Diffuse large B-cell lymphoma – proteomic and metabolomic studies on prognosis and treatment failure

Martin Stenson

Department of Internal Medicine and Clinical Nutrition
Institute of Medicine
Sahlgrenska Academy, University of Gothenburg

UNIVERSITY OF GOTHENBURG

Gothenburg 2018

Cover illustration: Bohusgranit. Utterholmarna 18/7-2018.

”Det var ett kullarnas land och höjdernas, följaktligen dalgångarnas och bäckarnas. Kullarna bestodo av granit, men liksom sedd genom förstoringsglas och polariserat ljus, så att de eljes små kristallerna visade sig i större skala och i sönderdelade färger. Glimmern tedde sig som stora fjäll av guld och silver, kvartsen som ljusblå eller rosenröda sexhörningar och den köttfärgade fältspaten i streckade gyttringar som gesimser. Och de simplaste stenarter erbjödo ett evigt växlande skådespel för ögat; inblandade korn av det värdelösa hornbländet tedde sig som ädla stenar i vassgrönt och djup purpur, den banala olivinen i svagt gult som gurkmeja eller safflor.”

- August Strindberg, ur Armageddon (Början till en roman)

Diffuse large B-cell lymphoma – proteomic and metabolomic studies on prognosis and treatment failure

martin.stenson@vgregion.se

ISBN 978-91-7833-153-6 (PRINT)
ISBN 978-91-7833-154-3 (PDF)
http://hdl.handle.net/2077/56918

Printed in Gothenburg, Sweden 2018
Printed by BrandFactory

Without deviation from the norm, progress is not possible

-Frank Zappa

Diffuse large B-cell lymphoma – proteomic and metabolomic studies on prognosis and treatment failure

Martin Stenson

Department of Internal Medicine and Clinical Nutrition, Institute of Medicine
Sahlgrenska Academy, University of Gothenburg, Sweden

ABSTRACT

Background and aim: Every year almost 600 patients in Sweden are diagnosed with diffuse large B-cell lymphoma (DLBCL), the most common lymphoma, and with immunochemotherapy, approximately 60 % are cured. Yet, for patients with primary refractory disease or early relapse, the prognosis is very poor. Despite advances in molecular subclassification of DLBCL, the major tool used to risk stratify patients is the clinically based International Prognostic Index (IPI). However, there is still no available system that with precision can identify the individual patients at highest risk of treatment failure. The aim of this thesis was to search for novel prognostic and predictive biomarkers, and also investigate the mechanisms behind chemoresistance in DLBCL.

Patients and methods: In paper I and III, tumor tissue from two groups of DLBCL patients; (i) patients with primary refractory disease or early relapse (REF/REL; paper I: n=5, paper III: n=44); and (ii) long-term progression-free patients, clinically considered cured (CURED; paper I: n=5, paper III: n=53), was examined with mass spectrometry proteomic approaches to explore possible differences in global protein expression, but also with the aim to reveal new mechanisms involved in immunochemotherapy resistance. In paper II, metabolomic examination with nuclear magnetic resonance spectroscopy was performed on serum from REF/REL (n=27) and CURED (n=60) DLBCL patients, to determine if differences in clinical outcome could be correlated to diverse metabolomic profiles.

Results and Conclusions: In paper I, a large number of proteins could be identified and quantified. Overexpression of actin-related proteins was found among the CURED patients, a finding that appeared to be confirmed in paper III, where in addition a novel discovery regarding overexpression of multiple ribosomal proteins in the REF/REL group was made. The findings suggest previously undescribed mechanisms for immunochemotherapy resistance in DLBCL patients. In paper II, differences in the serum metabolome was found between the two groups, that could be separated with multivariate statistical analyses. Even though the results are encouraging they need to be confirmed in larger unselected studies, with aims of further exploring actin-related and ribosomal proteins, not only as possible prognostic/predictive biomarkers, but also regarding their functional role in treatment resistance in DLBCL.

Keywords: DLBCL, prognostics, proteomics, metabolomics

ISBN 978-91-7833-153-6 (PRINT), ISBN 978-91-7833-154-3 (PDF), http://hdl.handle.net/2077/56918

SAMMANFATTNING PÅ SVENSKA

Bakgrund: Diffust storcelligt B-cellslymfom (DLBCL), den vanligaste lymfomtypen, drabbar 500-600 patienter i Sverige varje år. DLBCL är en aggressiv sjukdom där överlevnaden är kort utan behandling. Med hjälp av immunokemoterapi, dvs dosintensiv cytostatikabehandling (CHOP) med tillägg av rituximab (en monoklonal anti-CD20 antikropp), botar man idag drygt 60% av patienterna, men i de fall där behandling har dålig effekt (behandlingsrefraktär sjukdom), eller där sjukdomen snabbt återkommer efter avslutad behandling, är prognosen mycket dålig. Trots att det på senare år nåtts framgångar inom molekylär subklassifiering av DLBCL är fortfarande det kliniskt baserade IPI (International Prognostic Index), det prognostiska instrument som används i klinisk vardag. Men IPI har stora svårigheter att förutsäga vilka patienter som har störst risk för dåligt behandlingssvar, dvs behandlingsresistens och det finns ett behov av att hitta biologiska riskmarkörer som bättre kan identifiera individuella patienter med högst risk.

