phenotypes to improve clinical outcomes[52]. The COO

classification paved way to increasing understanding on DLBCL biology and kept its prognostic value, as recently reported[53].

1.3.3. DLBCL genetic subtypes

DLBCL tumor transcriptional phenotypes hold prognostic value but cannot entirely resolve DLBCL biological heterogeneity. ABC and GCB tumors share genetic abnormalities that, mostly based on a shared GC passage, contribute for their arrested late stage GC lymphoma phenotype. Diverse genetic background observed within COO subtypes incited a DLBCL genetic taxonomy[54]. Shared recurrent genetic alterations observed allowed the identification of 4 major DLBCL genetic subtypes. These, associate to distinctive gene expression profiles and, importantly, correlated differently to R-CHOP treatment prognosis[55]. The biological drivers of DLBC lymphomagenesis and tumor progression underlying genetic subtypes are under examination[56]. Identification of genetic subtypes comes with serious clinical implications. DLBCL patients are likely to benefit from therapies targeting subtype-specific phenotypes that result from genetic alterations. A recently developed probabilistic tool could classify large cohorts of DLBCL tumors (~54% of cases) into 7 genetic subtypes that subdivide clinical outcomes within the ABC and GCB subtypes[57]. Importantly, disparate survival characteristics were reported in these genetic groups to depend on their transcriptionally defined COO subtype (Figure 6). Besides feeding discussion for the rational design of future clinical trials, these results support DLBCL cell of origin subtyping still holds prognostic value. Determination of COO subtype is, thus, mandatory to effectively diagnose DLBCL patients.

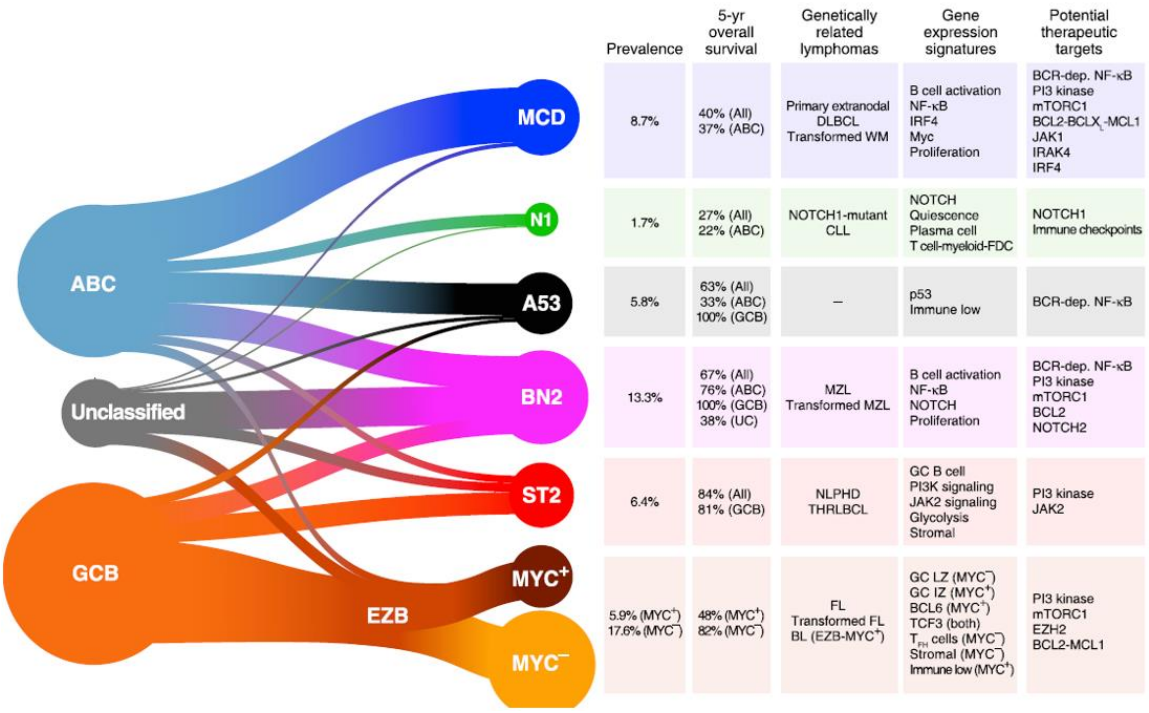

Figure 6. Diffuse large b cell lymphoma genetic subtypes and association with cell of origin subtypes. *Diffuse large b cell lymphoma tumors share genetic alterations and can be divided into 7 genetic subtypes. These, associate to different gene expression profiles and survival rates revealing that the cell of origin molecular phenotypes hold prognostic value within genetic subtypes. Targeted therapies focus on the molecular pathways and cellular phenotypes that result from these genetic aberrations. (adapted from reference 57)*

1.4. A role for intercellular communication

B cells are quite motile and interactive to their surroundings during developmental and maturation stages. Upon the start of the GC reaction, B cells rely on microenvironmental cues for survival, activation, and differentiation - critical for their adaptive immune function. These interactions involve the secretion and uptake[58] of soluble factors, direct cell to cell contacts[59], and extracellular vesicles (EVs) mediated transfer of proteins[60,61], microRNAs[62] and others. The remarkable complexity of signaling networks between malignant and "normal" cells is an emerging cancer hallmark and a promising research avenue[63]. It is increasingly evident that DLBCL cells alter their microenvironment to favor tumor formation, disease progression and treatment resistance[64]. Insights into how reprogramed communication competences, or lack of, are functionally integrated and selected along DLBCL progression and dissemination are unclear.

1.4.2. DLBCL Tumor microenvironment (TME)

The influence of the tumor microenvironment (TME) in germinal center lymphomas remains fairly unexplored[65]. In 1987 a study used patient derived lymphoma cells and revealed DLBCL cells were easily cultured *in vitro* compared to other lymphoma types[66]. This hinted DLBCL cells as less dependent on their TME. However, growing evidence suggest there is a dynamic network involving immune and stromal cells, blood vessels, soluble factors and extracellular matrix components supporting DLBCL cell survival[67]. Additionally, malignant B cells can avoid tumor immunogenicity and may convey stromal and/or immune cells in favor of tumor progression and dissemination[65]. Studying other intercellular interactions is ongoing[67,68], although the focus on DLBCL (cells) communication to other immune cells is strongly justified by its pathophysiology. Namely, their development in secondary lymphoid organs, where they need to escape tight regulation of the GC reaction. Also, because of their known capacity to invade and establish tumors in immune privileged locations such as the BM. These functional roles impact patient clinical outcomes[69,70].

1.4.3. Immune evasion

Several genetic aberrations target genes involved in the intercellular interactions between DLBCL (cells) and other immune cells[71]. These, ultimately result in 3 fundamentally distinct but related mechanisms whereby DLBCL cells can evade immune surveillance[31]. First, DLBCL cells commonly feature loss of surface MHC I and/or II, B2M and other core components at the immune synapse to prevent detection by macrophages, T and NK cells[22,71]. Second, DLBCL cells present altered expression and/or regulation of immunomodulators to reduce effector immune cell-mediated anti-tumor immunity. Accordingly, immune checkpoint blockade (PD1-PD-L1 and CTLA-4) as well as chimeric antigen receptor (CAR) T cell therapies are under clinical evaluation of their efficacy in DLBCL patients[72,73]. Another example is CD47 overexpression at DLBCL cell membrane, a "don't eat me signal" that inhibits phagocytic activity of macrophages and DCs. Treatment using anti-CD47 monoclonal antibody enhanced macrophage phagocytosis of DLBCL cells and synergized with Rituximab (anti-CD20) to completely eradicate DLBCL tumors *in vivo*[74]. Combination "5F9" (anti-CD47 antibody) immune checkpoint inhibitor and Rituximab showed promising results in relapsed DLBCL patients[75], in 2018. At last, apart from reduced detection and activity of immune effectors towards DLBCL tumors, malignant cells recruit or "reeducate" other immune cells into providers of supportive TMEs. Neutrophils attracted by CXCL8 producing DLBCL cells send back pro survival signals as secreted proliferation-inducing ligand APRIL, which overexpression associates to poor clinical outcomes[76]. Additionally, DLBCL derived interleukin-8 (IL8) binds neutrophil CXCR2 to induce the formation of neutrophil extracellular traps (NETs). Formed NETs directly activate TLR9 signaling back in DLBCL cells, increasing NF-κB and STAT3 signaling, contributing for tumor progression[77]. A recent study showed DLBCL tumors

are infiltrated by functionally different T cell subsets that don't feature exhausted phenotypes, hinting at putative local supportive interactions[78].

1.4.4. Targeting the DLBCL TME

Gene expression profiling (GEP) identified 2 stromal-cell derived transcriptional signatures in DLBCL tumor samples and revealed the TME as important prognostic factor to R-CHOP treatment[45]. The "Stromal-1" signature incorporate genes involved in extracellular-matrix deposition and histiocytic infiltration and associates to favorable prognosis. In contrast, the "stromal-2" signature seem to reflect a prognostically unfavorable pro-angiogenic TME. The need to consider the TME encouraged development of improved immunohistochemistry (IHC) based algorithms[79] to efficiently stratify DLBCL patients and reduce selection bias in clinical trials. Although cheaper and faster than GEP, IHC approaches still lack the sensitivity and specificity to precisely classify DLBCL tumors into cell of origin subtypes[80,81,82]. Interest grows for less invasive diagnostic modalities that circumvent the need for cumbersome collection of DLBCL tissue biopsies[83].

1.4.5. Growing interest in liquid biopsies

The use of liquid biopsies to obtain valuable clinical information is currently an active focus in cancer research[84,85]. DLBCL patient blood samples offer the possibility to indirectly access patient status and monitor therapeutic care, and allow further laboratorial inquiry[86]. Physiological juxtaposition between lymphatic and blood circulatory systems adds to the easy collection of blood samples and raises interest in the identification of potential circulating DLBCL biomarkers. The plasma from these patients contain circulating extracellular vesicles (EVs) secreted by DLBCL cells[87]. Although historically underscored, circulating EVs in human fluids and their contents can be probed in observational cohorts for their diagnostic and prognostic potential[88].

1.4.6. Small extracellular vesicles (sEVs), exosomes and more

Virtually all human cells can secrete and uptake small vesicular particles carrying biomolecules (proteins, DNA, RNAs, metabolites and lipids) that facilitate intercellular communications under normal physiology. Initial evidence came from studying the release by sheep reticulocytes of the transferrin receptor in small vesicular particles[89] and slightly after their (sEVs) biogenesis and secretion[90]. Germinal center B cells secrete antigen-presenting vesicles (containing peptide bound MHC class II) that can induce T cell responses[60]. Engagement of B cells with (antigen) specific CD4+ T cells increased exosome release, revealing a role of B cell sEVs at the GC immune synapse[61]. Additionally, EVs proteomes seem to contain relevant information into their cell of origin biological features.[91]

Several proposed classes of extracellular vesicles range from Exomeres[92] ~30-50 nanometers (nm), exosomes (50-200 nm), microvesicles and apoptotic bodies (~100-1000 nm), as well as large cancer oncosomes[93]. This classification is driven by current understanding of the molecular regulation underlying EVs biogenesis[94,95,96], cargo loading[97,98,99,100], secretion and uptake[101] (Figure 7). Small extracellular vesicles (sEVs) participate in local tumor biology[102] and in tumor cell dissemination to distinct tissues. Tumor exosomes were shown to induce pre-metastatic niche formation in the lung and BM[103], and liver[104], of naïve mice favoring metastatic progression. This brought light into long-term questions on distant tumor dissemination, paving way for functional assessment of molecular cargo (integrin proteins) driving organ tropism of disseminated tumor cells[105]. Evidence accumulates on the putative roles of tumor secreted small extracellular vesicles (sEVs) in lymphoma pathology, including in NHL[106].

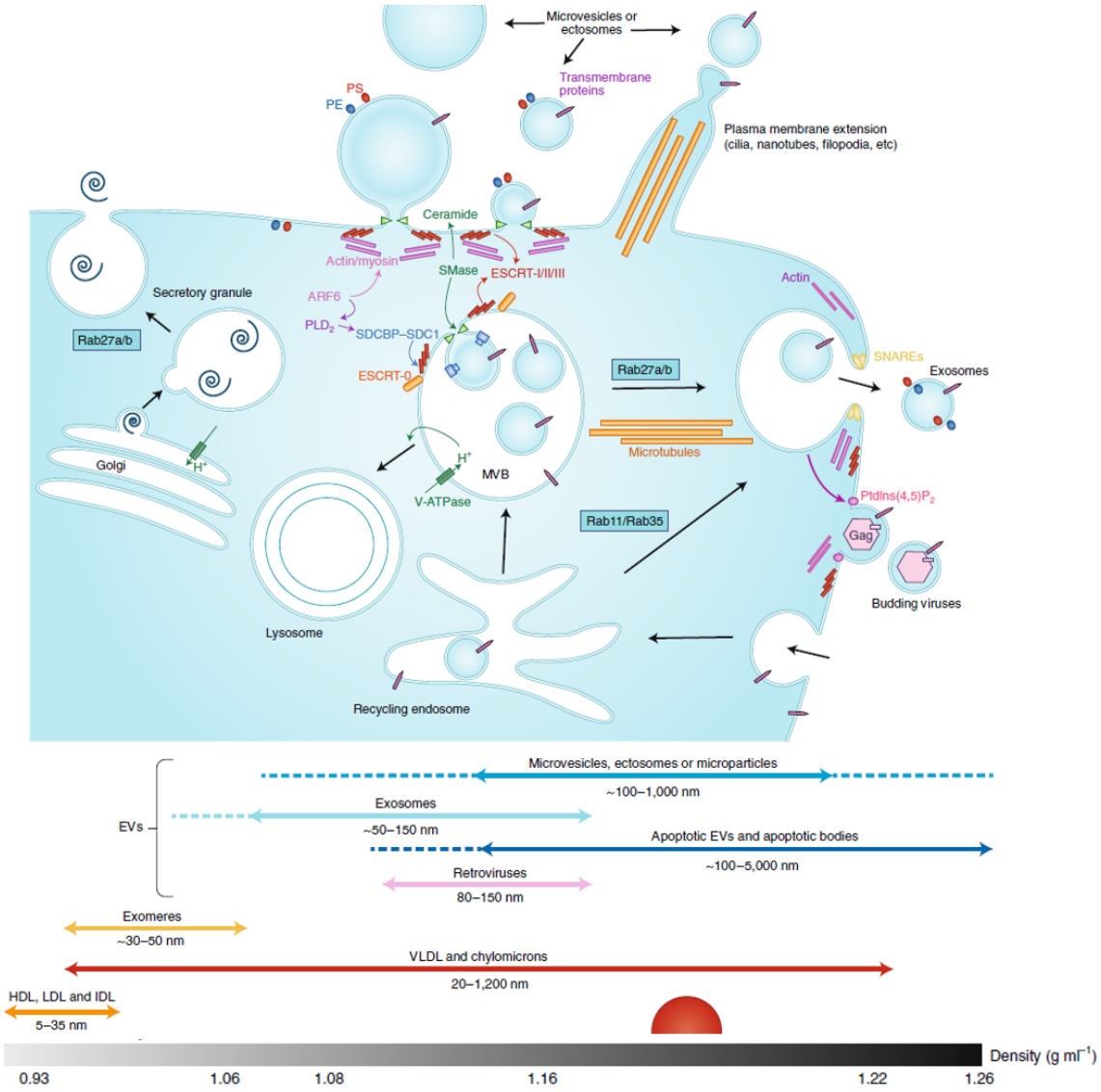

Figure 7. Classes of extracellular vesicles (EVs), their biogenesis and secretion. *Human cells secrete, and uptake, small vesicular particles containing proteins, nucleic acids, and others, to exchange molecular information under normal physiology. (A) Extracellular vesicles differ on their biological properties such as their size, density, and molecular cargo. (B) Small extracellular vesicles (sEVs) are divided into two main classes, endocytic derived Exosomes; and plasma membrane shed Microvesicles. Growing evidence suggest sEVs play active roles in cancer which poses blood circulating sEVs as attractive source of cancer biomarkers. (adapted from reference 100)*

1.4.7. sEVs in DLBCL pathophysiology

Functional studies focused on DLBCL secreted sEVs and exposed their potential roles in tumor progression[107] and treatment response[108] as well as their heterogeneous molecular cargo[109]. The need to determine sEVs secreted proteins was evidenced by the presence of CD20 (protein target of Rituximab) in sEVs isolated from DLBCL cell lines, and from patient blood[87]. Rituximab binds DLBCL exosomes and lead to complement fixation and consumption *in vitro*. CD20 secretion in sEVs may work as a resistance mechanism against humoral immunotherapy by shielding target cells from Rituximab, or simply by reducing CD20 cellular levels. Interestingly, the ABC transporter A3 (ABCA3) protein mediates cellular extrusion of cytostatic drugs (anthracyclines – CHOP drug backbone) inside DLBCL sEVs[110]. Since DLBCL sEVs may reach the

blood circulation, a study monitored their uptake by peripheral blood cells. Unlike NK and T cells, B lymphocytes and monocytes rapidly and preferentially internalized DLBCL sEVs, *in vitro*[111]. Other tissues emerge as candidate acceptors for circulating DLBCL sEVs, with attention shifted to bone marrow resident cells. DLBCL cells propagate constitutive active (pro-inflammatory) MYD88 signaling by transferring a mutated isoform into BM-derived macrophages and mast cells via sEVs[70]. Also, their direct injection into mice femurs could induce a pro-inflammatory BM microenvironment and recapitulate systemic effects of the disease. Interestingly, DLBCL sEVs injection into mice promoted DLBCL xenograft tumor growth, while increasing DC mediated antitumor immunity *in vitro*[112]. Recently, sEVs mediated transfer of neuron-specific enolase (ENO2) from DLBCL cells into mouse macrophages induced M2 polarization and modulated tumor progression *in vivo*[113] and *in vitro*[114]. To understand the putative roles of tumor secreted EVs in cancer pathophysiology and their value as clinical biomarkers it is imperative to characterize their molecular contents.

1.5. Mass spectrometry based proteomic profiling of cancer EVs

Recently, we conducted a meta-analysis using National Cancer Institute (NCI) mass spectrometry based proteomic data gathered from 60 human cancer cell lines and respectively secreted EVs[115]. Given their potential as clinical cancer biomarkers, we clustered EVs protein expression profiles and mined FDA approved drugs targeting abundant proteins across the panel. We pinpoint novel potential therapeutic targets and argue that secretion of drug targets (proteins) in cancer EVs may reduce therapy efficacy, as previously reported for DLBCL sEVs[87]. Our results indicate cancer EV proteomes partially mirror their cell-of-origin, representing potential clinical utility. A caveat is the EVs purification protocol used in the original study[116], which ignores existence of distinct EVs subtypes (Figure 8) and their biological properties[117]. A mass spectrometry-based analysis of blood circulating EVs proteins recently allowed the distinction of cancers[118], as previously reported for breast cancer tumor subtypes[119]. Such considerations highlight liquid biopsy EVs protein expression profiling to aid in clinical management of cancer patients. Importantly, studies using EVs should acknowledge limitations of EV purification and characterization methods and thoroughly report them to prevent flawed interpretation of data[120].

Besides their emerging roles in DLBCL pathophysiology, we speculate DLBCL secreted sEVs carry protein cargo that, as described previously using cell lines[121,122] and tumor samples[123–125], may be used to distinguish their cell of origin prognostic subtypes, GCB and ABC. Work described here takes part in Fundação para Ciência e Tecnologia (FCT) funded project "Portugal 2020, 02/SAICT/2017, 30087, Impact of B-cell deregulated pathways on tumor immune evasion - relevance for Diffuse Large B-Cell Lymphoma therapy response-LED". We comprehensively described the sEVs proteomes of in vitro cultured DLBCL cell lines, in parallel to their cellular proteomes and inform on their potential utility for patient subtyping. Additionally, proteomics data offers rational to generate novel hypothesis to investigate the contribution of secreted sEVs, and their protein contents, in DLBCL.

2. Motivation and objectives

2.1. Motivation/Abstract

The work herein described focuses on the potential of tumor secreted small extracellular vesicles (sEVs) as a source for protein biomarkers to identify the prognostic diffuse large b cell lymphoma (DLBCL) cell-of-origin subtypes. Namely, the activated B cell like (ABC) and germinal center B cell like (GCB) entities initially determined by gene expression profiling (Transcriptomics)[38,43,45]. Mass spectrometry (MS)-based protein expression profiling

(Proteomics) could, posteriorly segregate ABC and GCB DLBCLs based on their cellular proteomes[121,123]. These studies offered clinicians novel and more accurate targets for the cheaper, and globally applied, immunohistochemistry methods used in patient triage. However, most diagnostic modalities perform at the tissue level, which puts heavy burden on DLBCL patients and clinicians of surgical intervention to collect tumor biopsies. Recent studies suggest sEVs present in liquid biopsies carry molecular features (such as proteins, nucleic acids, and others) that can be useful for cancer subtyping and patient follow-up[118,119]. Until now, the protein expression profiles in DLBCL sEVs and their association to the cell of origin subtypes were not comprehensively studied. Acknowledging close physiological juxtaposition between the lymph and blood, isolating sEVs from patient blood may work as less invasive approach to achieve DLBCL prognostic subtyping. We speculate the proteome of sEVs can be used to stratify DLBCL patients into their prognostic subtypes. We aim to provide proof of principle that DLBCL cell-of-origin subtypes can be determined using their tumor secreted sEVs protein expression profiles.