Patienter och metoder: I delarbete I och III gjordes proteomikanalyser på sparad tumörvävnad från DLBCL-patienter från två distinkt olika grupper; i) patienter med primärt behandlingsrefraktär sjukdom eller som fått återfall inom 1 år efter diagnos (REF/REL), och ii) botade patienter, dvs patienter som ej fått återfall under 5 års uppföljning efter behandling (CURED). Med proteomik undersöks det globala proteinmönstret i en vävnad, och målet med studierna var att se om man kunde hitta skillnader i proteinuttryck mellan grupperna, dels för att kunna använda dessa skillnader som prognostiska verktyg, men också för att leta efter okända biologiska mekanismer bakom behandlingsresistens.

I delarbete II gjordes i stället en metabolomikanalys på sparat patientserum från samma patientgrupper. Med metabolomik undersöks hur mönstret av olika metaboliter och andra mindre molekyler ser ut i ett prov, och syftet med delarbete III var att se om skillnaderna i behandlingsresultat mellan de två patientgrupperna kunde kopplas till skillnader i detta metabolitmönster.

Resultat och konklusion: I delarbete I kunde en stor mängd proteiner identifieras och kvantifieras i de sammanlagt 10 proverna, som utgjordes av färskfrusna lymfkörtelpreparat från DLBCL-patienter. Uttrycket av sammanlagt 87 proteiner skilde sig mellan de två patientgrupperna, som med hjälp av multivariata statistikmetoder kunde separeras utifrån sina proteinuttryck. I gruppen med långtidsöverlevande sågs ett överuttryck av proteiner kopplade till cellens actincytoskelett. I delarbete III, det andra proteomikarbetet, kunde återigen en stor mängd proteiner identifieras och kvantifieras i sammanlagt 97 prover, som denna gång utgjordes av

formalinfixerade, paraffininbäddade tumörpreparat. Fyndet av överuttryck av actinrelaterade protein hos CURED-patienterna verkade stå sig från delarbete I, och dessutom hittades ett överuttryck av ribosomala proteiner hos REF/REL-patienterna. Ribosomala proteiner är kopplade till ribosomerna, som är de små cellulära maskiner som tillverkar proteiner, dvs översätter genetisk information och sätter ihop aminosyror till proteiner utifrån DNA-mallen i arvsmassan. Fynden gällande actinrelaterade och ribosomala proteiner kan, förutom möjligheterna till prognostisk information, ge en helt ny inblick i möjliga förklaringsmodeller bakom behandlingsresistens vid DLBCL, och kan också öka möjligheterna att finna nya mål för riktade behandlingar. I delarbete II sågs skillnader i serummetabolomet vid jämförelse mellan de två patientgrupperna, skillnader med potential att kunna ge prognostisk information men som behöver konfirmeras i större studier.

LIST OF PAPERS

This thesis is based on the following studies, referred to in the text by their Roman numerals.

I. Rüetschi U*, Stenson M*, Hasselblom S, Nilsson-Ehle H, Hansson U, Fagman H, Andersson P-O. SILAC-based quantitative proteomic analysis of diffuse large B-cell lymphoma Patients. Int J Proteomics vol. 2015: Article ID 841769, 12 pages. **These authors contributed equally.*

II. Stenson M, Pedersen A, Hasselblom S, Nilsson-Ehle H, Karlsson G, Pinto R, Andersson P-O. Serum nuclear magnetic resonance-based metabolomics and outcome in diffuse large B-cell lymphoma patients - a pilot study. Leukemia and Lymphoma 2016; 57:8, 1814-1822.

III. Bram Ednersson S*, Stenson M*, Stern M, Enblad G, Fagman H, Nilsson-Ehle H, Hasselblom S, Andersson, P-O. Expression of ribosomal and actin network proteins and immunochemotherapy resistance in diffuse large B cell lymphoma patients. Br J Haematol 2018; 181: 770-781. **These authors contributed equally.*

CONTENT

ABBREVIATIONS

^{1}H NMR	Proton Nuclear Magnetic Resonance
aaIPI	Age-adjusted International Prognostic Index
ABC	Activated B-cell like
ADCC	Antibody-dependent cellular cytotoxicity
AID	Activation-induced cytidine deaminase
ALL	Acute lymphoblastic leukemia
ASCT	Autologous stem cell transplantation
BCR	B-cell receptor
BH	Benjamini-Hochberg (statistical method)
BM	Bone marrow
CAR	Chimeric antigen receptor
COO	Cell-of-origin
CR	Complete remission/response
CT	Computed tomography
DEL	Double expressor lymphomas
DHL	Double hit lymphomas
DLBCL	Diffuse large B-cell lymphoma
FFPE	Formalin-fixed paraffin-embedded
GC	Germinal center
GCB	Germinal center B-like
GEP	Gene expression profiling
H chains	Heavy chains
HDAC	Histone deacetylase