2.2. Motivação/Abstract (PT)

O foco deste trabalho está na potencial utilização de vesiculas extracelulares (sEVs) como fonte de biomarcadores (proteínas) para identificar os subtipos biológicos de lymphoma difuso de grandes células B (DLBCL), com valor prognóstico. Nomeadamente, as entidades DLBCL de células B activadas (ABC) e de células B de centro germinal (GCB), inicialmente determinadas pelos seus perfis de expressão genética (Transcriptómica). Posteriormente, foi possivel determinar estes subtipos de acordo com os perfis de expressão proteica (Proteómica), ao nivel celular. Estes estudos possibilitaram a optimização de técnicas mais baratas e aplicadas universalmente, como immunohistoquimica, na triagem destes pacientes. No entanto, estas modalidades de diagnóstico funcionam ao nivel do tecido, sendo necessário remover cirurgicamente a biópsia tumoral. Estudos recentes sugerem que sEVs presentes em amostras liquidas (como o sangue) contém conteúdo molecular (proteínas, ácidos nucleicos e outros) com potencial utilização no diagnóstico de diferentes cancros, dos seus subtipos e no seguimento clinico destes pacientes. Até agora, a expressão proteica em sEVs de DLBCL, bem como a sua potencial associação com os subtipos biológicos não foram compreensivamente investigados. Nós pensamos que o proteoma secretado por células de DLBCL poderá ter aplicação na estratificação prognóstica destes pacientes. A nossa intenção é demonstrar que os subtipos biológicos de DLBCL podem ser determinados utilizando os perfis de expressão proteica secretados em sEVs.

2.3. Hypothesis and objectives

Can we identify DLBCL cell-of-origin subtypes based on their sEVs protein expression profiles?

Our Hypothesis is that the protein expression profiles of sEVs from *in vitro* cultured DLBCL cell lines reflect their cell-of-origin biological subtypes. To test this, we selected 4 commercially available cell line models of diffuse large B cell lymphoma. Namely, DB and HT cell lines annotated as germinal center B cell like (GCB) and RIVA and OCI-ly3 cell lines as activated B cell like (ABC) DLBCLs. We cultured cells in suspension, following supplier's instructions, until appropriate conditioned media volume is reached to obtain representative sEVs preparations. Next, we separated cells from their conditioned media and isolated sEVs using differential (ultra)centrifugation followed by discontinuous gradient ultracentrifugation[126]. Biological properties of DLBCL sEVs were studied using nanoparticle analysis, electron microscopy and others, in accordance with latest guidelines. Cells and sEVs were lysed and digested overnight using trypsin protease and resulting peptides analyzed using state-of-the-art Orbitrap Exploris

(Thermo Scientific™) mass spectrometer. Mass spectrometry-based quantitative proteomic profiles of DLBCL cell lysates and sEVs preparations were compared using well established and robust bioinformatics analysis strategies[127].

3. Methods

We cultured, *in vitro,* 4 commercial DLBCL cell lines and isolated sEVs from their culture media to perform a detailed mass spectrometry based proteomic analysis. DLBCL subtypes were each represented by 2 cell lines: DB and HT represent the GCB DLBCL; RIVA and OCI-ly3 the ABC DLBCL. We collected cells and respectively purified sEVs fractions from 3 independent cultures of each cell line model. Cell lines were tested for mycoplasma contamination and cultured following suppliers' (DSMZ) instructions to ensure reproducibility of our results. Explicitly, seed in/out cell densities, growth in suspension, and the percentage of FBS (v/v) in culture media. Antibiotic solution (P/S) containing Penicillin and Streptomycin was added as 1% total culture volume. Our results should be interpreted, and comparable to others, considering the experimental conditions and sEVs purification protocol detailed below.

3.1. DLBCL cell culture start

We thawed DLBCL cell line stock tubes (stored at second or third split at -150 ºC), added 15 mL RPMI, 20% FBS, 1% P/S and centrifuged at 200 x g, 5 min to clear from DMSO. Pellet cells were resuspended in RPMI, 1% P/S solution and started at the recommended % (v/v) FBS and cell densities. Cultures were kept at 37 ºC, 5% CO_2.

3.1.2. Cell count and viability estimation

DLBCL cell number estimation was performed at cell split and at the time of sEVs isolation. This allowed estimation of cell growth rates and cell "viability", using the Trypan blue test. Neubauer chamber was used to distinguish, under brightfield microscope, the cells with compromised plasma membrane (stained in blue) from "viable" cells (clear cytoplasm).

3.1.3. DLBCL cell culturing

DLBCL replicas were cultured in RPMI, 1% P/S and appropriate FBS % (v/v). Split time was 48 h. Cells were cultured to reach 220 mL culture media to ensure we could obtain representative sEVs preparations. In the last 48 h (before sEVs isolation) DLBCL cells were cultured in RPMI, 1% P/S and FBS_D in 11 x T75 flasks (20 mL each). FBS_D ("FBS-exosome-free") was obtained by O/N ultracentrifugation of diluted (1:4) FBS to prevent possible FBS contaminant vesicles in DLBCL sEVs preparations.[120,126] We did not detect any morphological, proliferative or viability changes on cultured cells upon the shift to FBS_D.

3.2. DLBCL sEVs isolation

After the last 48 h, DLBCL cells were collected to 50 mL tubes and centrifuged at 300 x g, 10 min, to pellet cells. Conditioned media (roughly 220 mL supernatant) was subjected to differential centrifugation steps as follows. First, centrifugation at 3000 x g, 20 min, was performed followed by 12 000 x g centrifugation, 1 h, to remove unwanted dead cells and cellular debris. After, ultracentrifugation at 110 000 x g, 2h10, was performed to pellet a partially purified sEVs fraction (Ex100k). These fractions (from 6 tubes) were resuspended in 1200 µL ice-cold PBS. Immediately after, extra ultracentrifugation was done using discontinuous gradient, as follows. Briefly, partially purified sEVs were resuspended in 32 mL PBS and gently loaded, without disturbing the interface, above 6 mL of a 30 % Sucrose Tris/D_2O solution and (ultra)centrifuged at 110 000 x g, 1h10. Using 10 mL syringe, the sucrose cushion containing purified sEVs was collected, diluted to 38 mL PBS, and subjected to a final ultracentrifugation at 110 000 x g, 16 h,

to pellet purified sEVs. Finally, purified sEVs were resuspended in 150 µL PBS, subjected to the BCA (colorimetric) assay and stored at -80 ºC until further use. All steps were performed at 4 ºC.

3.2.2. DLBCL cells collection

Pelleted cells from each (50 mL) tube were resuspended in 40 mL warm PBS and centrifuged at 300 x g, 10 min, to clear media contaminants. Cells were gently resuspended in warm PBS, pooled in a 50 mL tube, and counted. Cells were then split into 15 mL tubes, 5 x 10^7 cells per tube, and centrifuged at 300 x g, 10 min. Supernatant was discarded, and cell pellets stored at -80 ºC.

3.3. Protein estimation

Overall protein quantity in DLBCL cell lysates and respectively isolated sEVs was estimated using the BCA assay (Bio-Rad), according to the supplier's instructions. Briefly, samples were diluted in H_2O (1:2) to fit the standard BCA curve, incubated for 30 min at 37 ºC with BCA reagents, and the absorbance was measured using Tecan Instrument.

3.4. Nano particle tracking analysis (NTA)

Particle number and size distribution in DLBCL sEVs preparations were measured at the Champalimaud center for the unknown institute, with the Systems Oncology group. Briefly, representative DLBCL sEVs preparations were resuspended (1:500) in sterile PBS and subjected to NanoSight NS300 analysis (Malvern Instruments, Inc., Westborough, MA, USA).

3.5. Transmission electron microscopy (TEM) imaging

Imaging of DLBCL sEVs was performed at the Instituto Gulbenkian de Ciência (IGC) electron microscopy facility. Briefly, purified sEVs were loaded on a copper-palladium 100 mesh grid coated with 1% formvar in chloroform and carbon. Grids were glow-discharged before the adhesion of samples. DLBCL sEVs (DB cell line) were fixed using 2% formaldehyde in PBS, washed 10 times with distilled water, and stained using 2% uranyl acetate in ddH2O. Tecnai G2 Spirit BioTWIN from FEI operating at 120 keV and equipped with an Olympus-SIS Veleta CCD Camera was used for imaging.

3.6. Peptide preparation for mass spectrometry

DLBCL cells and sEVs were lysed using urea buffer, loaded onto spin filters, and washed exhaustively with urea (8M) in HEPES buffer. Next, proteins were reduced using DTT and alkylated with IAA^{128}. Protein digestion (overnight) was then performed using trypsin sequencing grade (Promega).

3.7. Liquid chromatography – mass spectrometry (LC-MS)

Peptide samples were analysed by nano-LC-MSMS (Dionex RSLCnano 3000) coupled to a Q-Exactive Orbitrap mass spectrometer (Thermo Scientific). Briefly, the samples (5 µL) were loaded onto a custom made fused capillary pre-column (2 cm length, 360 µm OD, 75 µm ID) with a flow of 5 µL per minute for 7 minutes. Trapped peptides were separated on a custom made fused capillary column (20 cm length, 360 µm outer diameter, 75 µm inner diameter) packed with ReproSil Pur C18 3-µm resin (Dr. Maish, Ammerbuch-Entringen, Germany) with a flow of 300 nL per minute using a linear gradient from 92 % A (0.1% formic acid) to 28 % B (0,1% formic acid in 100 acetonitrile) over 75 min followed by a linear gradient from 28 % B to 35 % B over 10 min at a flowrate of 300 nL per minute. Mass spectra were acquired in positive ion mode applying automatic data-dependent switch between one Orbitrap survey MS scan in the mass range of 400 to 1200 m/z followed by HCD fragmentation and Orbitrap detection of the 15 most intense ions observed in the MS scan. Target value in the Orbitrap for MS scan was 1,000,000 ions at a resolution of 70,000 at m/z 200. Fragmentation in the HCD cell was performed at normalized collision energy of 31 eV. Ion selection threshold was set to 25,000 counts and maximum

injection time was 100 ms for MS scans and 300 and 500 ms for MSMS scans. Selected sequenced ions were dynamically excluded for 45 seconds.

3.7.2. Database search

The obtained data from the 24 LC-MS runs were searched using VEMS[129,130] and MaxQuant[131]. A standard human proteome database from UniProt (3AUP000005640) and permutated protein sequences, where Arg and Lys were not permutated, were included in the database. Trypsin cleavage allowing a maximum of 4 missed cleavages was used. Carbamidomethyl cysteine was included as fixed modification. Methionine oxidation, N-terminal protein acetylation, deamidation of asparagine and glutamine was included as variable modifications. 5 ppm mass accuracy was specified for precursor ions and 0.01 m/z for fragment ions. The false discovery rate (FDR) for protein identification was set to 1% for peptide and protein identifications. No restriction was applied for minimal peptide length for VEMS search. Identified proteins were divided into evidence groups as defined by Matthiesen et al[130].

3.7.3. Quantitative analysis:

Quantitative iBAQ data from MaxQuant and VEMS were analyzed in R statistical programming language. The quantitative values were all subjected to statistical analysis using R package limma[132], with contrast between ABC versus GCB cell lines on both the cellular and sEVs level. Correction for multiple testing was applied using the method of Benjamini & Hochberg[133]. The result from PCA analysis of all data clearly separated the data into sample groups. VEMS spectral counts were preprocessed by removing common MS contaminants followed by log2(x + 1) transformation and quantile normalization. Quantitative outputs were subjected to statistical and functional analysis, described in the next section.

3.8. Data analysis, statistics, and reproducibility

We developed a data analysis pipeline to address whether the proteome regulation in DLBCL sEVs can provide insights into the biological subtype of their cell of origin. Proteomics data presented here is based on 2 quantitative matrixes. The first contains the spectral counts of mass spectrometry identified peptides (identification based on MS/MS scans). Peptides were collapsed into their encoding genes and used to estimate the total number of proteins identified in each fraction (proteome complexity) and unique proteins. Using the FunRich software[134], we performed functional enrichment analysis of DLBCL cellular and sEVs proteomes and compared them to the Vesiclepedia and Exocarta databases to pinpoint novel sEVs proteins and explore the regulation of common sEVs markers. The second matrix results from statistical comparison, based on quantitative analysis of DLBCL cell and sEVs proteomes (quantitative analysis based on ion counts from MS scans). Proteins were searched against the Uniprot database for their gene symbols, filtered for duplicates (using Python), and subjected to statistical and reproducibility analysis. Heatmaps were plotted using the R package ComplexHeatmap[135] and the scale() R base function, for scaling and Z-scoring the quantitative matrix when appropriate. Reproducibility was tested using unsupervised clustering analysis of quantitative proteomes (R package dendextend[136]) and analyzing Pearson's correlation coefficients comparing DLBCL cell and sEVs replicates (using Python). Principal component analysis (PCA), using R package FactomineR[137], was used to validate the reproducibility of biological and technical replicates and explore differences and similarities in the proteomes of DLBCL cells and sEVs. K-means cluster analysis was performed using the kmeans() R base function. Volcano plots were pictured using Enhancedvolcano[138] R package. Linear regression analysis was performed (using Python) to explore the extent of linear correlation between the (quantitative) proteomes of DLBCL cells and sEVs. Finally, the quantitative matrix was subjected to "complete functional regulation analysis",

described by our group[139], to dissect which functional annotations are shared, or differ, between DLBCL cells and sEVs, focusing on the comparison between DLBCL cell of origin subtypes.

4. Results

4.1. Research rational/workflow

In this study, we seek to identify the protein expression profiles of *in vitro* cultured DLBCL cells and of sEVs isolated from their conditioned media using a quantitative mass spectrometry (MS)-based proteomic analysis. Besides significant debate on the available methodologies to isolate EVs from cell conditioned media, differential centrifugation followed by ultracentrifugation (DUC) is widely used to purify sEVs. But, contaminants such as non-vesicular membranes, protein complexes and others may co-precipitate with sEVs. Trying to prevent this, an extra purification procedure using discontinuous sucrose cushion ultracentrifugation is recommended and frequently reported. The density of sEVs is equivalent to that of a 30% sucrose-D_2O solution (1.12-1.18 g/mL), which exerts "cushioning" effect on sEVs (that float in this fraction) while high density contaminants precipitate. So that sEVs proteins identified/quantified in this study serve as potential subtype specific biomarkers we coupled DUC and sucrose cushion ultracentrifugation to purify DLBCL sEVs (Figure 8).

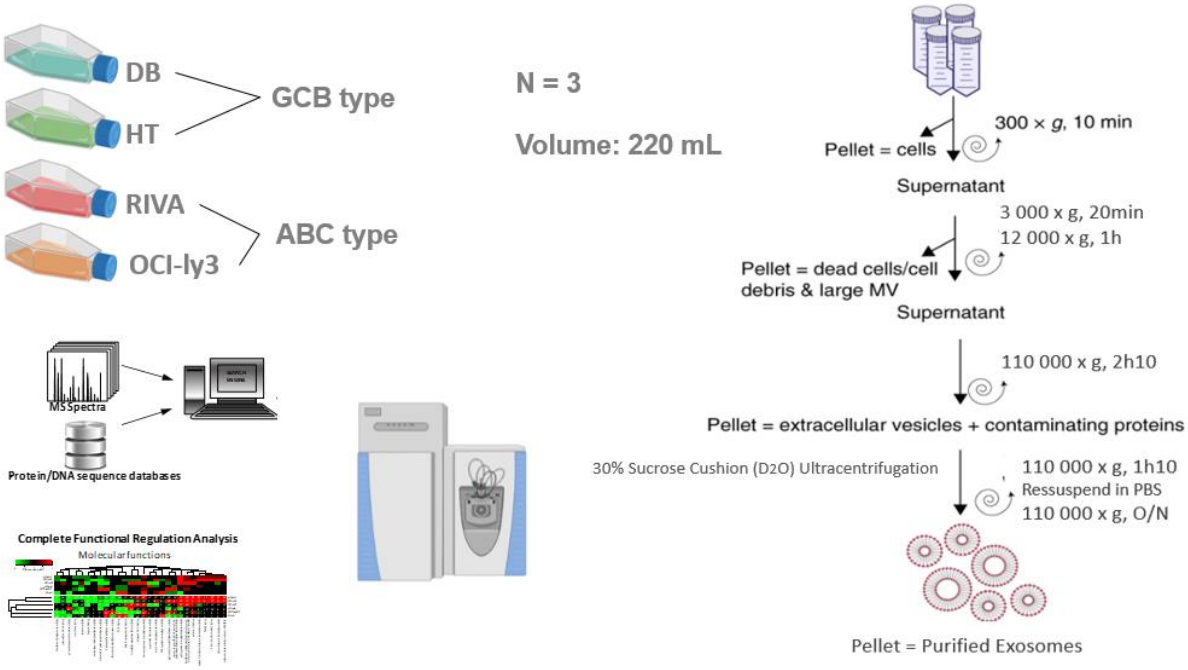

Figure 8. Experimental workflow. *We cultured 4 diffuse large b cell lymphoma cell lines and purified small extracellular vesicles (sEVs) from their conditioned media using differential (ultra)centrifugation followed by discontinuous gradient ultracentrifugation. Cells and sEVs were lysed and digested using trypsin protease. Resulting tryptic peptides were quantified using mass spectrometry. The quantitative proteomes of cells and sEVs were compared using robust bioinformatic strategies to inform on the potential of sEVs proteins to identify their cell of origin prognostic subtypes (ABC vs GCB).*

The yield of sEVs purification increase with the starting volume of conditioned media. To understand if we could obtain representative sEVs preparations to analyze biological properties and proteomic cargo, we first cultured DB cells to 220 mL of culture media. Cells were collected and two EVs fractions separated from conditioned media. These were obtained pre- and post-discontinuous sucrose cushion ultracentrifugation and subjected to BCA (colorimetric) assay to

estimate protein concentration in EVs preparations. Detailed in Table 1, protein concentration in sEVs (after 30% cushion ultracentrifugation) is below 1/4 of that observed in the Ex100k fraction (before cushion). We recovered ~28 ug total protein in the final DB sEVs fraction. Approximately 10-20 µg protein is sufficient for a robust LC-MS based analysis with deep proteome coverage. Additionally, a few µL from each sEVs preparation allow, to some extent, for characterization of biological properties. We considered 220 mL a suited final volume for the isolation of DLBCL sEVs.

Cell line	Volume (mL)	Total cells (millions)	Exo100k BCA estimate (ug/uL)	sEVs BCA estimate (ug/uL)	sEVs Total protein (ug)
DB	220	360	1,087626934	0,191425371	28,7138056

Table 1. Characteristics of small extracellular vesicle isolation. *To ensure we obtained representative small extracellular vesicle (sEVs) fractions to analyze their biological properties and proteomic cargo, we tried isolating DB cell line sEVs. We compared fractions obtained pre (Ex100k) and post (sEVs) 30% sucrose cushion purification. Protein estimates post purification are less than 1/4 of the previous fraction (equal resuspension volume). Total estimated 28 ug protein recovered allows posterior mass spectrometry analysis and characterization of biological properties.*

We cultured 3 biological replicates of each cell line model of DLBCL and respectively isolated sEVs from their conditioned media. Estimated growth rates of DLBCL cell lines are depicted in Figure 9. DLBCL proliferation fitted the expected exponential curves and appeared quite similar, independently of their biological subtype (ABC or GCB). Average cell viability estimated at isolation is reported in Figure 10A. Based on these results we suppose differences in sEVs number or protein contents should not result from stressful conditions nor (intracellular organelle) contamination due to excessive cell death/lysis. Our scale up culture of DLBCL cell lines seems reproducible and therefore adequate to isolate of sEVs from their conditioned media.

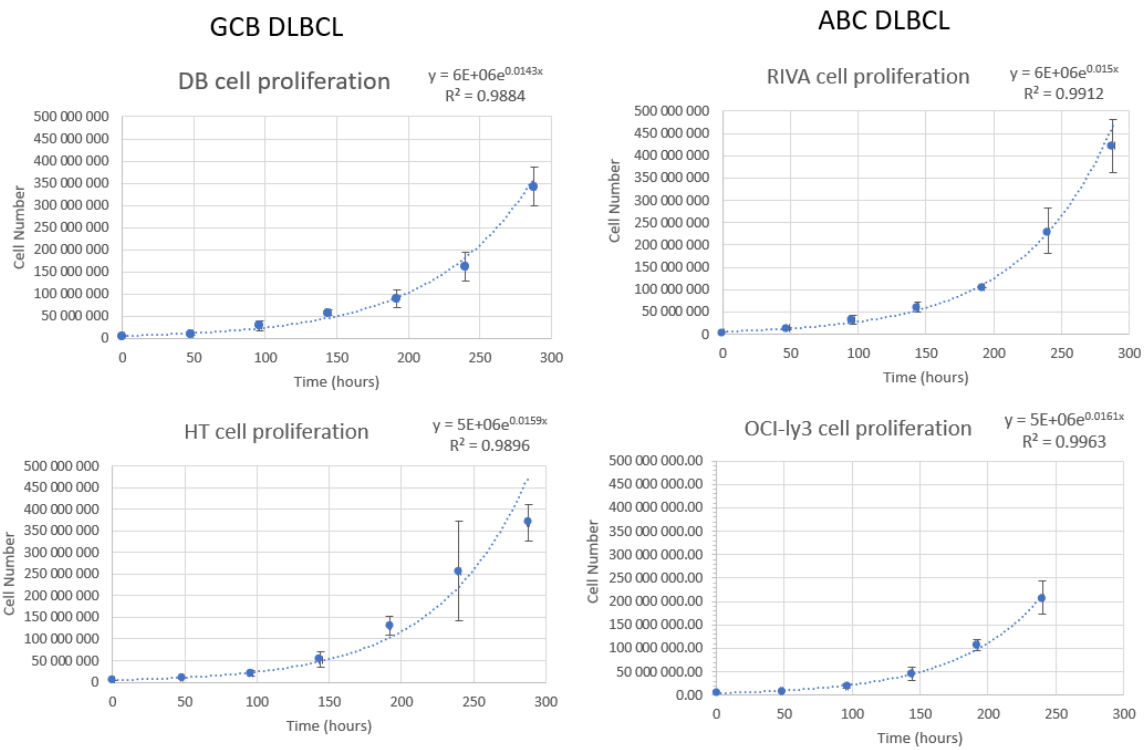

Figure 9. Cell proliferation of diffuse large b cell lymphoma cell lines. *We cultured 3 biological replicas for each diffuse large b cell lymphoma cell line. Estimated proliferative rates fitted the exponential growth rates. OCI-ly3 cells reach the final volume for isolation of small extracellular vesicles 48h before the other cell lines. No differences were observed between the 2 cell of origin subtypes (ABC vs GCB).*

Finally, sEVs fractions obtained according to Figure 9 were probed using the BCA assay and immediately stored at -80 ºC. Average estimated total protein (µg) in DLBCL sEVs fractions is plotted in Figure 10C. Estimates of total protein ranged from 21,13-57,63 µg suggesting consistency in our isolation protocol. The next section focus is on biochemical and biophysical features studied in DLBCL sEVs isolates, their associations between and to their respective cellular fractions.

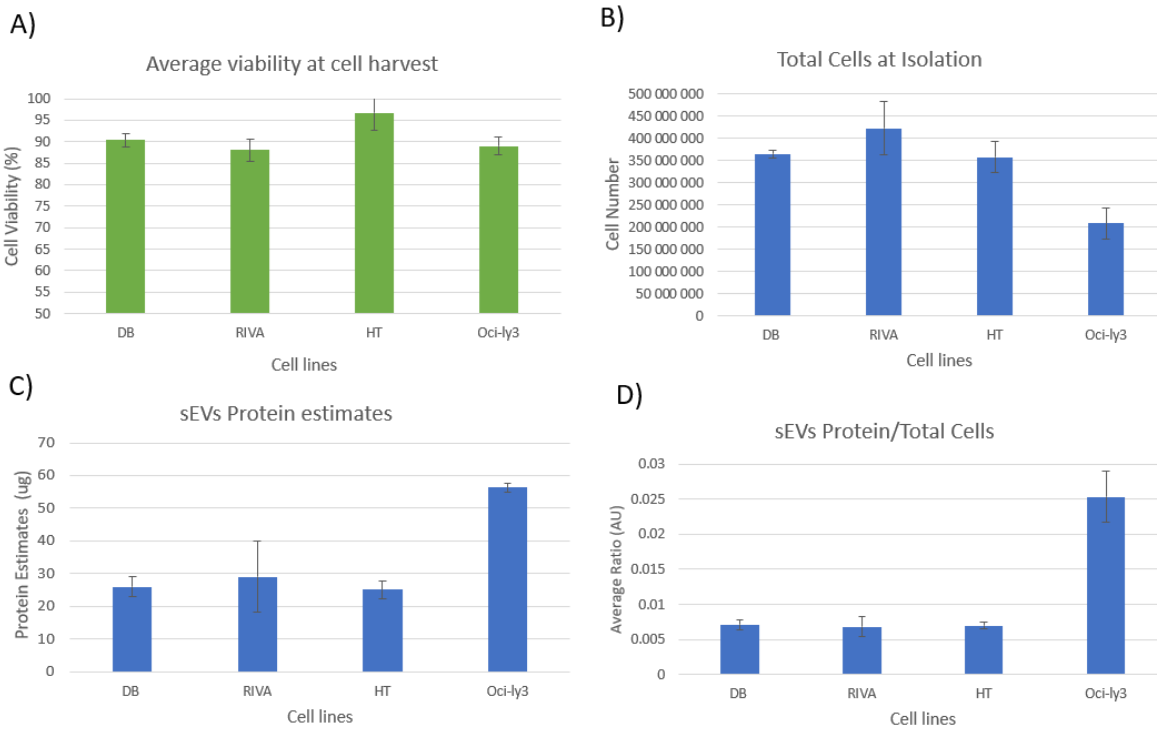

Figure 10. Characteristics of diffuse large b cell lymphoma small extracellular vesicles isolation. *(A) Viability estimates of diffuse large b cell lymphoma cell lines at small extracellular vesicles (sEVs) isolation. Similar results and low variability in each cell line suggest our scale-up culture is reproducible. (B) BCA (colorimetric) assay protein estimates in sEVs fractions. OCI-ly3 sEVs protein estimates were superior to other DLBCL sEVs. (C) Number of collected cells at sEVs isolation. OCI-ly3 cell number at isolation is about half other cell lines. (D) Ratio between sEVs protein estimates and total number of cells collected at isolation. Oci-ly3 protein-to-total cells is significantly superior to other cell lines.*

4.2. Biological properties of DLBCL sEVs

We share concerns raised by the international society for extracellular vesicles (ISEV) on the importance of thorough reporting experimental conditions and isolation protocols used in sEVs purification. First, we compare the average cell number collected at isolation day in Figure 10B. OCI-ly3 cells reach the final volume for sEVs isolation at about half the cell density – a split before – than that of other DLBCL cell lines. Interestingly, the average estimated protein concentration of OCI-ly3 sEVs was superior to that of other DLBCL sEVs (Figure 10C). Difference is kept when normalizing sEVs protein estimates to the total number of cells collected at isolation (Figure 10D). It seems that DLBCL sEVs total protein estimates do not strictly relate to the total number of cells collected at isolation day.

Next, we used nanoparticle tracking analysis (NTA) to assess the number and size distribution of isolated DLBCL sEVs. The NanoSight instrument detects and quantitates (nano)particles (30-1000 nanometers) moving in suspension under “Brownian” motion[140]. Representative DLBCL sEVs preparations, from each cell line, were diluted in filtered PBS and processed for NTA. The size distribution of particles detected in DLBCL sEVs preparations is shown in Figure 11. A distinct “peak” of particle concentration is perceived at around 100 nm in all sEVs preparations. We found significant enrichment in small size particles - within the inferior detection limit (~30 nm) and 200 nm – in all tested DLBCL sEVs preparations.

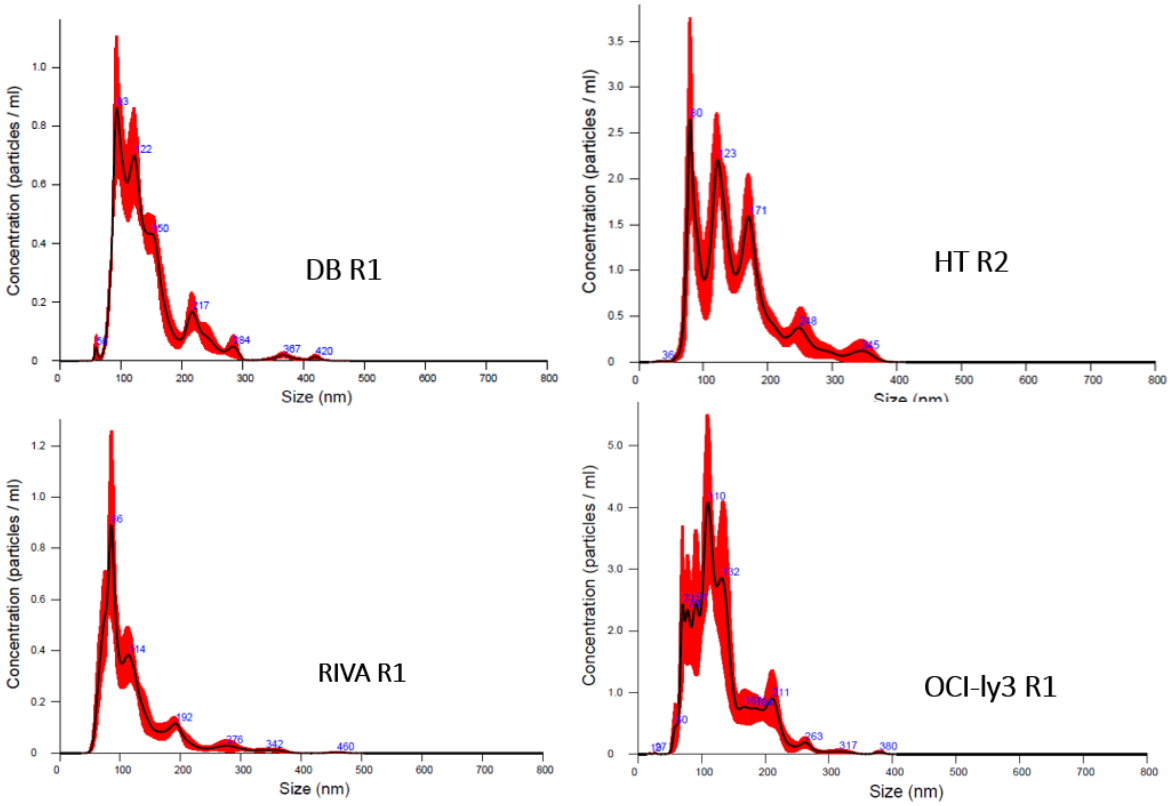

Figure 11. Diffuse large b cell lymphoma small extracellular vesicles size distribution. *Representative preparations of diffuse large b cell lymphoma small extracellular vesicles (sEVs) were detected and quantitated by nano particle tracking analysis. A "peak" of particle concentration is always seen around 100 nanometers (nm). No differences in sEVs size distribution associated with cell of origin subtypes. Our sEVs isolation protocol retrieved small sized particles, located mainly between the inferior detection limit (~30 nm) and 200 nm.*

Figure 12A shows the average particle concentration estimated (6 technical replicates) in each representative DLBCL sEVs preparation. Particle concentration ranged between 1-3 x 10^{11} particles/mL. DB sEVs showed the most abundant particle concentration. Normalizing particle concentration to the total cell number collected at isolation revealed that the OCI-ly3 particle-to-total cells ratio is alike that of DB (Figure 12B). The ratio between nanoparticle count (NTA) and total protein estimate (BCA) is proposed as a general tool for assessing purity of sEVs preparations[141]. Since our detected range of particle per mL (NTA) is within that of their study, we calculated the particle-to-total protein ratio(s) in our representative DLBCL sEVs preparations (Figure 12C). DLBCL sEVs isolates showed particle-to-protein ratios between 4-16 x 10^{10} particles per µg protein. Being in the same magnitude of particle concentration, it seems our protocol yields similar amount of small size particles in the final sEVs fractions.

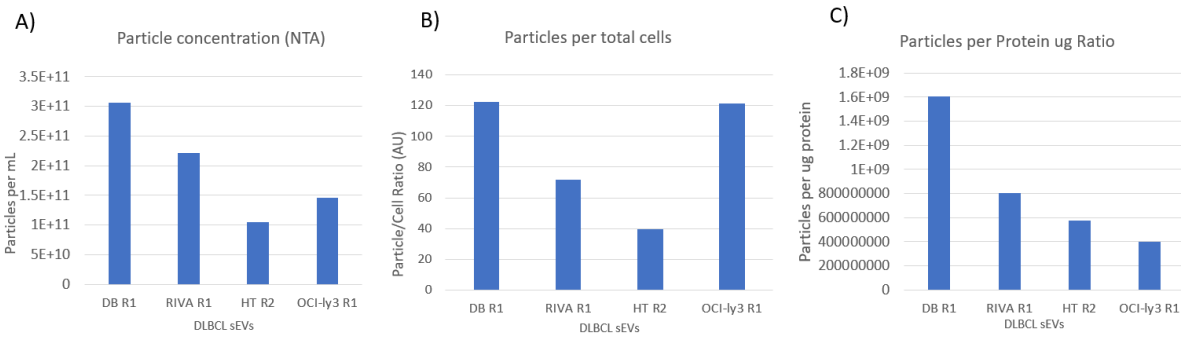

Figure 12. Particle concentration of diffuse large b cell lymphoma small extracellular vesicles. *(A) Diffuse large b cell lymphoma small extracellular vesicles (sEVs) were quantitated by nano particle tracking analysis. DB sEVs showed the highest particle concentration. (B) Normalization of particle concentration to the total number of collected cells. Oci-ly3 and DB particle-to-total cells ratios seem similar. (C) Normalization of particle concentration to the total protein (ug) estimates in sEVs preparations.*

To achieve visual inference of isolated particles, we processed a few microliters (µL) from a representative DLBCL sEVs preparation (DB cell line) for transmission electron microscopy (TEM) imaging. We fixed sEVs in formaldehyde and stained them using uranyl acetate as detailed in the methods section. DB sEVs appear as “cup-like shaped”, characteristic of vesicle membrane morphology display upon fixation and dehydration for TEM (Figure 13). Observations suggest that our experimental settings and purification protocol yield particles with morphological features and size distribution consistent with previous studies.

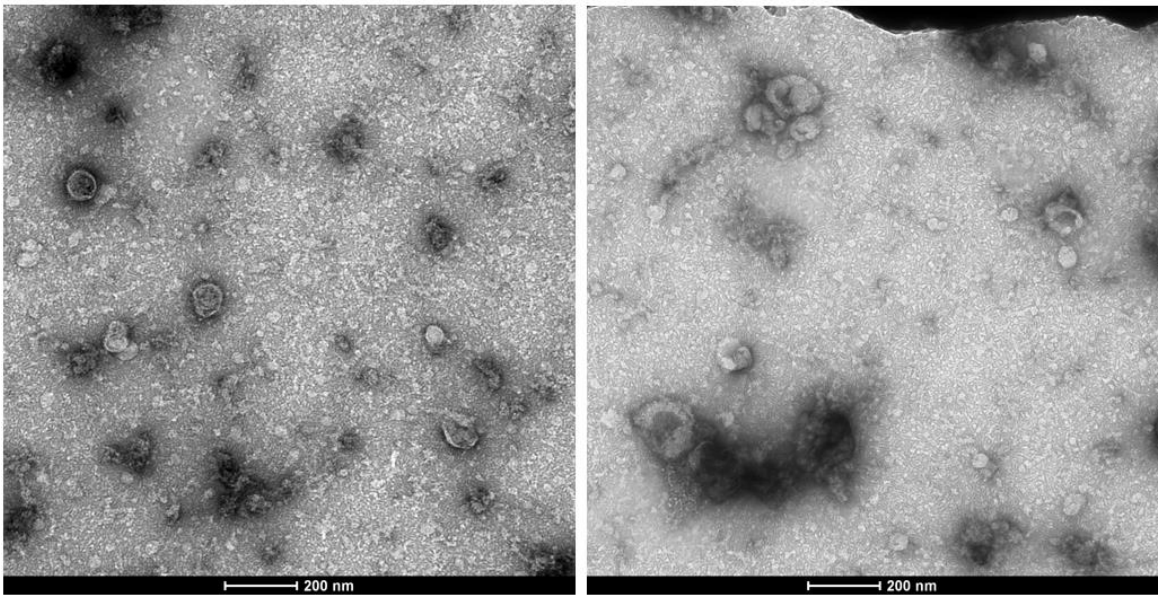

Figure 13. Imaging diffuse large b cell lymphoma small extracellular vesicles. *Transmission electron microscopy (TEM) was used to visualize diffuse large b cell lymphoma small extracellular vesicles (sEVs). A representative sEVs preparation (DB cell line) was fixed and stained for TEM imaging. sEVs show a “cup-like shaped” morphology, characteristic of vesicular membranes after fixation and dehydration for TEM.*

4.3. Reproducibility of MS-based proteomics

To study the proteome regulation of DLBCL cells and sEVs, we analyzed 2 technical replicas by LC-MS of each biological replica to ensure deep and reproducible coverage of proteome profiles. Proteomics data presented next refers to the DB and RIVA cell lines, and respective sEVs. Mass spectrometry results for the HT and OCI-ly3 cells and sEVs arrived shortly before thesis submission and is currently under analysis.

Figure 14 displays the scaled quantitative protein expression profiles of DB and RIVA cells. The reproducibility of our technical and biological replicas is high based on the MS-based proteomics analysis. Besides promptly observed similarity between technical and biological replicates of DB and RIVA, there is large overlap between the proteomes in terms of identified proteins (see Figure 20A). An expected result since these cell lines model the same disease - DLBCL. Differences across expression profiles should reflect subtype or cell line specific molecular features, and cellular phenotypes. In Figure 15 we compare the (quantitative) proteome profiles of DB and RIVA sEVs. High replicability is also seen in the expression profiles of DB and RIVA sEVs.

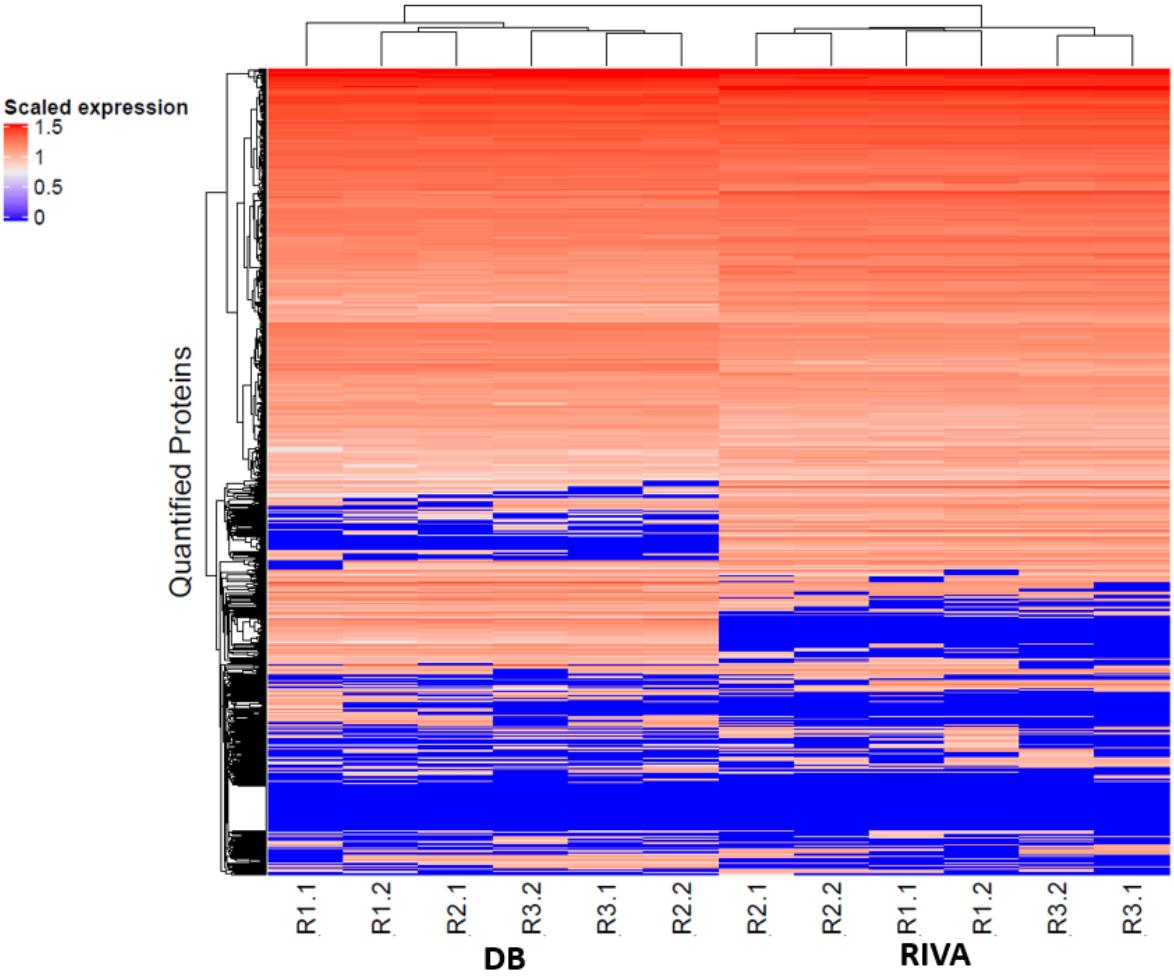

Figure 14. Proteome regulation of DB and RIVA cells. *Quantitative protein expression profiles (scaled) of diffuse large b cell lymphoma cell lines. Two technical replicas for each of the 3 biological replicas were analyzed by mass spectrometry. High similarity is observed across replicates suggesting high reproducibility in our proteomic analysis. DB and RIVA cellular proteomes show large overlap, expectedly since both cells model diffuse large b cell lymphoma.*