Ig	Immunoglobulin
IHC	Immunohistochemistry
IMiD	Immunomodulating drug
IPI	International Prognostic Index
L chains	Light chains
LC-MS/MS	Liquid chromatography – tandem mass spectrometry
LDH	Lactate dehydrogenase (also LD)
MALT	Mucosa-associated lymphoid tissue
MS	Mass spectrometry
MZ	Marginal Zone
OPLS-DA	Orthogonal projection to latent structures discriminant analysis (statistical method)
OS	Overall survival
PCA	Principal components analysis (statistical method)
PET	Positron emission tomography
PFS	Progression free survival
PLS-DA	Partial least-squares discriminant analysis (statistical method)
RP	Ribosomal protein
SHM	Somatic hypermutation
THL	Triple hit lymphomas
TMA	Tissue micro array
TMT	Tandem mass tag

DEFINITIONS IN SHORT

CURED	Patients that were long-term progression-free with a follow-up, from diagnosis, of at least 5 years, clinically considered cured from disease.
ECOG Performance Status	Scale of performance status developed by the Eastern Cooperative Oncology Group, that describes the patient's level of functioning from 0 (fully active) to 4 (completely disabled).
Metabolomics	Large scale study or investigation of the metabolome, i.e. the complete set of small-molecule metabolites that is found within a biological sample.
Proteomics	Large scale study or investigation of the proteome, i.e. the entire set of proteins that are produced or modified by an organism or system.
R-CHOP	Immunochemotherapy regimen, containing the chemotherapy agents cyclophosphamide (C), doxorubicin (H) and oncovin (O), high dose corticosteroids, i.e. prednisone (P), and the monoclonal anti-CD20 antibody rituximab (R).
REF/REL	Patients with primary refractory disease or relapse within 1 year after diagnosis.
SILAC	*Stable isotope labeling of amino acids in cell culture*, a proteomic method in which an isotope-labeled mix of proteins is analyzed alongside the protein sample, to create an internal reference for quantification.

1 INTRODUCTION

Lymphomas make up approximately 4% of all cancer in the western world. More than 90 different variations of mature lymphoid neoplasms, i.e. lymphomas, are listed in the recently updated WHO classification of hematological malignances (1). Lymphomas are very heterogeneous diseases. They differ greatly in their clinical manifestations, their histological and molecular appearances and not least their prognosis. The great majority of lymphomas are derived from B-cells (rather than T- or NK-cells) (2), and have clinical courses spanning from very indolent, chronic and even asymptomatic (3) to very aggressive with a multitude of symptoms and very bad prognosis with short-time mortality of 100 % if left untreated.

The most common lymphoma, diffuse large B-cell lymphoma (DLBCL), make up 20-25% of all lymphomas (2), and is in itself a heterogeneous disease, that comprises around 15 subgroups among the mature B-cell neoplasms in the WHO-classification. In Sweden the annual incidence of DLBCL is 5-6/100.000, which means that 500 to 600 people are diagnosed every year (4). DLBCL is an aggressive lymphoma that mostly presents with enlarged lymph nodes and sometimes associated systemic symptoms (fever, weight loss, night sweats). Treatment with standard immunochemotherapy can cure well over 50% of patients, but those with tumors that are refractory to initial treatment or have a quick relapse, have a very poor prognosis (5, 6).

Despite great progress in molecular subclassification of DLBCL, much is still uncertain regarding the prognostication of the individual patient, and there are still mainly clinical variables, i.e. the IPI-score (7) (age, performance status, disease stage, extranodal disease, and lactate dehydrogenase in serum) that are used to risk-stratify patients newly diagnosed with DLBCL.

In this thesis tumor tissue or serum from DLBCL-patients with different clinical outcome; either cured patients or patients with refractory or relapsed disease, have been retrospectively analyzed in search for differences in protein expression or metabolite concentration, with the aim of finding new insights in the molecular background of the disease, explanations for resistance to immunochemotherapy and also new prognostic or predictive biomarkers.

1.1 THE B-CELL

B-cells, lymphocytes that are responsible for the production of antibodies, are key players in the adaptive immune system including the humoral immune response and the immunologic memory. Different subsets of B-cells manages different phases of the immune response. For example, in the marginal zone of the spleen, non T-cell dependent B-cells are responsible for the first line defense against blood-borne pathogens with low affinity IgM-antibodies, whereas the later immune response with high affinity IgG- or IgA-antibodies and development of B-memory cells comes from B-cells having matured, with the aid of T-cells, in the germinal centers of lymph nodes. The ontogenesis of B-cells is a complex course of events where they go through programmed mutational processes in which their DNA step-wise is altered to acquire more specific recognition of foreign antigens, a process with powerful elements of selection to which most B-cells succumb due to non-functioning antibodies or too strong recognition of self-antigens. The inborn DNA-modification capacity of B-cells also makes them prone to undergo malignant transformation, resulting in different types of B-cell lymphomas depending on in which stage of B-cell development the oncogenic DNA-changes occur (8-10).