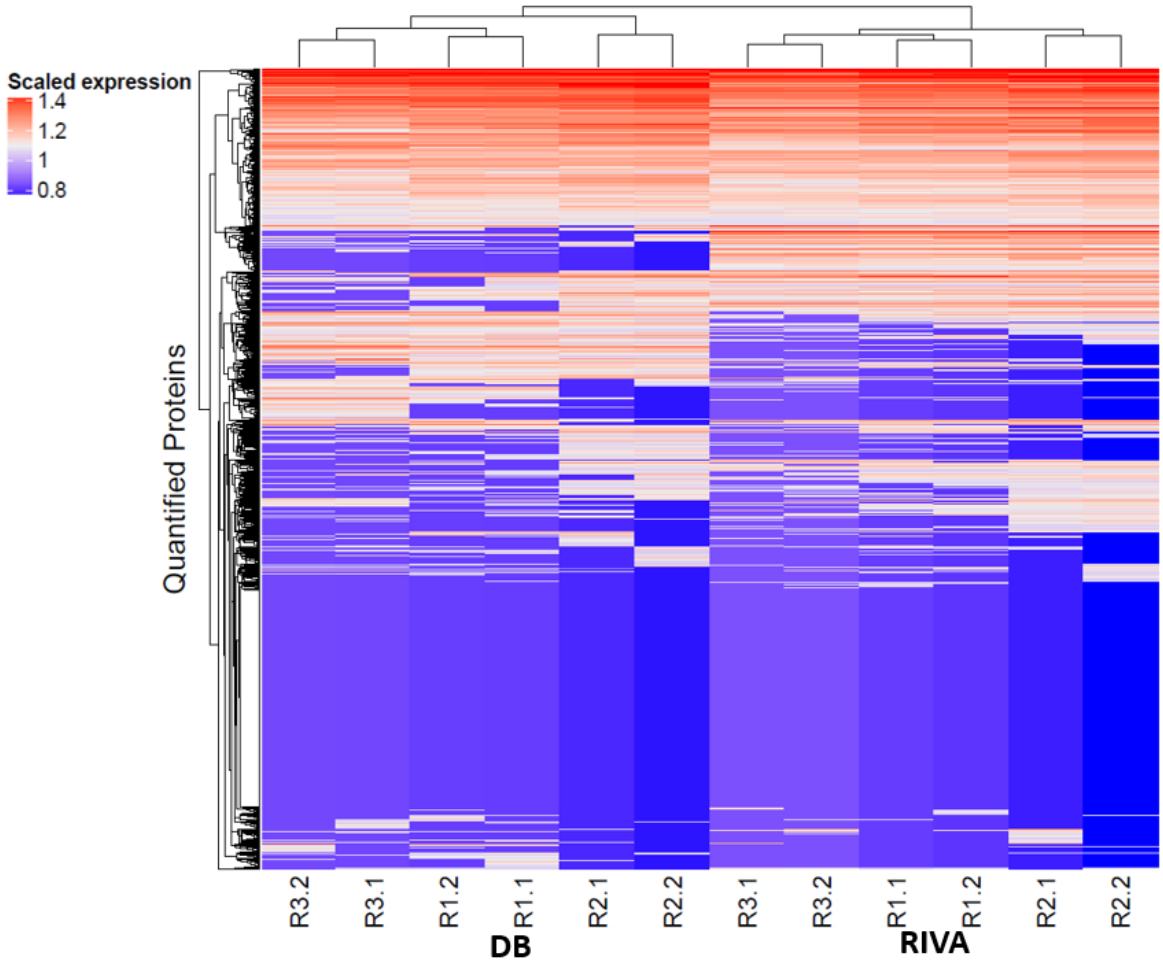

Figure 15. Proteome regulation of DB and RIVA small extracellular vesicles. *Quantitative protein expression profiles (scaled) of diffuse large b cell lymphoma cell lines. Two technical replicas for each of the 3 biological replicas were analyzed by mass spectrometry. High similarity is observed across replicates. The proteomes of diffuse large b cell lymphoma small extracellular vesicles do not seem to overlap as much as the cellular proteomes.*

A way to test the reproducibility of our proteomics profiling approach is to study Pearson's coefficients comparing our conditions (DB and RIVA cells and sEVs) and replicates. Pearson's coefficients range from -1 to 1, for which the latter represents highest similarity between conditions[142]. We calculated and compared using unsupervised hierarchical clustering the correlation coefficients across DB and RIVA cellular and sEVs proteomes, in Figure 16. Correlations between replicas clustered together, showing 0.8-1 correlation coefficients. DLBCL cellular and sEVs fractions also clustered together, suggesting the cellular proteomes are more identical to each other than to their own sEVs fractions. Comparing cellular proteomes, DB and RIVA replicates show 0.4-0.6 coefficients while comparing their sEVs proteomes results in higher coefficients (0.6-0.8). Inferior correlation coefficients were obtained comparing sEVs proteomes to the cell of origin proteomes. DB and RIVA sEVs proteomes appear more like each other than to their own cellular proteomes.

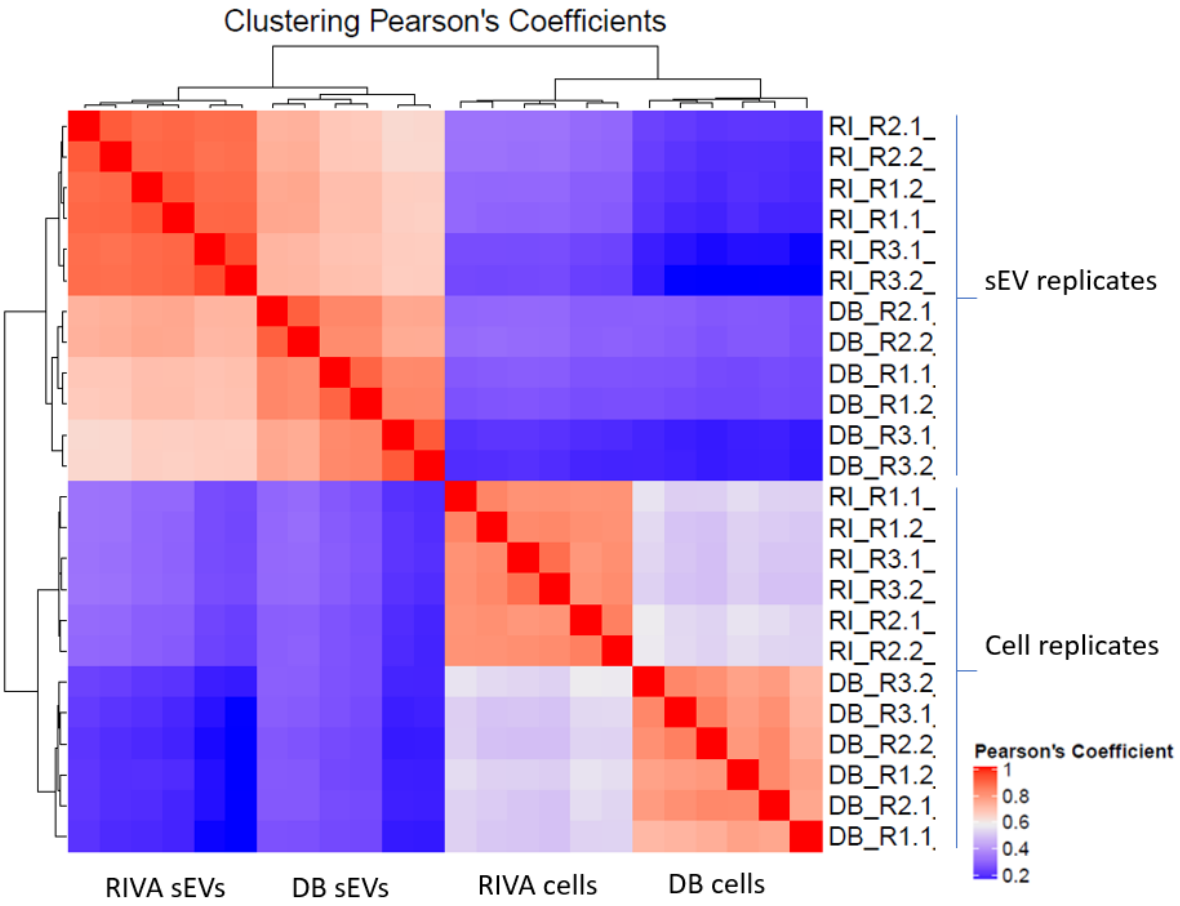

Figure 16. Hierarchical clustering of Pearson's coefficients. *Unsupervised hierarchical clustering of Pearson's coefficients comparing proteome replicas of diffuse large b cell lymphoma cells and small extracellular vesicles (sEVs). Correlation between replicates showed high similarity (0.8-1). Correlations between DB and RIVA sEVs proteomes cluster together (left), showing 0.6-0.8 coefficients. Correlations between DB and RIVA cellular proteomes cluster together (right), with 0.4-0.6 coefficients. DB and RIVA sEVs proteomes appear closer to each other than to their respective cellular proteomes. Our proteomic analysis seems highly reproducible.*

After collapsing protein isoforms into encoding genes, we found the expression - at protein level - of 5020 genes. Figure 17A shows the total proteins identified (proteome complexity) in DB and RIVA cells and sEVs as well as unique proteins found within each fraction (cells or sEVs). The cellular proteome of DB is more varied than that of RIVA, echoing in the number of unique proteins. DLBCL sEVs carry about half their cell of origin proteome complexity, however they hold several (sEVs) unique proteins. Figure 17B shows the dispersion of (average) protein expression quantified in DB and RIVA cellular and sEVs fractions. The cellular proteomes display wider range of quantitation compared to sEVs proteomes. Given sEVs preparations are end products to extensive enrichment and purification protocols, these results suggest we achieved good representation of DB and RIVA cellular and sEVs proteomes.

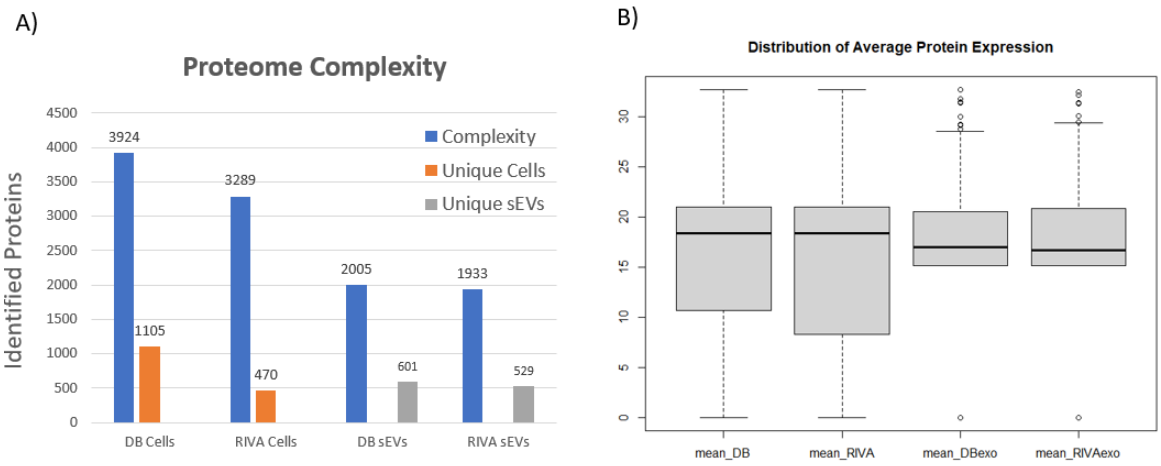

Figure 17. Proteome complexity and average quantitation distribution. *(A) Total number of identified proteins (complexity) in DB and RIVA cells and small extracellular vesicles (sEVs) and unique proteins found within each fraction. DB cellular proteome shows higher complexity, echoing in the number of unique proteins. sEVs proteomes carry about half the complexity of their cell of origin proteomes, but several unique proteins (not found on the other sEVs). (B) Distribution of average protein quantitation in DB and RIVA cells and sEVs. Small dispersion in protein quantitative values in sEVs fractions suggest an effective representation of cellular and sEVs proteomes.*

4.4. DLBCL cellular proteomes

Next, we performed principal component analysis (PCA) to study the quantitative, multivariate dataset of DB and RIVA cellular proteomes. PCA can be used to reduce dimensionality of large datasets, such as OMICs outputs, while retaining as much variation possible from the original dataset. Briefly, dimensionality reduction involves linear transformation of a quantitative matrix into a new set of variables (principal components), ordered according to the amount of variation they represent in the original data set[143]. Using the R package FactomineR[137], which performs pre-centering and scaling of the quantitative matrix, we explored differences and similarities between DLBCL cellular proteomes. In Figure 18A we plot the cumulative variance (%) - black dotted line - that is explained by the first 10 principal component dimensions generated by PCA – in light blue. Together, the first and second dimensions represent around 57% of the variation between cellular proteomes. In Figure 18B is the PCA representation of our individuals – DB and RIVA replicates – on the first 2 principal component dimensions. DB and RIVA cells show good separation on the first dimension. To gain perspective into which proteins are most responsible for the separation between DB and RIVA cell proteomes we plotted the quantified proteins fold change versus the p-value (from t-test comparison) of their mean expression values (Figure 19). The volcano plot, comparing cellular proteomes, shows in the right RIVA upregulated proteins, while in the left DB upregulated proteins.

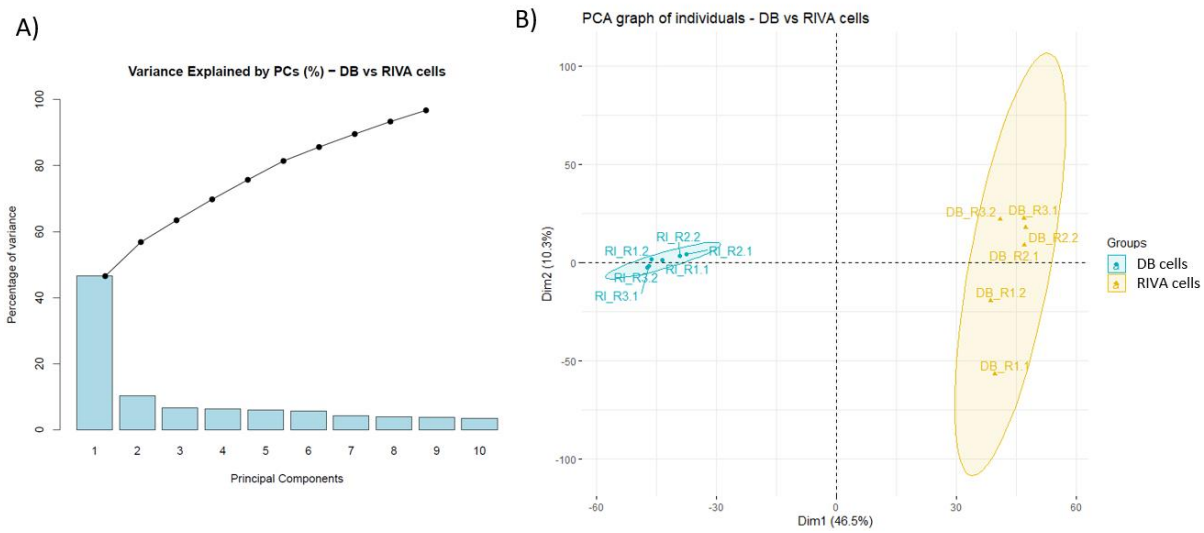

Figure 18. Principal component analysis of DB and RIVA cellular proteomes. *Dimensionality reduction was performed on the quantitative dataset of DB and RIVA cellular proteomes using principal component analysis (PCA). (A) Cumulative variance (%) explained by the 10 first principal component dimensions generated by PCA. First and second PCA dimensions represent ~57% of total variation in DB and RIVA cellular proteomes. (B) PCA space representation of DB (right) and RIVA (left) cellular replicates on the first 2 PCA dimensions. Good separation between DB and RIVA cell proteomes is seen on the first dimension. Using k-means clustering we found two main clusters that separate DB (yellow) and RIVA (blue) cell replicas. Elipses represent 95% confidence cluster intervals.*

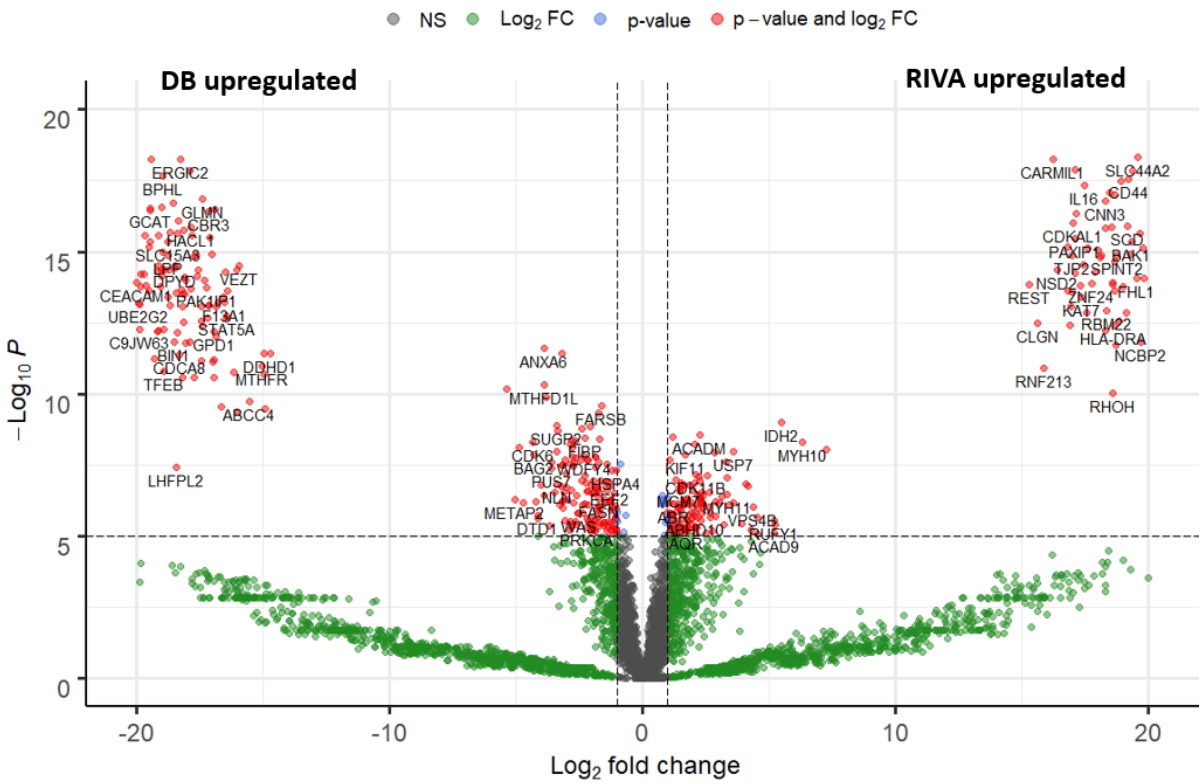

Figure 19. Volcano plot of DB and RIVA cells differential protein expression. *We plotted the difference of mean ratios (log2(fold change)) versus the t-test significance of mean ratios (log10(p-value)) comparing DB and RIVA cellular proteomes. RIVA upregulated proteins (on the right) such as IRF4, IL16 and HDAC21 are directly involved in the activated b cell like (ABC) subtype-specific molecular phenotype.*

DB upregulated proteins (on the left) such as TP53 and MEF2B strongly associate to the germinal center b cell like (GCB) subtype.

To gain further insight on the biological differences between DB and RIVA cells we used FunRich software[134] to perform functional analysis of their expressed proteins. Based on the cellular proteins identified (Figure 20A), cellular component show even enrichment in cytosolic and nuclear proteins, and in proteins annotated to the plasma membrane and main intra-cellular organelles (Figure 20B). Focusing on enriched biological processes, both cell models express proteins associated to immune function, such as innate immune responses and neutrophil degranulation and to the activity of RNA polymerase II and the spliceosome (Figure 21A). Of notice is the negative regulation of the apoptotic process, a shared hallmark by GCB and ABC subtypes. We questioned whether DB and RIVA unique (cellular) proteins could reveal information on their DLBCL subtype-specific biological activity. Figure 21B shows RIVA unique proteins are enriched in BCR signaling components and interferon-gamma signaling, in accordance with RIVA being a cell model of the activated B cell like (ABC) DLBCL subtype.

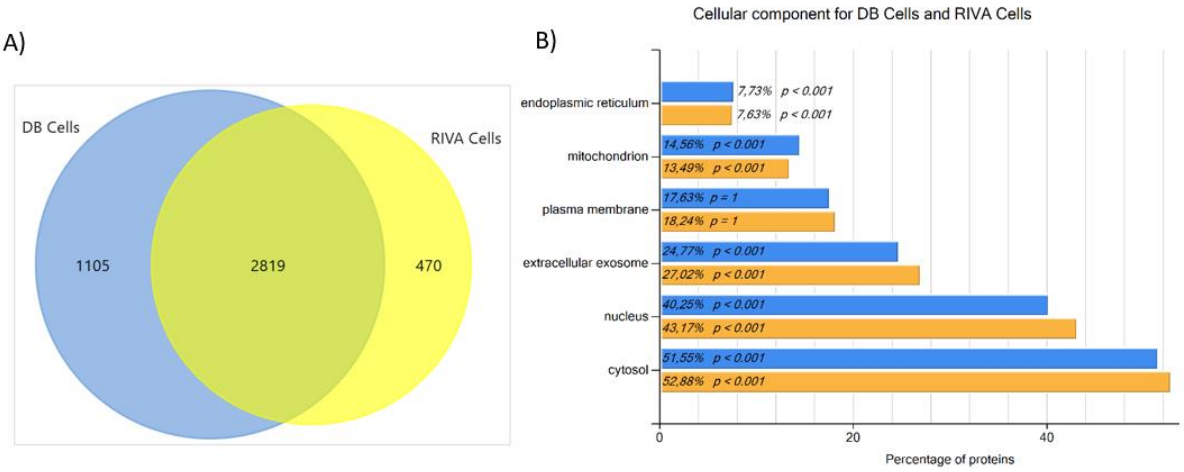

Figure 20. Enriched cellular components in DB and RIVA cellular proteomes. *(A) Overlap between proteins identified in DB and RIVA cellular fractions. (B) Both cell lines show similar enrichment in proteins annotated to main cellular locations (nucleus, cytosol, PM) and intracellular organelles (Mitochondrion, ER, Exosomes).*

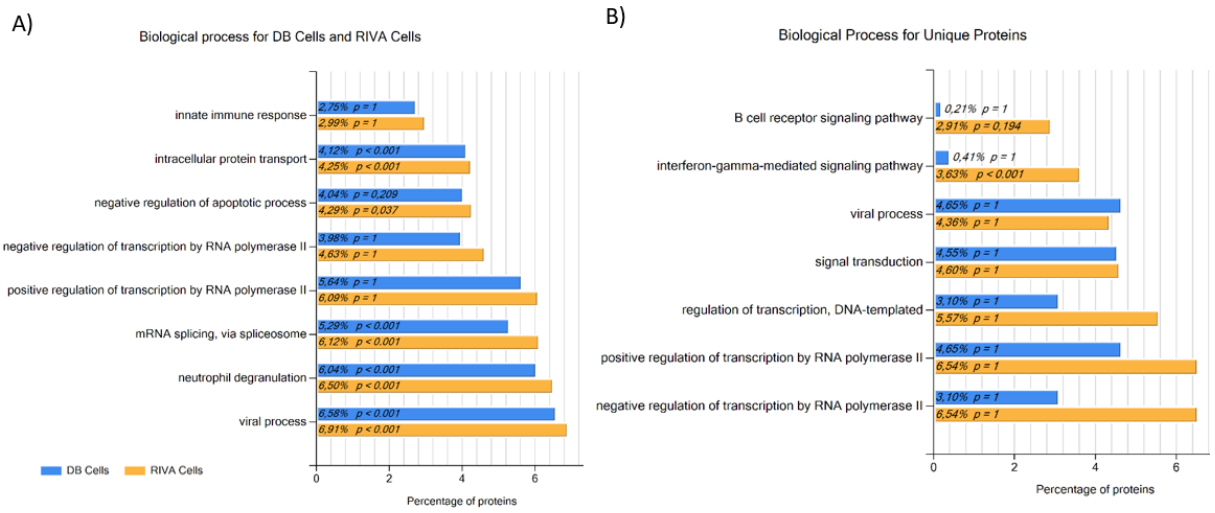

Figure 21. Enriched biological processes in DB and RIVA cellular proteomes. *(A) DB and RIVA cell proteomes share enrichment in proteins associated to immune function (innate immune response and neutrophil degranulation) and to the activity of RNA pol II and spliceosome. Two latter are involved in antibody diversification that occurs in germinal centers, physiological origin of diffuse large b cell lymphomas. (B) Focus on DB and RIVA unique (cellular) proteins revealed B cell receptor signaling and interferon-gamma mediated signaling enrichment, supporting the annotation of RIVA cells as activated B cell like (ABC) diffuse large b cell lymphoma.*

4.5. The proteome of DLBCL sEVs

In this section we report the proteins identified in DLBCL sEVs. We start by looking at the overlap between the proteomes of DLBCL cells and sEVs, in Figure 22A. All conditions contain unique proteins - not detected on any of the others. DLBCL cells and sEVs share numerous proteins (1177) while some were only present in RIVA fractions (69) or exclusively found in DB fractions (80). Further, we questioned if proteins expressed in DLBCL sEVs had previously been reported as secreted on EVs. Figure 22B shows that many proteins detected in both DB and RIVA sEVs (1339/1404) had already been reported in Vesiclepedia[144,145] database. We identified 177 proteins not yet described to be present in human EVs.

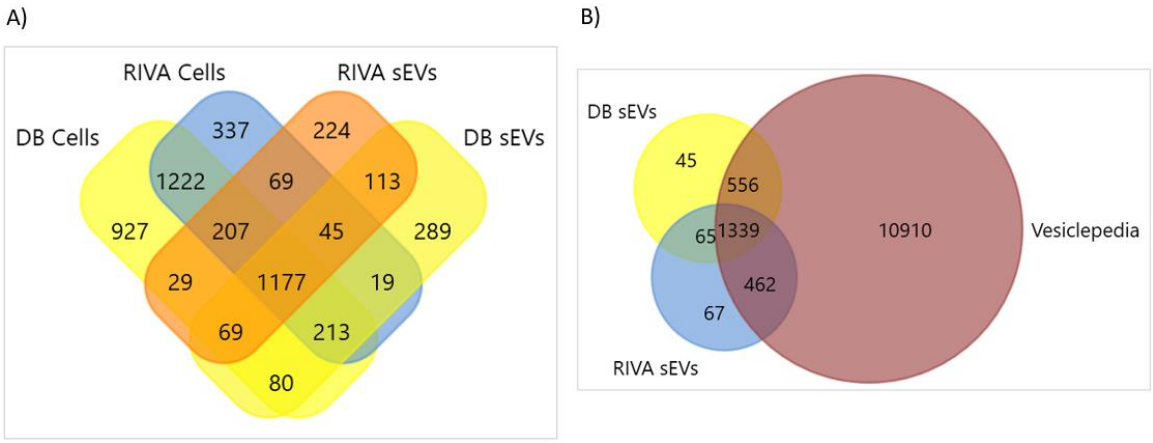

Figure 22. Overview of DLBCL small extracellular vesicle proteomes. *(A) Overlap between the proteomes of DB and RIVA cells and small extracellular vesicles (sEVs). 80 proteins were uniquely found in both DB fractions and 69 in both RIVA fractions. (B) Overlap between DB and RIVA sEVs proteomes and the Vesiclepedia database. 65 proteins found in both DB and RIVA sEVs were not previously identified in*

human extracellular vesicles (EVs) raising interest in their use to purify diffuse large b cell lymphoma sEVs from heterogeneous EVs populations.

Proteins identified in DB and RIVA sEVs (and cells) were compared to the publicly available EV compendium Exocarta[146] to examine the expression of frequently reported EV proteins. In Figure 23A we display the overlap between Exocarta's top 100 most reported exosome proteins and all identified proteins in DLBCL cells and sEVs. DLBCL sEVs carry 7 proteins from Exocarta's top 100 that were not found in their cellular proteomes. These (CD63, THBS1, ANXA4, MVP, ITGA6, LGALS3BP and MFGE8) all function in the endocytic pathway or in cell adhesion processes. To gain additional insights we plotted the average expression of those 100 proteins in DB and RIVA cells and respective sEVs (Figure 23B). CD9, a frequently reported sEVs marker was enriched only in RIVA sEVs. Additionally, CD81 appears to be inversely regulated in sEVs compared to its cellular expression. Given these results, we propose the quantification of several proteins, in parallel, to improve the accuracy of sEVs characterization approaches.

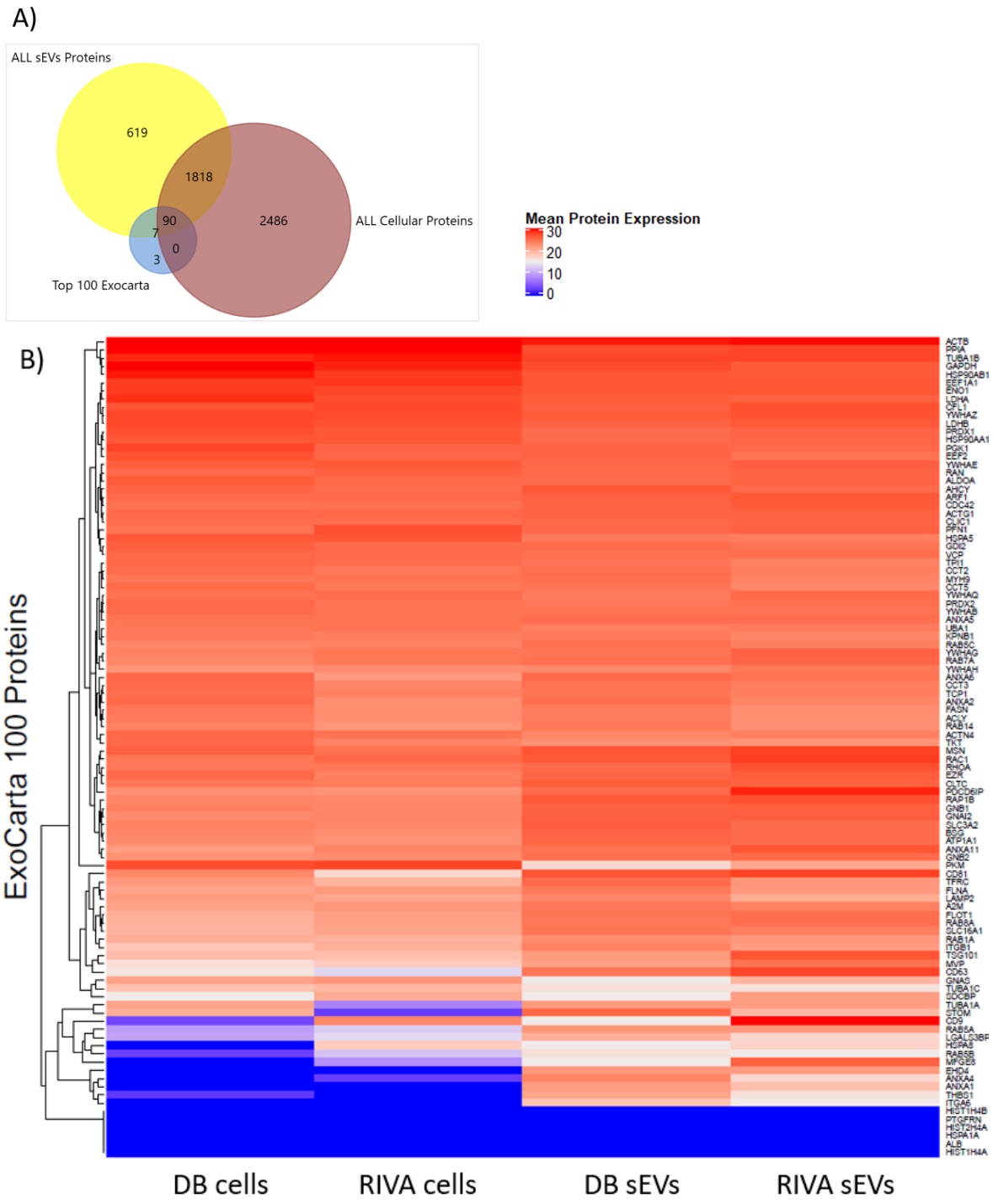

Figure 23. Overlap and quantitative expression of ExoCarta top 100 proteins. *(A) Overlap between the combined proteomes of diffuse large b cell lymphoma cells and small extracellular vesicles with the 100 most frequently reported exosome proteins. We identified 7 proteins exclusively in sEVs fractions. These proteins associate to endosomal trafficking and/or cell adhesion processes. (B) Average quantitation of Exocarta's top 100 proteins in DB and RIVA cells and sEVs.*

We interrogated (using FunRich) the combined proteomes of DLBCL cells, and of sEVs, for their most represented biological functions and cellular components, in Figure 24. DLBCL sEVs seem enriched in proteins annotated to the extracellular space, extracellular exosomes, and plasma membrane (Figure 24A). In contrast, proteins annotated to mitochondrion and nucleus seem

better represented in the cellular proteomes. Figure 24A confirms DLBCL cells express proteins involved in viral infection and in the regulation of spliceosome function and RNA polymerase II activity. Combining both DB and RIVA cell proteomes confirms DLBCL cells retain a germinal center B cell molecular phenotype. B cell receptor signaling, and overall signal transduction were well represented in sEV proteomes (Figure 24B). It is possible that DLBCL cells interchange via sEVs these fundamental protein intervenients and/or regulators of B cell biology. Enrichment in proteins associated to neutrophil degranulation and leucocyte migration on DLBCL sEVs (Figure 24B) is worth attention since the interplay between tumor and other immune cells need better understanding. Overall, DLBCL sEVs seem to carry proteins involved in fundamental B cell processes and in their intercellular communication to other immune cells.

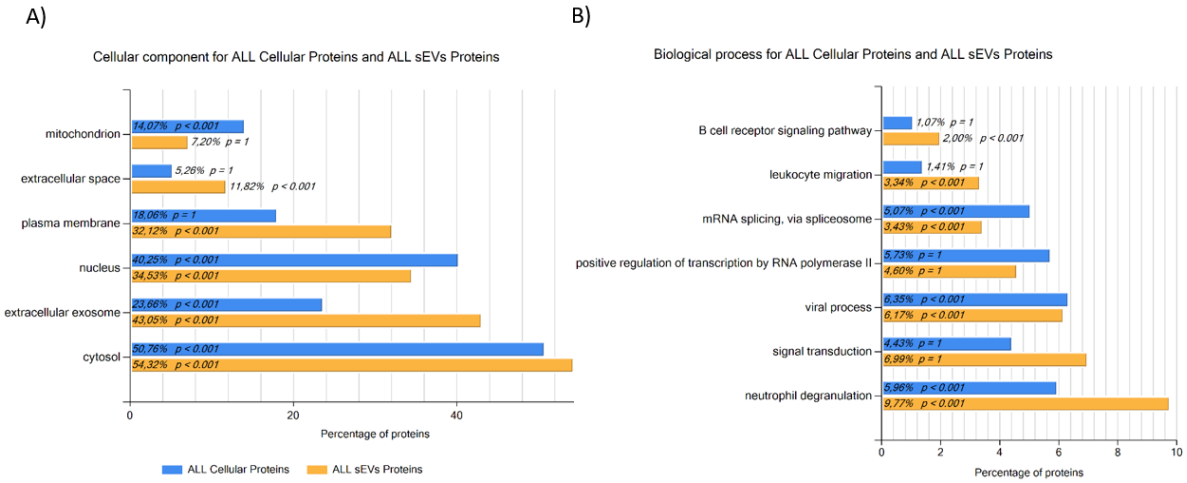

Figure 24. Functional analysis of the combined cellular and sEVs proteomes. *(A) Enriched cellular components in diffuse large b cell lymphoma cells and sEVs proteomes. sEVs proteins were highly annotated to the extracellular space and exosomes, whereas mitochondrion and nucleus were best represented in the cellular proteomes. (B) Enriched biological processes in cells and sEVs. B cell receptor signaling, and signal transduction were well represented in sEVs proteomes. Proteins related to intercellular communications with other immune cells (leukocytes and neutrophils) were enriched in sEVs.*

PCA was performed, as before, now comparing the quantitative expression dataset measured in DLBCL sEVs. In Figure 25A we plot the PCA representation of the individuals – DB and RIVA sEVs replicates – on the first 2 principal component dimensions. The first dimension generated by PCA is responsible for ~37% of the variance between DB and RIVA sEVs proteomes (Figure 25B). However, it mainly differentiates sEVs replicates (from one another) and not DB and RIVA sEVs, for which the separation is conserved along Dim1. It seems the second principal component dimension (dim2) separates DB and RIVA sEVs. As before, Figure 26 displays the volcano plot of fold change versus p-value comparing mean protein expression values measured in DB and RIVA sEVs. Among interesting protein regulation in sEVs, we highlight CD19 (B cell marker) and NOTCH1, highly expressed in RIVA sEVs. BCR complex and transduction (protein) components, were also quantitated in DLBCL sEVs.

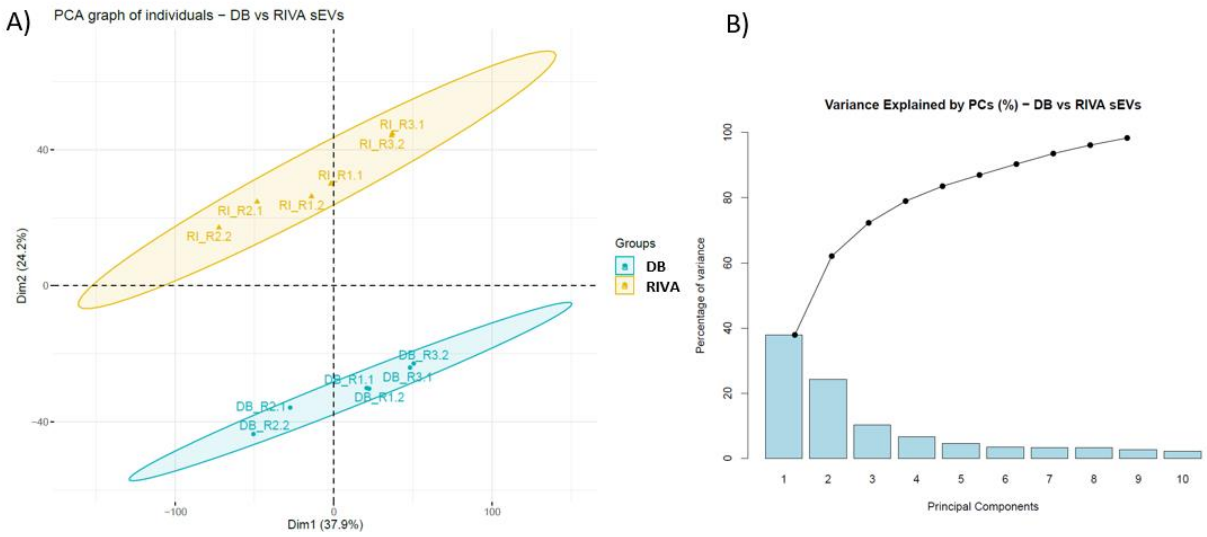

Figure 25. Principal component analysis of DB and RIVA small extracellular vesicles proteomes. *Dimensionality reduction was performed on the quantitative dataset of DB and RIVA small extracellular vesicles (sEVs) proteomes using principal component analysis (PCA). (A) PCA space representation of DB (below) and RIVA (above) sEVs replicates on the first 2 PCA dimensions. Separation between DB and RIVA sEVs proteomes is best described on the second dimension (dim2) Using k-means clustering we found two main clusters that separate DB (blue) and RIVA (blue) sEVs replicas. Ellipses represent 95% confidence cluster intervals. (B) Cumulative variance (%) explained by the 10 first principal component dimensions generated by PCA.*

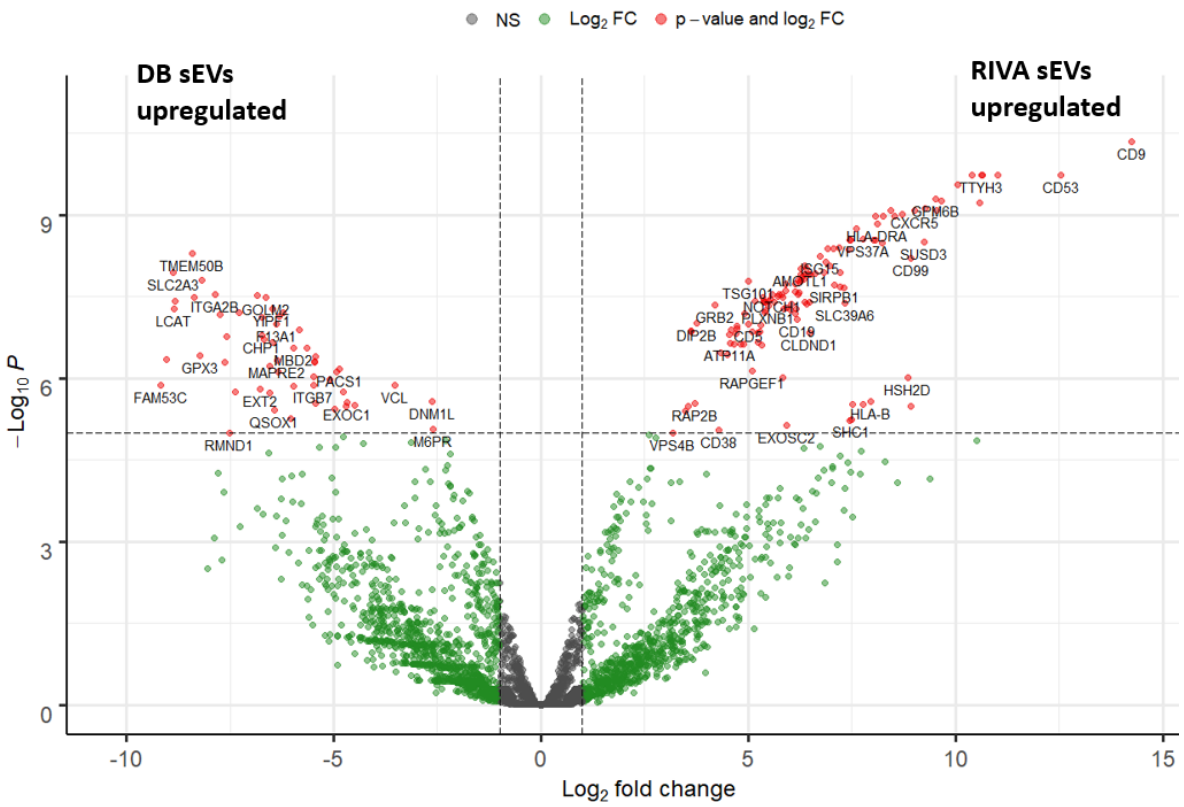

Figure 26. Volcano plot of DB and RIVA small extracellular vesicles differential protein expression. *We plotted the difference of mean ratios (log2(fold change)) versus the t-test significance of mean ratios (log10(p-value)) comparing DB and RIVA small extracellular vesicles (sEVs) proteomes. RIVA upregulated proteins (on the right) include common EVs markers CD9 and TSG101. Some Integrins and HLA proteins seem differentially expressed in DB and RIVA sEVs.*

We used FunRich, this time to explore which categorical biological processes were represented in DB and RIVA sEVs proteomes (Figure 27A). First, we focused on common proteins identified in both DLBCL sEVs (Figure 27B). The common sEVs proteome showed enrichment in membrane and protein trafficking, as well as in infection associated processes. Also, protein translation and RNA catabolic processes seem enriched in DLBCL sEVs. Previously observed enrichment in proteins associated to neutrophil degranulation seems shared by DB and RIVA sEVs. Analysis of unique proteins detected in (each) DB and RIVA sEVs showed that signal transduction is highly represented (Figure 27C). Given that these are (sEVs) unique proteins, DB and RIVA sEVs may, hypothetically, engage fundamentally distinct pathways in their recipient cells. DB sEVs unique proteins were best annotated to the regulation of cell adhesion and extracellular matrix organization, compared to RIVA. In contrast, RIVA sEVs unique proteome show enrichment in proteins associated to exogenous antigen presentation, interferon-gamma signaling and cell migration, compared to DB. DLBCL sEVs comprise functional groups of proteins that inform on their cell of origin molecular phenotypes.