1.1.1 Early B-cell receptor development in the bone marrow

Central in B-cell development, and necessary for B-cell survival, is the B-cell receptor (BCR), which is made up by an antibody, also called immunoglobulin (Ig), coupled to transmembrane proteins (CD79A and CD79B) with intracellular signaling capacity. Antibodies are Y-shaped proteins formed from two heavy chains (H chains) and two light chains (L chains), both with variable domains in the ends, domains that together make up the antigen-recognizing segments of the Ig. The non-variable ends of the H chains define the class or isotype of the antibody, that in the early stages always is of IgM-type. The initial DNA-changing events in the Ig-genes of the early B-cell occur in the bone marrow (BM). Three different segments of the variable portion of the H chain, the V (variable), D (diverse) and J (joining) segments, and two variable segments of the L chain, the V and J segments, are rearranged by the RAG1/2 recombinase. The human IgH (H chain gene), in its variable region contains a palette of 27 D_H gene segments, 6 J_H segments and 120 V_H segments, from which the actions of the RAG1/2 recombinase randomly select one of each to make the $V_H(D_H)J_H$ recombination. The recombined H chain is transiently coupled to surrogate L chains to form a pre-BCR that is tested for functionality, i.e. some kind of antigen recognizing capacity, before being allowed to progress into further development. B-cells with non-functioning pre-BCR:s undergo apoptosis, if they are not first saved

by available escape-mechanisms to rearrange the properties of the BCR (including switching to the other IgH-allele, or additional recombination including unused V_H-segments of the original IgH-allele). In the positively selected B-cells the surrogate L chains are replaced by the definitive light chains, in which the variable regions are similarly rearranged to form a V_LJ_L recombination, either from the κ-loci or λ-loci of the IgL (L chain gene), thus forming either a κ light chain or a λ light chain. The Ig of the BCR is now composed of functional heavy and light chains, but before leaving the BM as naïve B-cells the BCR must be tested for reactivity to autoantigens. B-cells with strong autoreactivity are either saved by further L-chain rearrangements, or forced to the path of apoptosis, the latter fate actually affecting a majority of B-cells (11-14).

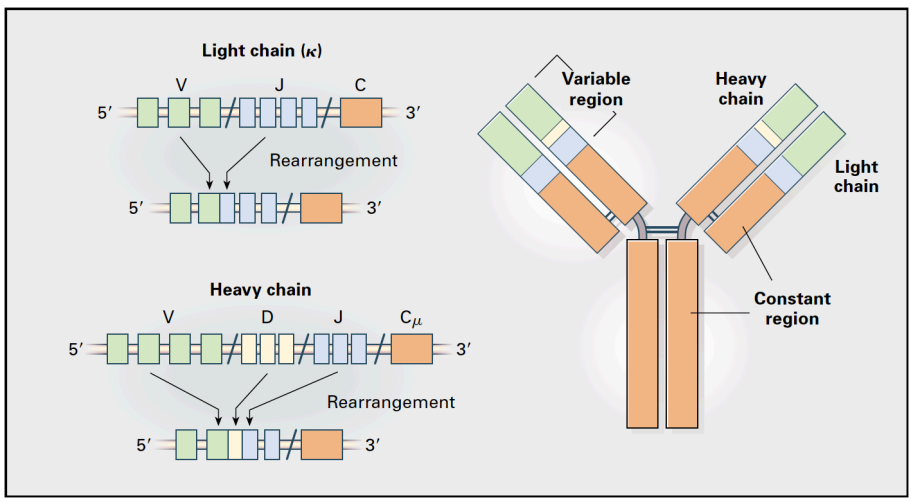

Figure 1. V(D)J recombination of the immunoglobulin during B-cell development. *Schematic presentation of how the variable regions of the heavy chain, consisting of variable (V), diversity (D) and joining (J) segments, and of the light chain (of either κ or λ type) with V and J segments, are assembled by recombination. In this process, the unused gene segments are deleted or inverted. The number of possible combinations of the different variable regions of the heavy and light chains is an astonishing 5 x 10^{13}(13). From (11). Reproduced with permission from Massachusetts Medical Society: The New England Journal of Medicine © 1999.*

1.1.2 Migration to the spleen

The B-cells that escape the negative selection leave the BM and migrate to the spleen, where they either become marginal zone (MZ) B-cells that finalize their maturation in the spleen, or are referred to secondary lymphoid organs, mainly lymph nodes, for further development. In the spleen, MZ B-cells form the first line of defense against blood borne pathogens by T-cell independent transformation into short lived plasma cells secreting low affinity antibodies of the IgM isotype (10).

1.1.3 Further development in the lymph node

In lymph nodes resting B-cells can recognize and thereby be activated by antigens, thus starting the second line of defense, a more specific B-cell response. The activated B-cells internalize the recognized antigens and present them to T-helper cells, which in turn starts a loop of reciprocal stimulation promoting further B-cell proliferation and induction of the somatic hypermutation (SHM) process, which strengthen the BCR affinity to the antigens (15).