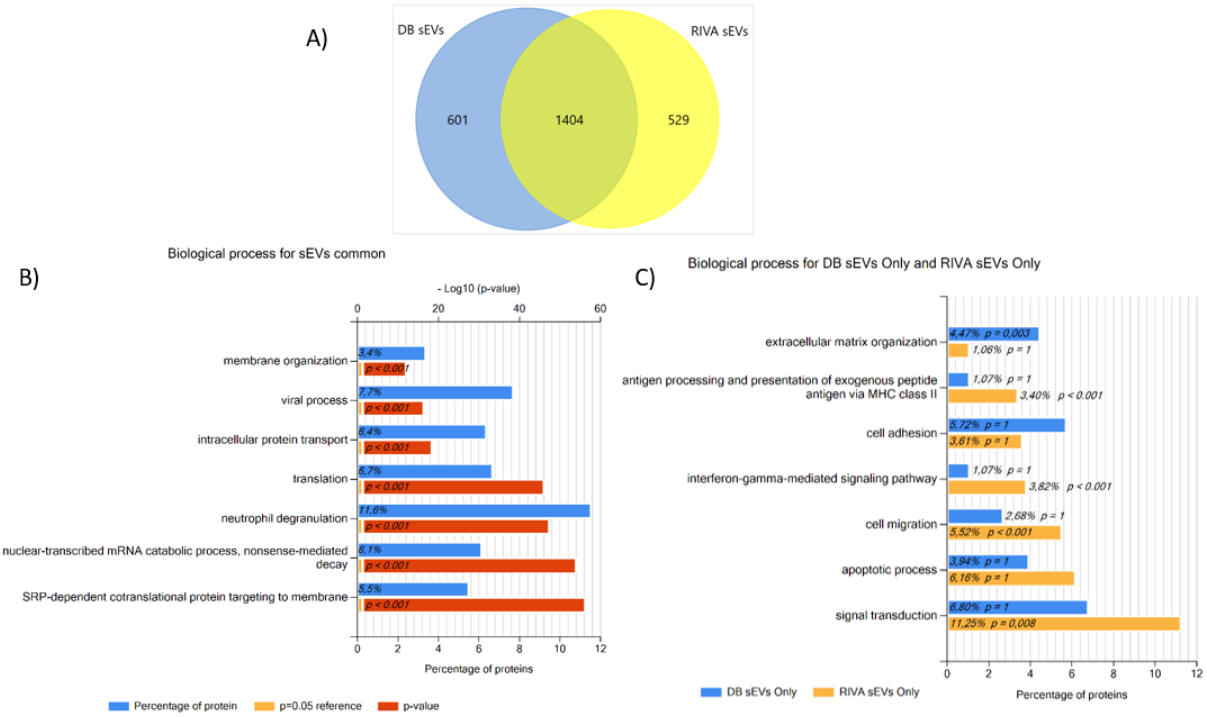

Figure 27. Enrichment analysis of DB and RIVA small extracellular vesicles proteins. *(A) Overlap between proteins identified in DB and RIVA small extracellular vesicles (sEVs). (B) Enriched biological processes in the sEVs common proteome (1404). Enrichment in proteins associated with neutrophil degranulation in shared between the 2 proteomes. (C) Enriched biological processes in DB and RIVA sEVs proteomes. Signal transduction is a well-represented category in sEVs unique proteins. RIVA sEVs unique proteins were best annotated to exogenous antigen presentation, interferon-gamma signaling and cell migration. DB sEVs unique proteins were best annotated to cell adhesion and extracellular matrix organization.*

4.6. sEVs proteomes reflect DLBCL cell-of-origin subtypes

In this section we question whether protein expression profiles of isolated sEVs provide enough evidence as to identify the biological subtype (GCB- or ABC-DLBCL) of their cell line of origin. We start by using linear regression analysis to understand if our conditions (DB and RIVA cells and respective sEVs) are linearly correlated based on their global (average) protein expression. In Figure 28 we show linear regression results comparing DB and RIVA cellular and sEVs proteomes. The coefficient of determination (r_value) provides a measure of how well the variation in the observed data is explained by the linear model. Quantitative proteomes of DB and RIVA sEVs correlate the most with r_value of ~0.73, while DB and RIVA cellular proteomes also correlate positively, but not as substantially (r_value ~0.55). Comparing DLBCL cellular proteomes to their respective sEVs showed positive correlation, but significantly inferior to previous assessments (0.34 - 0.42). In addition, extent of correlation was approximately the same when comparing sEVs proteomes to the non-respective cell of origin. These results suggest that DLBCL sEVs proteomes are, to some extent, shared, and thus to uncover their cell of origin biological subtypes one should focus on differentially expressed proteins.

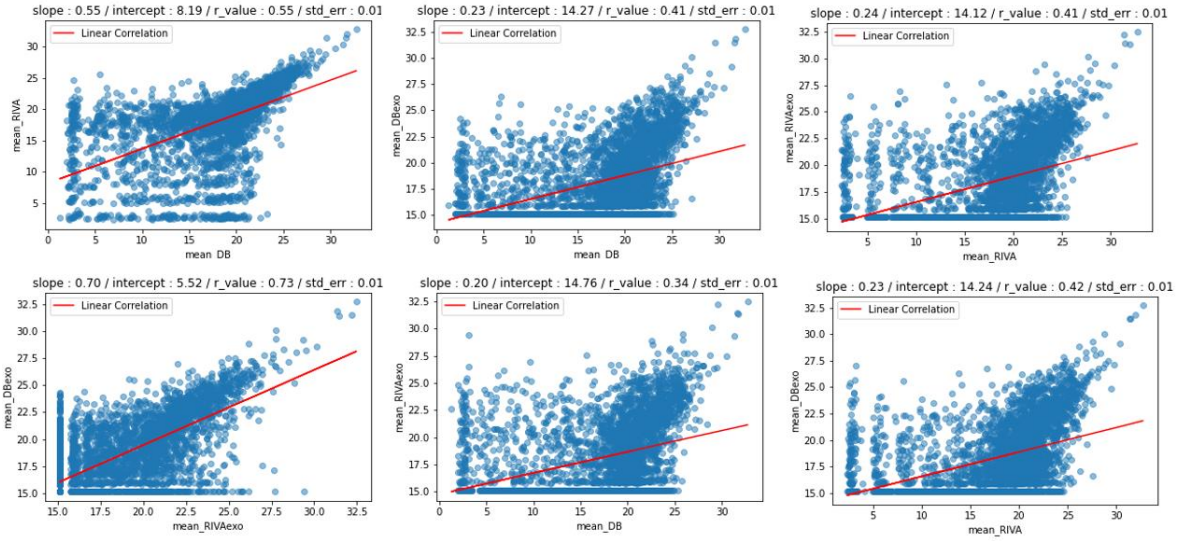

Figure 28. Linear regression analysis of DB and RIVA cellular and small extracellular vesicles proteomes. *Comparing the average protein expression in DB and RIVA cells and small extracellular vesicles (sEVs) shows positive correlation. DB and RIVA sEVs proteomes show the highest coefficients of determination (r_value). DB and RIVA cells show r_value of 0,55. Protein expression in sEVs show similar extent of correlation when compared to their own cellular proteome or to the other cell line proteome.*

Side by side in Figure 29 are the protein expression (average) profiles quantified in DB and RIVA whole-cells and sEVs fractions. We performed PCA, now including the full quantitative proteome dataset of DLBCL cells and sEVs. Representation of DB and RIVA cells and sEVs replicates in the PCA space can be seen in Figure 30A. DLBCL cells (on the left) and sEVs (on the right) replicates were nicely separated in the two first dimensions generated by PCA. The first dimension seems to distinguish DB from RIVA, and the second separates cells from sEVs. Because now we have all DLBCL cellular and sEVs replicates altogether, we speculate that proteins most responsible for separation along dim.2 should concern to DLBCL biological (or cell type specific) phenotypes. Meaning that these proteins may work as sEVs protein biomarkers to predict the biological subtype of their DLBCL cell of origin.

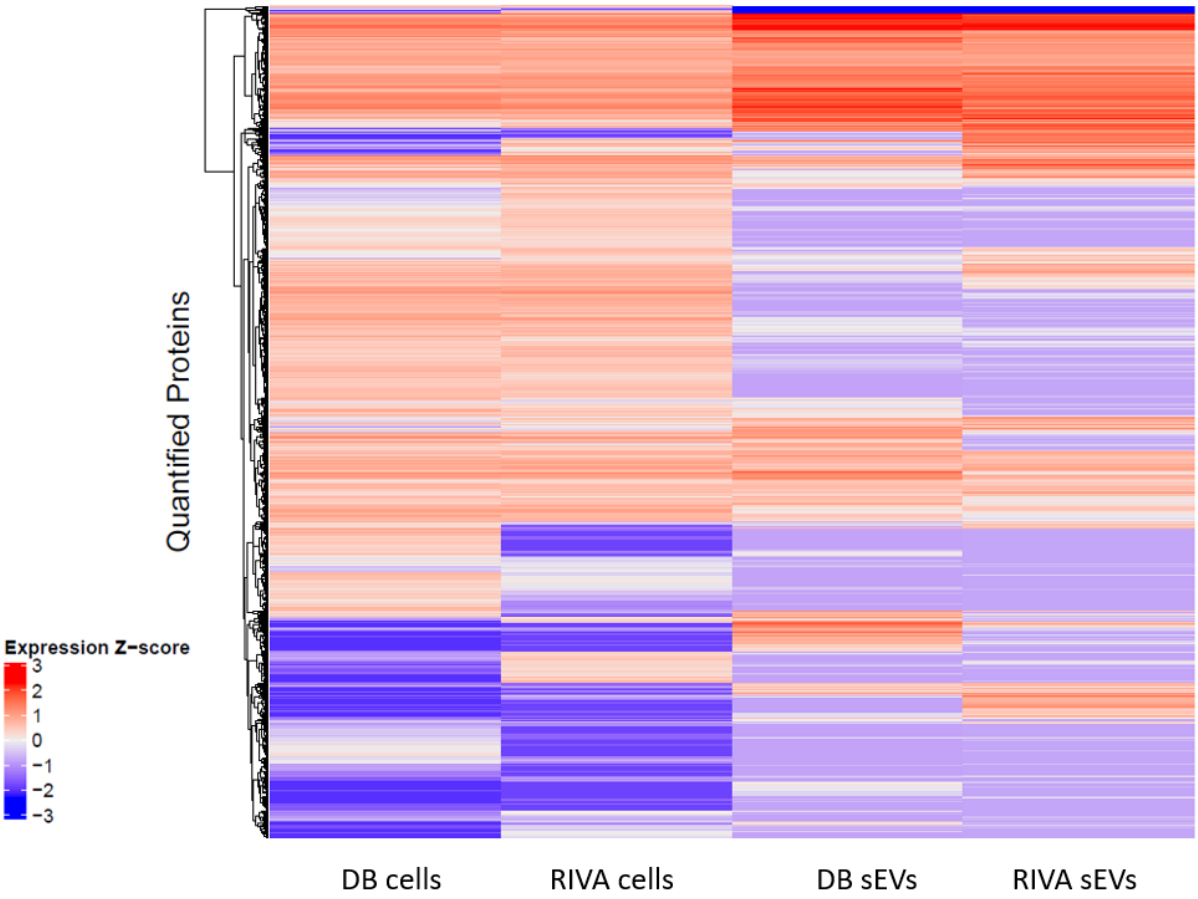

Figure 29. Proteome regulation of DB and RIVA cells and small extracellular vesicles. *Average quantitative protein expression profiles (Z-scored) of diffuse large b cell lymphoma cells and small extracellular vesicles (sEVs).*

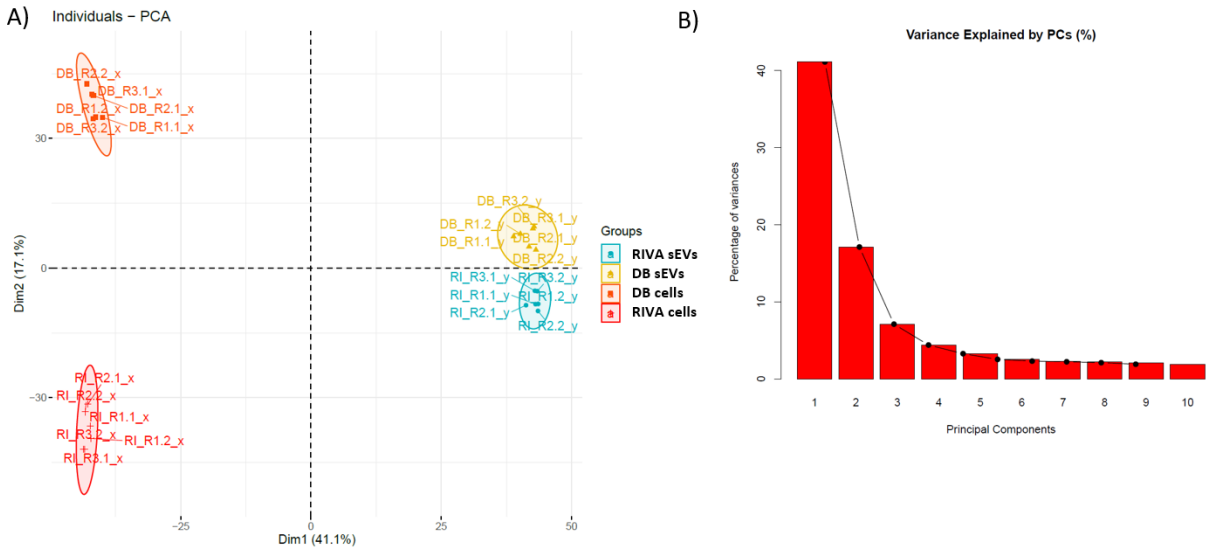

Figure 30. Principal component analysis of DB and RIVA cellular and small extracellular vesicles proteomes. *Dimensionality reduction was performed on the full dataset of DB and RIVA cellular and small*

extracellular vesicles (sEVs) proteomes using principal component analysis (PCA). (A) PCA space representation of DB and RIVA cellular and sEVs replicates on the first 2 PCA dimensions. All replicates clustered together. Using k-means clustering we found 4 main clusters that separate DB (yellow) and RIVA (blue) sEVs replicas on the right, from the DB (orange) and RIVA (red) cells on the left. Ellipses represent 95% confidence cluster intervals. The first dimension separates cells from sEVs, while the second dimension separates DB from RIVA. (B) Cumulative variance (%) explained by the 10 first principal component dimensions generated by PCA.

We wanted to address, in a quantitative manner, how well DLBCL sEVs protein contents reflect their cell of origin biological properties. Functional analysis, as performed before, provide useful information on which categorical processes (or components) are best represented. But it does not fully consider the quantitative nature of proteomic datasets together with the enrichment of the functional groups. This sometimes result in over (or under) representation of functional groups or in exclusion of non-filtered entities (proteins) from biological categories. In addition, it cannot combine distinct types of functional categories (at once). To address this, "complete functional regulation analysis" combines functional and enrichment analysis to compare several categorical types (in parallel) best represented by the proteins identified[129]. We performed complete functional regulation analysis of DLBCL cellular and sEVs proteomes focusing on the comparison DB versus RIVA. In Figure 31, we plot the functional regulation profiles of DLBCL cells and sEVs. The heatmap describes the direction of functional regulation (green for DB upregulated and red for RIVA upregulated) based on the quantitative proteomes of DLBCL cells and sEVs. In many cases, DLBCL sEVs proteomes share the direction of functional regulation (DB vs RIVA) of their cellular proteomes, suggesting sEVs proteins provide insights into their cell of origin molecular phenotypes.

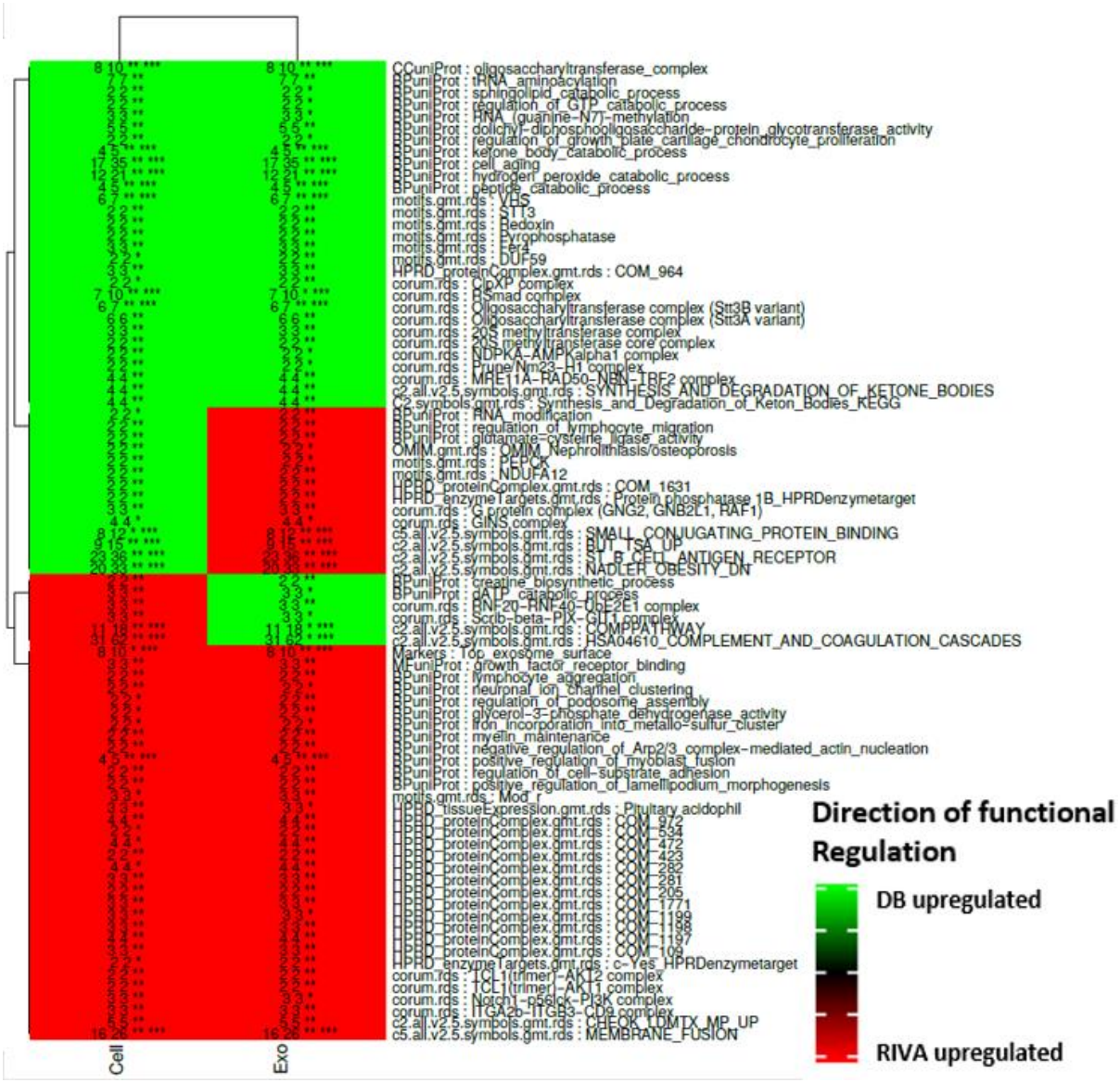

Figure 31. Complete functional regulation analysis of DB and RIVA cellular and small extracellular vesicles proteomes. *Functional regulation profiles of DB and RIVA cells and small extracellular vesicles (sEVs). Direction of functional regulation describes the comparison of DB vs RIVA. Red indicates RIVA upregulated functional categories, while green indicates DB upregulated functional categories. sEVs proteins largely follow the same direction of functional regulation of their cell of origin, suggesting that sEVs protein cargo describes their cell of origin biological properties.*

5. Discussion:

Here we implemented an experimental and computational workflow to study the protein expression profiles of small extracellular vesicles from *in vitro* cultured DLBCL cell lines. Our main question was whether the proteomes of small extracellular vesicles could be used to segregate the prognostic cell of origin subtypes of DLBCL. We selected 4 DLBCL cell line models, the DB and HT cells representing germinal center b cell like (GCB) and RIVA and OCI-ly3 cells as activated b cell like (ABC) DLBCLs. We cultured cells considering suppliers instructions to make sure our results are comparable to others using the same models. We tried to truthfully report our experimental conditions and sEVs purification protocol to prevent possible misinterpretation of

data. To our knowledge this is the first high accuracy, global, quantitative proteomics study to comprehensively compare the proteomes of sEVs derived from DLBCL cells. We considered the expression of other biomolecules in sEVs. However, proteins are main effectors of encoding genes and among the most frequently therapeutically targeted biomolecules. We decided to take advantage of recent developments in mass spectrometry instruments[147] and data analysis strategies[127] to investigate DLBCL sEVs proteomes.

Considering ongoing discussion on efficiency[148] and specificity[149] of sEVs isolation protocols[120,150,151], we expected that differential (ultra)centrifugation (DUC) followed by discontinuous gradient ultracentrifugation would provide decent compromise between particle yield and protein contents. Interestingly, protein concentration estimated after discontinuous gradient ultracentrifugation was significantly inferior to the previous fraction (after DUC). This suggests many proteins found after DUC may not strictly associate with sEVs but possibly result from non-sEVs protein contamination. As to precisely identify DLBCL sEVs proteins, potential tumor biomarkers, we used combination DUC and discontinuous gradient ultracentrifugation.

We cultured 3 DLBCL replicas for each cell line model, in a scale up manner, until 220 mL culture volume is reached to isolate sEVs from conditioned media. The average cell proliferation rates of DLBCL cells did not show differences between subtypes and the replicability of growth rates suggests our scale up culture was reproducible. Adding similar viability between replicas and cell lines we assume our experimental conditions did not promote intracellular (organelle) contamination by excessive cell lysis/death. OCI-ly3 final culturing cell density was half those of other cell lines but showed (consistently) higher sEVs total protein estimates compared to the others. Interestingly, these cells were previously shown to maintain an intra-clonal equilibrium of tumor cell populations *(in vitro),* mediated by sEVs transfer of Wnt signals between DLBCL cells[107]. We noticed they indeed grew in cluster islands, but not exclusively. DLBCL sEVs preparations held similar (total) protein amounts (Figure 10C) and were probed for their biological properties and proteomic cargo.