1.1.4 The germinal center

The quickly dividing B-cells, now termed centrocytes and centroblasts, do together with T-helper cells and follicular dendritic cells (stromal cells) make up the germinal centers of the lymph node follicles, areas that can be likened to busy immunological factories that fine tune and mass-produce highly specific immunologic responses to invading pathogens. The centroblasts constitute the dark zone of the GC, in which rapid proliferation and SHM take place. In the light zone, the centrocytes undergo positive or negative selection depending on the affinity of their BCR, and here also the Ig class switch and differentiation to either plasma cells or memory B-cells happens (16).

1.1.5 Somatic hypermutation

Somatic hypermutation is induced by an enzyme, activation-induced cytidine deaminase (AID), which causes small DNA-modifications such as point mutations, deletions and insertions into the rearranged IgV, i.e. the variable genes of the immunoglobulin. This changes the geometry of the antigen-recognizing segment of the antibody. If this results in a BCR with decreased affinity to the antigen or increased affinity to self-antigens the B-cell is counterselected and undergo apoptosis. However, if the BCR after SHM instead increases its affinity to the antigen the B-cell is positively selected to further proliferation including more cycles of SHM (13).

1.1.6 Ig class switch and final differentiation

Towards the end of the B-cell development in the germinal center, the enzyme AID is again involved in DNA modifying events, this time through the class switch recombination, in which the constant region of the antibody heavy chain in some (but not all) of the B-cells are changed to form new isotypes of antibodies, primarily from IgM to IgG or IgA. This changes the effector functions of the secreted antibody and also affects the intracellular signaling capacity of the BCR. Finally, the B-cell can leave the germinal center after having differentiated into either an antibody secreting plasma cell or a memory B-cell (16).

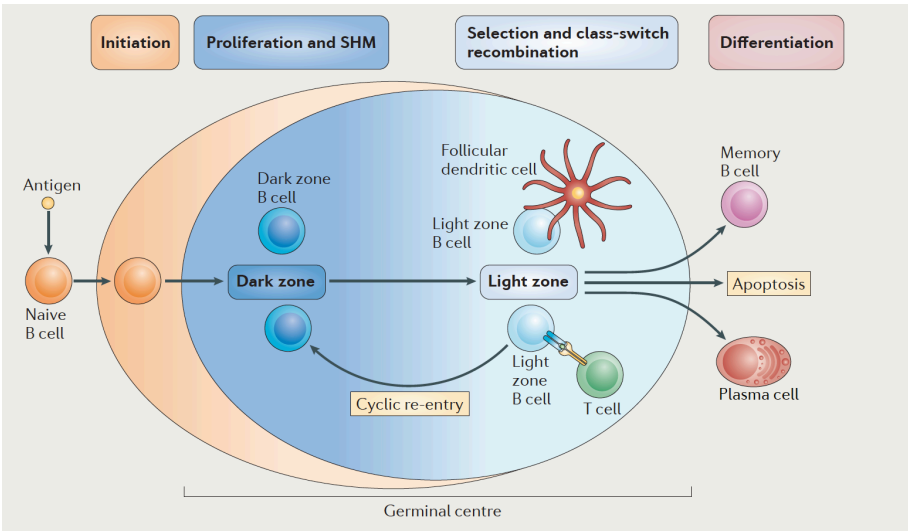

Figure 2. The germinal center (GC). *Naïve B-cells that are activated by antigens form the germinal center, in which proliferation and somatic hypermutation (SHM) take place in the dark zone, and selection and Ig class-switch recombination in the light zone. The B-cells cycle between the different compartments of the GC several times, to undergo repeated rounds of proliferation, SHM and selection, before finally exiting the GC as either memory B-cells or antibody secreting plasma cells. B-cells that are counterselected due to weak affinity of the BCR or too strong recognition of self-antigens undergo apoptosis. From (16). Reproduced with permission from Macmillan Publishers Ltd: Nat Rev Immunol © 2015.*

1.2 ORIGIN OF B-CELL LYMPHOMAS

As discussed earlier, the numerous DNA-changing events during normal B-cell development, together with the massive proliferation, make these cells prone to undergo malignant transformation. The histologic and molecular phenotype of lymphoma cells depend on the stage of normal B-cell development in which the malignant cells have their normal counterpart, i.e. at which stage the cells have made their final diversion from normal development and transformed into malignant cells. A vast majority of lymphomas have somatically mutated immunoglobulins in their genome, which means that they are stemming from germinal center (GC) B-cells (11). However, the first genetic changes probably happen earlier in B-cell development, but the last steps of the malignant transformation take place when the B-cell is exposed to antigens in the GC (16). Also, with few exceptions, most B-cell lymphomas seem to be dependent upon a functional BCR, most evident in lymphoma subtypes where proliferation is triggered by BCR-autoreactivity (as seen in cases of follicular lymphomas and MALT-lymphomas) and sometimes even where lymphomas are dependent on BCR-stimulation by a known pathogen (most studied in Hepatitis C-driven lymphomas) (14).