The efficiency and reproducibility of our sEVs separation protocol and the likelihood proteins identified/quantitated in our sEVs preparations associate with sEVs and not to other co-isolated materials was tested using differential methods. The particle abundance and size distribution in DLBCL sEVs preparations (Figure 11) suggest we achieved an enrichment in small (< 200 nanometers) vesicular particles (Figure 13), consistent with previous studies. Our quantitative results from BCA estimates and NTA particle counts, in sEVs preparations, suggest we obtained reasonable purification efficiency, in line with others[141]. We intend to test the presence of proposed sEVs protein markers as well as negative markers (intracellular components) in our sEVs preparations using immunoblot. However, due to limited sample amount (few µL left per sEVs replica) we await mass spectrometry data for the HT and OCI-ly3 sEVs to ensure we can simultaneously validate some of the proteomics data.

Quantitative protein expression profiles of DB and RIVA cells and sEVs were explored by performing 2 technical LC-MS runs of each biological replicate. These (Figure 14,15) revealed high reproducibility in our profiling approach, confirmed by Pearson's correlation coefficients between expression profiles in DB and RIVA cells and sEVs (Figure 16). We could identify a total of 5020 expressed genes (at the protein level) in DB and RIVA cells and sEVs. We dissected which proteins were expressed in DB and RIVA cells and sEVs, and those exclusively found within each fraction revealing the DB proteome is more varied than that of RIVA (Figure 17A). The number of proteins identified in sEVs was about half those perceived in their cellular proteomes. Being cell secreted organelles (extra-cellular vesicles) likely reasons why sEVs only partially reflect their

full cellular proteome. In addition, the average protein expression distribution (Figure 17B) was similar in DB and RIVA sEVs and less dispersed comparing to their cellular proteomes. Seemingly, we reached good representation of DB and RIVA sEVs (and cellular) proteomes using our experimental settings, isolation protocol, and mass spectrometry-based proteomics profiling.

The protein expression profiles of DB and RIVA cells (as well as HT and OCI-ly3) had previously been reported[121]. Our principal component analysis could also separate the 2 cell line proteomes (Figure 18). We focused on functional analyses of DB and RIVA cellular proteomes as these were not previously reported. We found many proteins associate with immune cell functions and antibody diversification processes (CSR and SHM) typical of germinal center B cells, cellular origin of most non-Hodgkin's lymphomas as DLBCL. Additionally, proteins involved in the negative regulation of the apoptotic process seem enriched in both DB and RIVA proteomes (Figure 21A). The BCL2 antiapoptotic protein was overexpressed in RIVA cells, which carry BCL2 gene amplification[152]. Studying unique (cellular) proteins (Figure 21B) highlighted BCR signaling as enriched biological process in RIVA cells, the model of the activated B cell like DLBCL subtype.

We scrutinized sEVs proteomes to understand if the identification and quantitative regulation of proteins expressed in these compartments can inform on their cellular biology, particularly on their subtype specific (ABC vs GCB) phenotypes. We found several proteins not previously reported in human EVs (Figure 22B). We will report our experimental conditions, sEVs isolation protocols and proteomics data into their fitting knowledgebases, such as EV-TRACK[153] and PRIDE[154]. Newly identified sEVs proteins may serve as potential targets for immunoaffinity-based isolation of DLBCL sEVs from patients' blood, presumably containing sEVs originating from various cell kinds. Studying protein topology[155] in DLBCL sEVs preparations should aid determining the best sEVs proteins to specifically pull-down DLBCL tumor sEVs. Focusing on expression of common EV markers revealed DLBCL sEVs carry many (97/100) frequently reported EVs protein markers. We did not address sEVs biogenesis nor their secretion from DLBCL cells. However, enrichment (Figure 23) in proteins associated to endosomal (and membrane) trafficking such as Rab GTPases (Rab1A, 5A, 5B, 5C, 7A, 8A and 14), tetraspanins (CD9, CD81, CD63) and others (FLOT1, TSG101, MVP, GAPDH, ACTB and CLCT) suggest the endosomal pathway is well represented in our DLBCL sEVs preparations. Albumin (ALB), a common contaminant in EVs preparations was not detected in our DLBCL sEVs.

Functional analysis of the combined cellular and sEVs proteomes (DB and RIVA) nicely reflects our sEVs fractionation strategy (Figure 24). DLBCL cells were enriched in proteins annotated to most intracellular locations and organelles. In contrast, DLBCL sEVs proteomes were best associated to the extracellular space, including extracellular vesicles. The combined (cellular) proteome enrichment in proteins involved in BCR signaling, viral infection and antibody diversification further support germinal center B cells as cell of origin to both cell lines. The identification and enrichment in proteins associated to BCR and overall signal transduction in DLBCL sEVs needs attention. Signal transduction was highly represented in our previous cancer EVs proteomics meta-analysis[115]. The putative roles for the expression of these proteins in DLBCL sEVs can now be explored. In addition, discovery of proteins involved in communication to neutrophils and leucocytes in DLBCL sEVs hint these as candidate acceptor cells for DLBCL sEVs. Given the prognostic value of neutrophil infiltration in DLBCL tumors[156], we propose DLBCL cells may possibly attract and/or sustain their presence and "supporting activity" by sEVs-mediated protein transfer.

DLBCL sEVs quantitative proteomes were compared by PCA and t-test comparison of their mean expression values. We see good separation between sEVs proteomes (DB and RIVA) by PCA

(Figure 25). Data from HT and OCI-ly3 sEVs proteomes should aid define proteins (and their regulation profiles) as potential cell line or subtype (ABC vs GCB) specific biomarkers. With this in mind, we performed functional analysis of DLBCL sEVs proteomes to understand which biological processes were best represented. The common sEVs proteome seemed enriched in components associated to the endosomal pathway, as exemplified by membrane organization, protein trafficking and viral associated processes. We found the previously spotted DLBCL sEVs enrichment in proteins involved in neutrophil degranulation (Figure 24) appears shared by DB and RIVA (Figure 27B) suggesting the interaction of DLBCL cells and neutrophils (via sEVs) may occur in a subtype-independent manner or, hypothetically, DLBCL sEVs may influence distinct cellular pathways in recipient neutrophils. Accordingly, functional analysis of unique sEVs proteomes showed enrichment in proteins involved in signal transduction in both DB and RIVA. We are interested in addressing these questions using functional assays and perhaps more physiological model systems. Antigen presentation, BCR and interferon-gamma signaling (Figure 27C) were enriched in RIVA sEVs, in parallel to their cellular proteome. It seems DLBCL sEVs proteomes provide useful insights into their cell of origin phenotype and into DLBCL intercellular interactions.

We compared, using linear regression analysis, DB and RIVA cellular and sEVs quantitative proteomes to address the extent of linear correlation between them. Linear regression results (Figure 28) confirmed the hierarchical clustering of Pearson's coefficients. DB and RIVA sEVs (quantitative) proteomes correlate better with each other than to their own cellular proteomes. Cellular proteomes (DB vs RIVA) are less (linearly) correlated to each other than their sEVs (DB vs RIVA) suggesting that, overall, sEVs proteome regulation is quite similar. In our previous cancer proteomics meta-analysis[115], we showed that EVs from different cancer types (breast vs lung cancer) correlated better to their own cellular proteome than to those of other cancer types. We could not show this for DLBCL subtypes so far. Consequently, we focused on differentially expressed proteins in cells and sEVs by performing an inclusive principal component analysis. Dimensionality reduction of DB and RIVA cellular and sEVs quantitative proteomes revealed nice separation between cellular proteomes, as well as between sEVs proteomes (Figure 30). Interestingly, DB and RIVA appear separated on the second most representative (17% variability) dimension generated by PCA. To address whether the quantitative proteomes of DLBCL sEVs can inform on the biology of their cell of origin we used an integrative approach described by our group named "complete functional regulation analysis". By combining several functional categories (cellular component, biological process, molecular function…) and acknowledging protein quantitation values this approach allows in-depth comparison between conditions (DB vs RIVA). Our results (Figure 31) show DLBCL sEVs carry proteins associated with the functional categories enriched in their cell of origin proteomes (DB vs RIVA), confirming previous functional analysis results and supporting their potential to inform on DLBCL biology. Thus, we expect complete functional regulation analysis on the 4 DLBCL cell lines (and sEVs) proteomics dataset should provide additional insights into the cell biology underlying DLBCL biological subtypes (ABC vs GCB).

Although proteomics data described informs on the biology of these *in vitro* cultured DLBCL cell lines, our analysis is not complete. Upon receiving mass spectrometry results for the HT and OCI-ly3 cell lines (and respective sEVs) we will re-address statistical and functional analysis on the full dataset. This should provide more accurate assessments into differences between cell lines (and sEVs) belonging to the 2 biological subtypes (ABC vs GCB). Comparison of our full dataset (cellular proteomes) with previously reported[121] DLBCL cellular proteomes should work nicely to validate our proteomics profiling approach. Additionally, and driven by results so far, we plan to

elaborate our analysis by focusing two important protein families in the context of cancer sEVs. Namely, Integrins and antigen presentation proteins (HLA proteins). Integrins regulate sEVs biodistribution under physiological and pathological[105] conditions, being involved in their adhesion to recipient cells and/or tissue extracellular components.[157] We anticipate this will inform into which cells or tissues would preferably receive DLBCL sEVs. Given the involvement of exosomal HLA protein secretion in B cell immunological networks, it is appealing to study their participation in DLBCL, admitting its characteristic immune evasion capacities. The most important follow-up to this study will be to compare our sEVs protein expression data with the proteome profiles of DLBCL patients' blood-derived sEVs. This will address whether circulating sEVs proteins can be used to segregate DLBCL patients into their prognostic biological subtypes. Our exploratory approach, using mass spectrometry, should recognize potential sEVs protein biomarkers to probe in liquid biopsies using cheaper and faster targeted strategies as ELISA. Our long-term objective is to contribute for the development of novel and improved (diagnostic) tools to aid in clinical management of DLBCL and, hopefully, lessen the need for surgical intervention.

In addition, we are interested in further exploring DLBCL biology. Recently, we started a comprehensive subcellular fractionation approach (requiring large initial amounts of cells for MS profiling) to address potential (protein expression) differences in these cellular locations (nucleus, microsomes, mitochondria…). As in previous findings[87], CD20 was present and enriched in our DLBCL sEVs preparations (DB and RIVA). We are working to apply Rituximab (anti-CD20 mAb) in global drug profiling assays[139], starting by comparing sEVs protein expression profiles of treated vs non-treat cells and follow-up with functional studies to better understand sEVs protein secretion in the context of DLBCL treatment response. Intriguingly we detected CD47 protein "don't eat me signal" in both DB and RIVA sEVs proteomes, which may indicate a previously unknown form of immune evasion by DLBCL cells against macrophage phagocytic activity, mediated by sEVs. Also, we identified ENO2 in DLBCL sEVs, a previously reported modulator of macrophage anti-tumor activity, transferred from DLBCL cells to macrophages by sEVs.[113] We definitely highlight the presence of BCR components (CD79A, CD79B, SRC, LYN, B2M, BTK, and Immunoglobulins) in DLBCL sEVs. Interestingly, BCL2 was also highly expressed in RIVA sEVs, in parallel to its cellular expression, which may work to support tumor cells survival in the TME. Our results provide rational for the generation of novel hypothesis to assess the roles of these sEVs proteins in DLBCL pathology. Overall, this study validates protein expression profiling, by mass spectrometry, as a resourceful method to study DLBCL sEVs secreted proteins. Upon completion, we expect to provide the knowledge base with a comprehensive description of DLBCL sEVs proteomes, their association with cell of origin prognostic subtypes and hopefully explore their contributions in DLBCL pathophysiology.

6. Acknowledgements:

I'll start by state my deep gratitude to Rune Matthiesen and Ana Sofia Carvalho for allowing me to work on this project and for your friendship. To Ana, I admire your working capacity and highly appreciate the long hours we spent together in the lab and your continuous availability to help me, and other lab members, progress in our research. To Rune, I'm happy to be a part of your team and I will always remember the early discussions that ended up becoming this master's thesis. I thank you for always believing in me, in my ideas, and trusting me to work independently (after great efforts, by Ana) on cell culture and sEVs isolation and in the data analysis. To make it explicit I think your mentoring and friendship were crucial for my development as a biomedical researcher and, hopefully, I will preserve them.

I'm thankful to the funding, as well as to the CEDOC cell culture facility and all the community for being fundamental parts in this project. To the rest of the CEB lab, specially Andreia and Ivo, thank you for your interest in me, in my work, for your insightful comments, but, mainly for your friendship. To José Ramalho and Paulo Pereira for always providing me with their apparatuses and expertise. Also, to the Systems Oncology group at Champalimaud for valuable help with the NTA. To CEDOC's Histology facility (Ana Farinho) for assistance with sample preparation, and the IGC Electron Microscopy Facility for imaging sEVs. Finally, to Hans Christian Beck at the University of Southern Denmark (SDU), for the mass spectrometry sample analysis.

I would like to thank the NBR's course coordinators (Paulo, Duarte, Rita, Cláudia, Helena, and others) for selecting me to this master program. I believe you are doing a great job in providing future scientists with the best working skills (literature review; teamwork; writing, developing, and presenting scientific work; and many others…) to enhance our potential as biomedical researchers. Also, a thank your note to my NBR colleagues for everything we shared during these times, I really owe you. I hope everyone preserves the love for science that compels us to keep searching, to keep looking, to be curious. To god mode for exquisite resourcefulness and disposition.

Finally, to my family, in particular my brother Pedro, I really hope things will get better, I know I can be. To my mother, thank you. To my friends of the Rijura, we'll keep it up, and always thrive.

7. References:

1. Blackstone, N. W. Molecular Biology of the Cell . Fourth Edition. By Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts , and , Peter Walter. New York: Garland Science. $102.00. xxxiv + 1463 p; ill.; glossary (G:1–G:36); index (I:1–I:49); tables (T:1). ISBN: 0–8153–3218–1. [CD-ROM included.] 2002. . *Q. Rev. Biol.* **78**, 91–92 (2003).

2. Young, R. M., Phelan, J. D., Wilson, W. H. & Staudt, L. M. Pathogenic B-cell receptor signaling in lymphoid malignancies: New insights to improve treatment. *Immunol. Rev.* **291**, 190–213 (2019).

3. Shaffer, A. L., Young, R. M. & Staudt, L. M. Pathogenesis of Human B Cell Lymphomas. *Annu. Rev. Immunol.* **30**, 565–610 (2012).

4. Oettinger, M. A., Schatz, D. G., Gorka, C. & Baltimore, D. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. *Science (80-.).* **248**, 1517–1523 (1990).

5. Alves-Pereira, C. F. *et al.* Independent recruitment of Igh alleles in V(D)J recombination. *Nat. Commun.* **5**, 1–15 (2014).

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

7. Pereira, J. P., Cyster, J. G. & Xu, Y. A role for S1P and S1P1 in immature-B cell egress from mouse bone marrow. *PLoS One* **5**, (2010).

8. Boyce, J. A., Finkelman, F., Shearer, W. T., Pieper, K. & Grimbacher, B. Mechanisms of allergic diseases B-cell biology and development. *J. Allergy Clin. Immunol.* **131**, 959–971 (2013).

9. Arnon, T. I., Horton, R. M., Grigorova, I. L. & Cyster, J. G. Visualization of splenic marginal zone B-cell shuttling and follicular B-cell egress. *Nature* **493**, 684–688 (2013).

10. Allman, D. & Pillai, S. Peripheral B cell subsets. *Current Opinion in Immunology* **20**, 149–157 (2008).

11. Chaplin, D. D. Overview of the immune response. *J. Allergy Clin. Immunol.* **125**, S3 (2010).

12. Willard-Mack, C. L. Normal structure, function, and histology of lymph nodes. *Toxicol. Pathol.* **34**, 409–24 (2006).

13. Cyster, J. G. & Allen, C. D. C. Leading Edge Review B Cell Responses: Cell Interaction Dynamics and Decisions. (2019). doi:10.1016/j.cell.2019.03.016

14. Allen, C. D. C., Okada, T., Tang, H. L. & Cyster, J. G. Imaging of germinal center selection events during affinity maturation. *Science (80-.).* **315**, 528–531 (2007).

15. Barreto, V., Reina-San-Martin, B., Ramiro, A. R., McBride, K. M. & Nussenzweig, M. C. C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion. *Mol. Cell* **12**, 501–508 (2003).

16. Pavri, R. *et al.* Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell* **143**, 122–133 (2010).

17. Casellas, R. *et al.* Mutations, kataegis and translocations in B cells: Understanding AID promiscuous activity. *Nature Reviews Immunology* **16**, 164–176 (2016).

18. Pasqualucci, L. *et al.* AID is required for germinal center-derived lymphomagenesis. *Nat. Genet.* **40**, 108–112 (2008).

19. Roco, J. A. *et al.* Class-Switch Recombination Occurs Infrequently in Germinal Centers. *Immunity* **51**, 337-350.e7 (2019).

20. Pasqualucci, L. Molecular pathogenesis of germinal center-derived B cell lymphomas. *Immunol. Rev.* **288**, 240–261 (2019).

21. Peterson, L. F. *et al.* Acute myeloid leukemia with the 8q22;21q22 translocation: Secondary mutational events and alternative t(8;21) transcripts. *Blood* **110**, 799–805 (2007).

22. Basso, K. & Dalla-Favera, R. Germinal centres and B cell lymphomagenesis. *Nat. Rev. Immunol.* **15**, 172–184 (2015).

23. Küppers, R. The biology of Hodgkin's lymphoma. *Nature Reviews Cancer* **9**, 15–27 (2009).

24. Nakamura, S. *et al.* Translocations involving the immunoglobulin heavy chain gene locus predict better survival in gastric diffuse large B-Cell lymphoma. *Clin. Cancer Res.* **14**, 3002–3010 (2008).

25. Seifert, M., Scholtysik, R. & Küppers, R. Origin and pathogenesis of B cell lymphomas. *Methods Mol. Biol.* **971**, 1–25 (2013).

26. Efremov, D. G., Turkalj, S. & Laurenti, L. Mechanisms of B Cell Receptor Activation and Responses to B Cell Receptor Inhibitors in B Cell Malignancies. *Cancers (Basel).* **12**, 1396 (2020).

27. Conde, L. *et al.* Genome-wide association study of follicular lymphoma identifies a risk locus at 6p21.32. *Nat. Genet.* **42**, 661–664 (2010).

28. Cerhan, J. R. *et al.* Genome-wide association study identifies multiple susceptibility loci for diffuse large B cell lymphoma. *Nat. Genet.* **46**, 1233–1238 (2014).

29. Castillo, J. J. *et al.* Obesity is associated with increased relative risk of diffuse large B-cell lymphoma: A meta-analysis of observational studies. *Clin. Lymphoma, Myeloma Leuk.* **14**, 122–130 (2014).

30. Armitage, J. O., Gascoyne, R. D., Lunning, M. A. & Cavalli, F. Non-Hodgkin lymphoma. *The Lancet* **390**, 298–310 (2017).

31. Mlynarczyk, C., Fontán, L. & Melnick, A. Germinal center-derived lymphomas: The darkest side of humoral immunity. *Immunological Reviews* **288**, 214–239 (2019).

32. Shiga, T. *et al.* Long-Term Observation of the Progression From Nodal Marginal Zone Lymphoma to Diffuse Large B-Cell Lymphoma in a Dog: *https://doi.org/10.1177/0300985820932143* 030098582093214 (2020). doi:10.1177/0300985820932143

33. Young, R. M. *et al.* Taming the Heterogeneity of Aggressive Lymphomas for Precision Therapy. *Annu. Rev. Cancer Biol.* **3**, 429–455 (2019).

34. Bajaj, A. Voluminous and Vesicular: Diffuse Large B Cell Lymphoma (DLBCL). (2018). doi:10.4172/2329-6771.1000215

35. Tilly, H. *et al.* Diffuse large B-cell lymphoma (DLBCL): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* **26**, v116–v125 (2015).

36. Vitolo, U. *et al.* Extranodal diffuse large B-cell lymphoma (DLBCL) and primary mediastinal B-cell lymphoma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **27**, v91–v102 (2016).

37. Li, S., Young, K. H. & Medeiros, L. J. Diffuse large B-cell lymphoma. *Pathology* **50**, 74–87 (2018).

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

39. Davis, R. E., Brown, K. D., Siebenlist, U. & Staudt, L. M. *Constitutive Nuclear Factor B Activity Is Required for Survival of Activated B Cell-like Diffuse Large B Cell Lymphoma Cells*. *The Journal of Experimental Medicine* • **194**, (2001).

40. Davis, R. E. *et al.* Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. *Nature* **463**, 88–92 (2010).

41. Phelan, J. D. *et al.* A multiprotein supercomplex controlling oncogenic signalling in lymphoma. *Nature* **560**, 387–391 (2018).

42. Fisher, R. I. *et al.* Comparison of a Standard Regimen (CHOP) with Three Intensive Chemotherapy Regimens for Advanced Non-Hodgkin's Lymphoma. *N. Engl. J. Med.* **328**, 1002–1006 (1993).

43. Rosenwald, A. *et al.* The Use of Molecular Profiling to Predict Survival after Chemotherapy for Diffuse Large-B-Cell Lymphoma. *N. Engl. J. Med.* **346**, 1937–1947 (2002).

44. Coiffier, B. *et al.* CHOP Chemotherapy plus Rituximab Compared with CHOP Alone in Elderly Patients with Diffuse Large-B-Cell Lymphoma. *N. Engl. J. Med.* **346**, 235–242 (2002).

45. Lenz, G. *et al.* Stromal Gene Signatures in Large-B-Cell Lymphomas. *N. Engl. J. Med.* **359**, 2313–2323 (2008).

46. Rougé, L. *et al.* Structure of CD20 in complex with the therapeutic monoclonal antibody rituximab. *Science (80-.).* **9356**, eaaz9356 (2020).

47. Minard-Colin, V. *et al.* Lymphoma depletion during CD20 immunotherapy in mice is mediated by macrophage FcγRI, FcγRIII, and FcγRIV. *Blood* **112**, 1205–1213 (2008).

48. Mohammed, R., Milne, A., Kayani, K. & Ojha, U. How the discovery of rituximab impacted the treatment of B-cell non-hodgkin's lymphomas. *Journal of Blood Medicine* **10**, 71–84 (2019).

49. Federico, M. *et al.* Rituximab and the risk of transformation of follicular lymphoma: a retrospective pooled analysis. *Lancet Haematol.* **5**, e359–e367 (2018).

50. Ngo, V. N. *et al.* A loss-of-function RNA interference screen for molecular targets in cancer. *Nature* **441**, 106–110 (2006).

51. Wilson, W. H. *et al.* Targeting B cell receptor signaling with ibrutinib in diffuse large B cell lymphoma. *Nat. Med.* **21**, 922–926 (2015).

52. Mondello, P. & Mian, M. Frontline treatment of diffuse large B-cell lymphoma: Beyond R-CHOP. *Hematol. Oncol.* hon.2613 (2019). doi:10.1002/hon.2613

53. Abdulla, M. *et al.* Cell-of-origin determined by both gene expression profiling and immunohistochemistry is the strongest predictor of survival in patients with diffuse large B-cell lymphoma. *Am. J. Hematol.* **95**, 57–67 (2020).

54. Schmitz, R. *et al.* Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma. *N. Engl. J. Med.* **378**, 1396–1407 (2018).

55. Chapuy, B. *et al.* Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. *Nat. Med.* **24**, 679–690 (2018).

56. Venturutti, L. *et al.* TBL1XR1 Mutations Drive Extranodal Lymphoma by Inducing a Pro-tumorigenic Memory Fate. *Cell* **182**, 297-316.e27 (2020).

57. George Wright, A. W. *et al.* A Probabilistic Classification Tool for Genetic Subtypes of Diffuse Large B Cell Lymphoma with Therapeutic Implications. (2020). doi:10.1016/j.ccell.2020.03.015

58. Moens, L. & Tangye, S. G. Cytokine-mediated regulation of plasma cell generation: IL-21 takes center stage. *Frontiers in Immunology* **5**, 65 (2014).

59. Suzuki, K., Grigorova, I., Phan, T. G., Kelly, L. M. & Cyster, J. G. Visualizing B cell capture of cognate antigen from follicular dendritic cells. *J. Exp. Med.* **206**, 1485–1493 (2009).

60. Raposo, G. *et al.* B lymphocytes secrete antigen-presenting vesicles. *J. Exp. Med.* **183**, 1161–1172 (1996).

61. Muntasell, A., Berger, A. C. & Roche, P. A. T cell-induced secretion of MHC class II–peptide complexes on B cell exosomes. *EMBO J.* **26**, 4263–4272 (2007).

62. Fernández-Messina, L. *et al.* Transfer of extracellular vesicle-micro <scp>RNA</scp> controls germinal center reaction and antibody production. *EMBO Rep.* **21**, (2020).

63. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–74 (2011).

64. Scott, D. W. & Gascoyne, R. D. The tumour microenvironment in B cell lymphomas. *Nature Reviews Cancer* **14**, 517–534 (2014).

65. Kumar, D. & Xu, M. L. Microenvironment Cell Contribution to Lymphoma Immunity. *Front. Oncol.* **8**, (2018).

66. Tweeddale, M. E. *et al.* The presence of clonogenic cells in high-grade malignant lymphoma: A prognostic factor. *Blood* **69**, 1307–1314 (1987).

67. Miao, Y., Medeiros, L. J., Xu-Monette, Z. Y., Li, J. & Young, K. H. Dysregulation of Cell Survival in Diffuse Large B Cell Lymphoma: Mechanisms and Therapeutic Targets. *Front. Oncol.* **9**, 107 (2019).

68. Mannino, R. G. *et al.* 3D microvascular model recapitulates the diffuse large B-cell lymphoma tumor microenvironment in vitro. *Lab Chip* **17**, 407–414 (2017).

69. Yao, Z. *et al.* Concordant bone marrow involvement of diffuse large B-cell lymphoma represents a distinct clinical and biological entity in the era of immunotherapy. *Leukemia* **32**, 353–363 (2018).

70. Manček-Keber, M. *et al.* Extracellular vesicle–mediated transfer of constitutively active MyD88 L265P engages MyD88 wt and activates signaling. *Blood* **131**, 1720–1729 (2018).

71. Pasqualucci, L. *et al.* Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat. Genet.* **43**, 830–837 (2011).

72. Zhang, J., Medeiros, L. J. & Young, K. H. Cancer immunotherapy in diffuse large B-cell lymphoma. *Frontiers in Oncology* **8**, 351 (2018).

73. Brudno, J. N. *et al.* Safety and feasibility of anti-CD19 CAR T cells with fully human binding domains in patients with B-cell lymphoma. *Nat. Med.* **26**, 270–280 (2020).

74. Chao, M. P. *et al.* Anti-CD47 Antibody Synergizes with Rituximab to Promote Phagocytosis and Eradicate Non-Hodgkin Lymphoma. *Cell* **142**, 699–713 (2010).

75. Advani, R. *et al.* CD47 Blockade by Hu5F9-G4 and Rituximab in Non-Hodgkin's Lymphoma. *N. Engl. J. Med.* **379**, 1711–1721 (2018).

76. Schwaller, J. *et al.* Neutrophil-derived APRIL concentrated in tumor lesions by proteoglycans correlates with human B-cell lymphoma aggressiveness. *Blood* **109**, 331–338 (2007).

77. Nie, M. *et al.* Neutrophil Extracellular Traps Induced by IL8 Promote Diffuse Large B-cell Lymphoma Progression via the TLR9 Signaling. *Clin. Cancer Res.* **25**, 1867–1879 (2019).

78. Greenbaum, A., Gopal, A. K., Fromm, J. R. & Houghton, A. M. Diffuse Large B-Cell Lymphoma Is Infiltrated with Functional CD8+ T-Cells Lacking the Hallmarks of Exhaustion. *Blood* **134**, 1518–1518 (2019).

79. Meyer, P. N. *et al.* Immunohistochemical methods for predicting cell of origin and survival in patients with diffuse large B-cell lymphoma treated with rituximab. *J. Clin. Oncol.* **29**, 200–207 (2011).

80. Coutinho, R. *et al.* Poor concordance among nine immunohistochemistry classifiers of cell-of-origin for diffuse large b-cell lymphoma: Implications for therapeutic strategies. *Clin. Cancer Res.* **19**, 6686–6695 (2013).

81. Coutinho, R. *et al.* Revisiting the immune microenvironment of diffuse large B-cell lymphoma using a tissue microarray and immunohistochemistry: Robust semi-automated analysis reveals CD3 and FoxP3 as potential predictors of response to R-CHOP. *Haematologica* **100**, 363–369 (2015).

82. Gutiérrez-García, G. *et al.* Gene-expression profiling and not immunophenotypic algorithms predicts

prognosis in patients with diffuse large B-cell lymphoma treated with immunochemotherapy. *Blood* **117**, 4836–4843 (2011).

83. Ofori, K., Bhagat, G. & Rai, A. Exosomes as Liquid Biopsy Biomarkers in Diffuse Large B-cell Lymphoma (DLBCL)- Current State-of-the-Art and Unmet Clinical Needs. *Authorea Prepr.* (2020). doi:10.22541/AU.159050477.74611924

84. Couto, N., Caja, S., Maia, J., Strano Moraes, M. C. & Costa-Silva, B. Exosomes as emerging players in cancer biology. *Biochimie* **155**, 2–10 (2018).

85. Carvalho, A. S. *et al.* Bronchoalveolar Lavage Proteomics in Patients with Suspected Lung Cancer. *Sci. Rep.* **7**, 1–13 (2017).

86. Scherer, F. *et al.* Distinct biological subtypes and patterns of genome evolution in lymphoma revealed by circulating tumor DNA. *Sci. Transl. Med.* **8**, (2016).

87. Aung, T. *et al.* Exosomal evasion of humoral immunotherapy in aggressive B-cell lymphoma modulated by ATP-binding cassette transporter A3. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 15336–41 (2011).

88. Carvalho, A. S. *et al.* Is the Proteome of Bronchoalveolar Lavage Extracellular Vesicles a Marker of Advanced Lung Cancer? *Cancers (Basel).* **12**, 3450 (2020).

89. Pan, B. T. & Johnstone, R. M. Fate of the transferrin receptor during maturation of sheep reticulocytes in vitro: Selective externalization of the receptor. *Cell* **33**, 967–978 (1983).

90. Pan, B. T., Teng, K., Wu, C., Adam, M. & Johnstone, R. M. Electron microscopic evidence for externalization of the transferrin receptor in vesicular form in sheep reticulocytes. *J. Cell Biol.* **101**, 942–948 (1985).

91. Conde-Vancells, J. *et al.* Characterization and comprehensive proteome profiling of exosomes secreted by hepatocytes. *J. Proteome Res.* **7**, 5157–66 (2008).

92. Zhang, H. *et al.* Identification of distinct nanoparticles and subsets of extracellular vesicles by asymmetric flow field-flow fractionation. *Nat. Cell Biol.* **20**, 332–343 (2018).

93. Minciacchi, V. R. *et al.* MYC mediates large oncosome-induced fibroblast reprogramming in prostate cancer. *Cancer Res.* **77**, 2306–2317 (2017).

94. Van Niel, G., D'Angelo, G. & Raposo, G. Shedding light on the cell biology of extracellular vesicles. *Nature Reviews Molecular Cell Biology* **19**, 213–228 (2018).

95. Escola, J. M. *et al.* Selective enrichment of tetraspan proteins on the internal vesicles of multivesicular endosomes and on exosomes secreted by human B-lymphocytes. *J. Biol. Chem.* **273**, 20121–20127 (1998).

96. Hessvik, N. P. & Llorente, A. Current knowledge on exosome biogenesis and release. *Cell. Mol. Life Sci.* **75**, 193–208 (2018).

97. Anand, S., Samuel, M., Kumar, S. & Mathivanan, S. Ticket to a bubble ride: Cargo sorting into exosomes and extracellular vesicles. *Biochim. Biophys. Acta - Proteins Proteomics* (2019).

98. Jeppesen, D. K. *et al.* Reassessment of Exosome Composition. *Cell* **177**, 428-445.e18 (2019).

99. Ferreira, J. V. *et al.* Exosomes and STUB1/CHIP cooperate to maintain intracellular proteostasis. *PLoS One* **14**, e0223790 (2019).

100. Moreno-Gonzalo, O., Fernandez-Delgado, I. & Sanchez-Madrid, F. Post-translational add-ons mark the path in exosomal protein sorting. *Cellular and Molecular Life Sciences* **75**, (2018).

101. Mathieu, M., Martin-Jaular, L., Lavieu, G. & Théry, C. Specificities of secretion and uptake of exosomes and other extracellular vesicles for cell-to-cell communication. *Nature Cell Biology* **21**, 9–17 (2019).

102. Couto, N., Costa-Silva, B., Caja, S., Strano Moraes, M. C. & Maia, J. Exosome-Based Cell-Cell Communication in the Tumor Microenvironment. *Front. Cell Dev. Biol.* **6**, 1–19 (2018).

103. Peinado, H. *et al.* Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET. *Nat. Med.* **18**, 883–891 (2012).

104. Costa-Silva, B. *et al.* Pancreatic cancer exosomes initiate pre-metastatic niche formation in the liver. *Nat. Cell Biol.* **17**, 816–826 (2015).

105. Hoshino, A. *et al.* Tumour exosome integrins determine organotropic metastasis. *Nature* **527**, 329–335 (2015).

106. Fernandes, M., Teixeira, A. L. & Medeiros, R. The opportunistic effect of exosomes on Non-Hodgkin Lymphoma microenvironment modulation. *Critical Reviews in Oncology/Hematology* **144**, 102825 (2019).

107. Koch, R. *et al.* Populational equilibrium through exosome-mediated Wnt signaling in tumor progression of diffuse large B-cell lymphoma. *Blood* **123**, 2189–2198 (2014).

108. Feng, Y. *et al.* The Role and Underlying Mechanism of Exosomal CA1 in Chemotherapy Resistance in Diffuse Large B Cell Lymphoma. *Mol. Ther. - Nucleic Acids* **21**, 452–463 (2020).

109. Oksvold, M. P. *et al.* Expression of B-Cell surface antigens in subpopulations of exosomes released from B-cell lymphoma cells. *Clin. Ther.* **36**, (2014).

110. Koch, R. *et al.* Cancer Therapy: Preclinical Nuclear Trapping through Inhibition of Exosomal Export by Indomethacin Increases Cytostatic Efficacy of Doxorubicin and Pixantrone. (2015). doi:10.1158/1078-0432.CCR-15-0577

111. Ferguson Bennit, H. R. *et al.* Uptake of lymphoma-derived exosomes by peripheral blood leukocytes. *Blood Lymphat. Cancer Targets Ther.* **Volume 7**, 9–23 (2017).

112. Chen, Z. *et al.* Dual effect of DLBCL-derived EXOs in lymphoma to improve DC vaccine efficacy in vitro while favor tumorgenesis in vivo. *J. Exp. Clin. Cancer Res.* **37**, 190 (2018).

113. Zhu, M. Y. *et al.* NSE from diffuse large B-cell lymphoma cells regulates macrophage polarization. *Cancer Manag. Res.* **11**, 4577–4595 (2019).

114. Liu, W., Zhu, M., Wang, H., Wang, W. & Lu, Y. Diffuse large B cell lymphoma-derived extracellular vesicles educate macrophages to promote tumours progression by increasing PGC-1β. *Scand. J. Immunol.* **91**, (2020).

115. Carvalho, A. S. *et al.* Extra-cellular vesicles carry proteome of cancer hallmarks. *Front. Biosci. - Landmark* **25**, 398–436 (2020).

116. Hurwitz, S. N. *et al.* Proteomic profiling of NCI-60 extracellular vesicles uncovers common protein cargo and cancer type-specific biomarkers. *Oncotarget* **7**, 86999–87015 (2016).

117. Yáñez-Mó, M. *et al.* Biological properties of extracellular vesicles and their physiological functions. *Journal of Extracellular Vesicles* **4**, 1–60 (2015).

118. Hoshino, A. *et al.* Extracellular Vesicle and Particle Biomarkers Define Multiple Human Cancers. *Cell* **182**, 1044-1061.e18 (2020).

119. Rontogianni, S. *et al.* Proteomic profiling of extracellular vesicles allows for human breast cancer subtyping. *Commun. Biol.* **2**, 325 (2019).

120. Théry, C. *et al.* Minimal information for studies of extracellular vesicles 2018 (MISEV2018): a position statement of the International Society for Extracellular Vesicles and update of the MISEV2014 guidelines. *J. Extracell. Vesicles* **7**, 1535750 (2018).

121. Deeb, S. J., D'Souza, R. C. J., Cox, J., Schmidt-Supprian, M. & Mann, M. Super-SILAC Allows Classification of Diffuse Large B-cell Lymphoma Subtypes by Their Protein Expression Profiles. *Mol. Cell. Proteomics* **11**, 77–89 (2012).

122. Deeb, S. J., Cox, J., Schmidt-Supprian, M. & Mann, M. N-linked glycosylation enrichment for in-depth cell surface proteomics of diffuse large B-cell lymphoma subtypes. *Mol. Cell. Proteomics* **13**, 240–251 (2014).

123. Deeb, S. J. *et al.* Machine Learning-based Classification of Diffuse Large B-cell Lymphoma Patients by Their Protein Expression Profiles. *Mol. Cell. Proteomics* **14**, 2947–60 (2015).

124. Kwiecińska, A. *et al.* Proteomic Profiling of Diffuse Large B-Cell Lymphomas. *Pathobiology* **85**, 211–219 (2018).

125. Van Der Meeren, L. E. *et al.* A super-SILAC based proteomics analysis of diffuse large B-cell lymphoma-NOS patient samples to identify new proteins that discriminate GCB and non-GCB lymphomas. **14**, e0223260 (2019).

126. Théry, C., Amigorena, S., Raposo, G. & Clayton, A. Isolation and Characterization of Exosomes from Cell

Culture Supernatants and Biological Fluids. *Curr. Protoc. Cell Biol.* **30**, 3.22.1-3.22.29 (2006).

127. Matthiesen, R. & Carvalho, A. S. Methods and algorithms for quantitative proteomics by mass spectrometry. in *Methods in Molecular Biology* **2051**, 161–197 (Humana Press Inc., 2020).

128. Wis, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. (2009). doi:10.1038/NMETH.1322

129. Voabil, P. *et al.* Global Mass Spectrometry and Transcriptomics Array Based Drug Profiling Provides Novel Insight into Glucosamine Induced Endoplasmic Reticulum Stress. *Mol. Cell. Proteomics* **13**, 3294–3307 (2014).

130. Matthiesen, R. *et al.* SIR: Deterministic protein inference from peptides assigned to MS data. *J. Proteomics* **75**, 4176–4183 (2012).

131. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).

132. Ritchie, M. E. *et al.* Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

133. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).

134. Pathan, M. *et al.* FunRich: An open access standalone functional enrichment and interaction network analysis tool. *Proteomics* **15**, 2597–2601 (2015).

135. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).

136. Galili, T. dendextend: An R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics* **31**, 3718–3720 (2015).

137. Lê, S., Josse, J., Rennes, A. & Husson, F. *FactoMineR: An R Package for Multivariate Analysis*. *JSS Journal of Statistical Software* **25**, (2008).

138. Blighe, K, S Rana, and M. L. EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. (2018). Available at: https://github.com/kevinblighe/EnhancedVolcano.

139. Carvalho, A. S., Molina, H. & Matthiesen, R. New insights into functional regulation in MS-based drug profiling. *Sci. Rep.* **6**, 1–11 (2016).

140. Filipe, V., Hawe, A. & Jiskoot, W. Critical evaluation of nanoparticle tracking analysis (NTA) by NanoSight for the measurement of nanoparticles and protein aggregates. *Pharm. Res.* **27**, 796–810 (2010).

141. Webber, J. & Clayton, A. How pure are your vesicles? *J. Extracell. Vesicles* **2**, 19861 (2013).

142. Frades, I. & Matthiesen, R. Overview on techniques in cluster analysis. *Methods in molecular biology (Clifton, N.J.)* **593**, 81–107 (2010).

143. Springer, I. T. J. *Principal Component Analysis, Second Edition*.

144. Kalra, H. *et al.* Vesiclepedia: A Compendium for Extracellular Vesicles with Continuous Community Annotation. *PLoS Biol.* **10**, e1001450 (2012).

145. Pathan, M. *et al.* Vesiclepedia 2019: A compendium of RNA, proteins, lipids and metabolites in extracellular vesicles. *Nucleic Acids Res.* **47**, D516–D519 (2019).

146. Mathivanan, S., Fahner, C. J., Reid, G. E. & Simpson, R. J. ExoCarta 2012: Database of exosomal proteins, RNA and lipids. *Nucleic Acids Res.* **40**, D1241–D1244 (2012).

147. Aebersold, R. & Mann, M. Mass-spectrometric exploration of proteome structure and function. *Nature* **537**, 347–355 (2016).

148. Tian, Y. *et al.* Quality and efficiency assessment of six extracellular vesicle isolation methods by nano-flow cytometry. *J. Extracell. Vesicles* **9**, (2020).

149. Kowal, J. *et al.* Proteomic comparison defines novel markers to characterize heterogeneous populations of extracellular vesicle subtypes. *Proc. Natl. Acad. Sci.* **113**, E968–E977 (2016).

150. Abramowicz, A. *et al.* Harmonization of exosome isolation from culture supernatants for optimized proteomics analysis. *PLoS One* **13**, e0205496 (2018).

151. Witwer, K. W. & Théry, C. Extracellular vesicles or exosomes? On primacy, precision, and popularity influencing a choice of nomenclature. *J. Extracell. Vesicles* **8**, 1648167 (2019).

152. Deng, W. *et al.* Variable Responses of MYC Translocation Positive Lymphoma Cell Lines To Different Combinations of Novel Agents: Impact of BCL2 Family Protein Expression. *Transl. Oncol.* **11**, 1147–1154 (2018).

153. Van Deun, J. *et al.* EV-TRACK: Transparent reporting and centralizing knowledge in extracellular vesicle research. *Nature Methods* **14**, 228–232 (2017).

154. Martens, L. *et al.* PRIDE: The proteomics identifications database. *Proteomics* **5**, 3537–3545 (2005).

155. Cvjetkovic, A. *et al.* Detailed Analysis of Protein Topology of Extracellular Vesicles–Evidence of Unconventional Membrane Protein Orientation. *Sci. Rep.* **6**, 36338 (2016).

156. Manfroi, B. *et al.* Tumor-associated neutrophils correlate with poor prognosis in diffuse large B-cell lymphoma patients. *Blood Cancer Journal* **8**, 66 (2018).

157. Nolte, M. A., Nolte-'t Hoen, E. N. M. & Margadant, C. Integrins Control Vesicular Trafficking; New Tricks for Old Dogs. *Trends in Biochemical Sciences* **46**, 124–137 (2021).