1.2.1 Examples of genetic alterations in B-cell lymphomas

Follicular lymphomas, that have an obvious GC origin, not least with regard to their histopathological growth pattern, have in 80 % of cases the typical t(14;18) translocation involving the anti-apoptotic protein BCL2. This translocation, which also is present in 35 % of germinal center B-like (GCB) DLBCL (a subtype introduced and explained in chapter 1.3.6), is caused by the (misdirected) actions of RAG1/2 recombinase during the V(D)J-translocation during early B-cell development in the BM, but contributes to lymphomagenesis later in the GC reaction (14, 17).

During the GC reaction, transient expression of MYC, an omnipotent transcription factor, is involved in the recycling of B-cells from the light zone into the dark zone for further proliferation and SHM. AID, the protein responsible for both SHM and Ig class switch, is also the culprit behind the t(8;14) translocation that involves MYC, that is present in about 10% of GCB-DLBCL and 100% of sporadic Burkitt lymphomas (8, 18).

BCL6, a protein with broad effects on transcription, mainly by repression, plays multiple roles in the GC process and is responsible for maintaining the B-cell in the GC process to promote further DNA modulation of the immunoglobulin genes, and repressing further differentiation into plasma cells or memory B-cells. BCL6 is present during the GC phase, and is

typically found in lymphomas with a GC phenotype, such as follicular lymphoma and GCB-DLBCL. BCL6 disappears during the final differentiation of the B-cell, and is typically absent in activated B-cell like (ABC) DLBCL (see chapter 1.3.6), a lymphoma subtype that has its probable natural counterpart in plasmablasts, i.e. B-cells having passed through the GC events and started to differentiate towards plasma cells (16, 19).

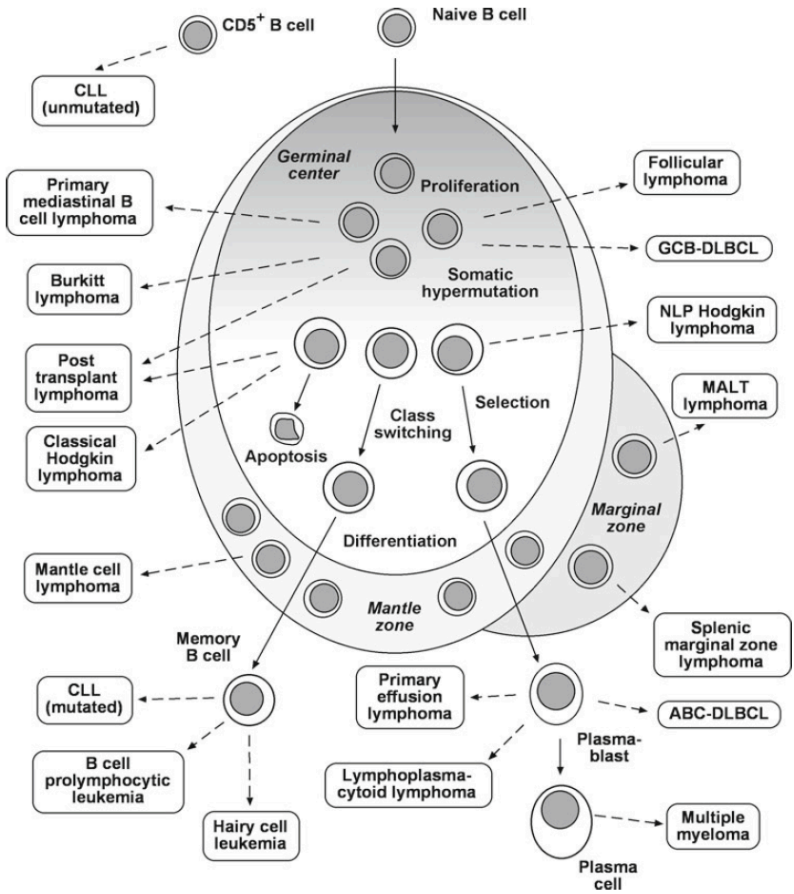

Figure 3. The cellular origin of B-cell lymphomas relative to the GC reaction. *Within the cell-of-origin concept of DLBCL, introduced in chapter 1.3.6, two major subgroups can be distinguished; germinal-center like (GCB) DLBCL, and activated B-cell like (ABC) DLBCL. From (19). Reproduced with permission from Springer Science+Business Media: Methods in Molecular Biology© 2013.*

1.3 DIFFUSE LARGE B-CELL LYMPHOMA

DLBCL is the most common lymphoma, making up about 20-25% of all lymphomas in the western world. The incidence of DLBCL has, for unknown reasons, steadily increased during the last 50 years (2, 20, 21). Five to six-hundred patients in Sweden are diagnosed with DLBCL every year (4), most of them with high age as the only known predisposing factor – the median age is 70 years at diagnosis (20, 21). Some cases of DLBCL arise after transformation from indolent lymphomas such as follicular lymphoma (1). There are also some other established factors that increase the risk of DLBCL, among them autoimmune and inflammatory diseases (22), HIV and other immune deficiency disorders, post-transplant situations with immunosuppression and previous radiation therapy. DLBCL is slightly more common in males, the male:female ratio being around 1.2:1 (21). Among first-degree relatives to DLBCL there is a 10-fold increase in relative risk for the same lymphoma type, implying an association of (unknown) specific germline genes. (23). There are no known life style factors that significantly influence the risk for DLBCL.

DLBCL is a heterogeneous disease, that comprises around 15 subgroups among the mature B-cell neoplasms in the updated WHO-classification (1). The term "DLBCL", has in this text been used as a collective term for the subgroups comprising the absolute majority of DLBCL cases, i.e. "Diffuse large B-cell lymphoma (DLBCL) Not Otherwise Specified (NOS)" with subgroups "Germinal center B-cell type (GCB)" and "Activated B-cell type (ABC)", and also "High-grade B-cell lymphoma, with MYC and BCL2 and/or BCL6 rearrangements" and "High-grade B-cell lymphoma, NOS".

1.3.1 Clinical presentation

Patients with DLBCL often present with enlarged lymph nodes or tumors in extranodal sites, and have frequently associated systemic symptoms (fever, weight loss, night sweats). The clinical course is aggressive, the symptoms often having evolved only during the last weeks or months. Without treatment the disease inevitably will have a fatal course, with fast growing tumor masses and quick deterioration of the patient´s general condition.

1.3.2 Diagnosis

The work-up of patients with suspected DLBCL includes a mandatory tumor biopsy that undergo histopathological examination and immunohistochemical staining for diagnosis and subclassification according to the WHO (1). For best results the pathological examination is made on a lymph node that is surgically excised in its entirety, but if that is not possible a needle biopsy can

often gather sufficient tumor material for correct diagnosis. DLBCL is, as the name implies, morphologically composed of large transformed lymphoid cells in a diffuse growth pattern, that disrupt or fully replaces the follicular appearance of the normal lymph node. Immunohistochemical staining usually detects pan B-cell markers such as CD19, CD20, CD22 and CD79a. Expression of other markers like CD30, CD5, CD10, BCL6, BCL2 and MUM1 can be seen in various proportions of cases. Proliferation measured as Ki-67 fraction is usually high, varying between 40% to over 90%.

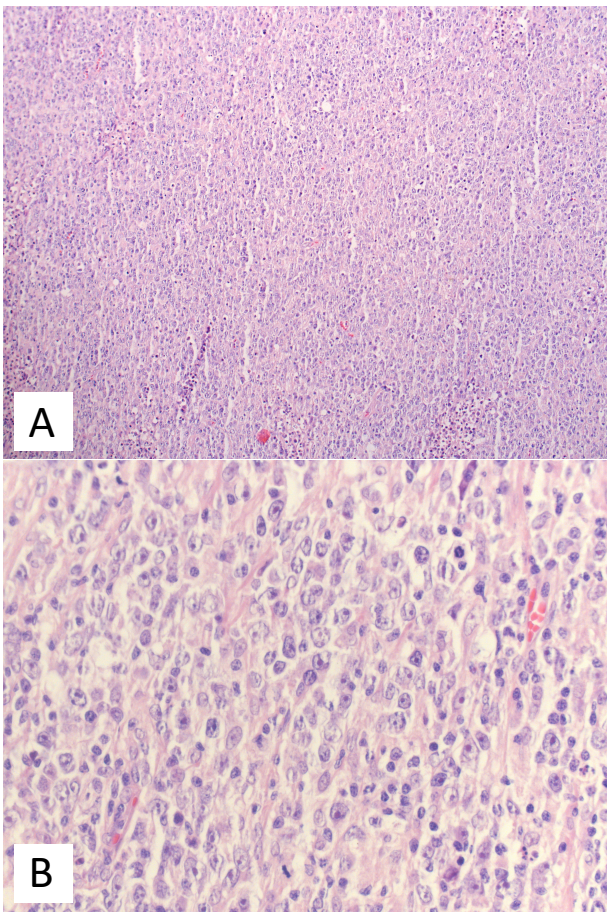

Figure 4. The centroblastic variant of DLBCL *is the predominant histologic type, accounting for approximately 80 % of cases. Here seen in two different magnifications; 10x (A) and 40x (B). The tumor is dominated by centroblasts, i.e. large cells with a moderate amount of cytoplasm, round to oval nuclei with 2-3 nucleoli, often peripherally located adjacent to the nuclear membrane. Reproduced with permission from dr. Bram Ednersson, Department of Pathology, Sahlgrenska University Hospital, Gothenburg.*

1.3.3 Staging

Staging of DLBCL is based on the Ann Arbor classification (24) and require, for full evaluation of the disease extension, investigation with CT-scan (or PET) and also a bone marrow biopsy. High-risk cases undergo examination of cerebrospinal fluid for assessment of possible CNS involvement. 45 % of patients present with stage I or II disease, the rest being in stage III or IV, i.e. having a more disseminated disease at diagnosis (21).

1.3.4 Treatment, and prognosis relative response

For many years the standard curative treatment for DLBCL patients was the CHOP-regimen, which combines the chemotherapy agents cyclophosphamide (C), doxorubicin (H) and oncovin (O) with high dose corticosteroids, i.e. prednisone (P). After the addition of the monoclonal anti-CD20 antibody rituximab (R) almost 20 years ago, the combined so-called *immunochemotherapy* regimen R-CHOP have been the cornerstone of DLBCL treatment, a regimen that cures approximately 60% of DLBCL patients (25). Several attempts to increase survival rates with more intensive frontline chemotherapy, mostly before rituximab was introduced, have not proved superior to R-CHOP (26, 27). The fast-growing nature and short doubling time of the tumor cells make DLBCL very sensitive to immunochemotherapy, and clinicians and patients can often see the tumors "melt away" quickly during the first courses of treatment. The goal is a total eradication of disease after the last course of the treatment, i.e. having a complete remission (CR), which is achieved in about 75% of patients (25).

Patients with primary refractory or relapsed DLBCL are in a difficult position. For younger patients the standard treatment, if the tumor is chemosensitive, is high-dose therapy with autologous stem cell transplantation (ASCT) (28). The most common treatments preceding the ASCT have been either ICE (ifosfamide, carboplatin and etoposide) or DHAP (dexamethasone, cytarabine and cisplatin) (29). Lately GDP (dexamethasone, gemcitabine and cisplatin) have proven to be as effective as DHAP regarding response to treatment and transplantation rates, however with less toxicity and superior quality of life, making it a more appealing treatment choice (30). Rituximab is added to the regimens if more than six months have passed since the last rituximab dose, but is omitted in cases with earlier relapses or primary refractory disease, since the tumor then is considered refractory to rituximab. Among the patients who relapse after ASCT the prognosis is dismal, but among patients responding to salvage therapy, there are cases that can obtain long remissions following allogenic stem cell transplantation (31). Also, with recent advances in chimeric antigen receptors (CARs), anti-CD19 CAR-T cells with very promising activity against chemotherapy-refractory DLBCL have been tested.

The risk of severe adverse effects with this treatment is well known, but chances of longer remissions are encouraging (32). Axicabtagene ciloleucel, an autologous anti-CD19 CAR-T cell therapy tested in a multicenter setting on DLBCL patients with refractory disease, is now approved for use in the EU, since it gave long remissions in a subset of patients; at the median follow-up time of 15 months, 40% were still in CR (33).

For older patients or patients that are less physically fit and consequently don't tolerate stronger treatment options such as ASCT, more palliative-oriented treatment options are chosen in the relapse or refractory situation.

The prognosis of patients with DLBCL is in many ways closely dependent on the response to the initial treatment. Patients with early relapse (within a year from diagnosis) (29) or primary progressive disease (5, 34) have a very poor prognosis, regardless if they are fit for ASCT or other intensive chemotherapy, or referred to more palliative treatment choices. The median overall survival (OS) in patients under 70 years with primary refractory disease is only 10 months; 85% are deceased within 18 months and only 7% of patients reaches a more prolonged remission (5). For elderly patients with primary refractory disease the median OS is only 3.3 months (34).

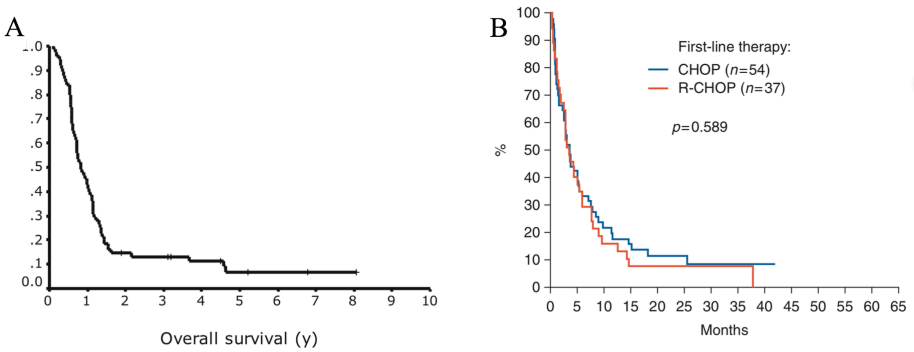

Figure 5. A) Overall survival (OS) of patients with primary refractory DLBCL who were 70 years of age or younger at the time of secondary progression. *5 year OS = 7%. Data from a population study of 1126 patients treated with R-CHOP, of whom 15 % had primary refractory disease. From (5). Reproduced with permission from Springer Berlin Heidelberg: Annals of Hematology© 2015.*
B) OS of elderly patients with primary refractory DLBCL*. Results from the RICOVER-60 study, that compared CHOP-14 with or without the addition of rituximab. The prognosis is equally poor regardless which treatment was given, with a median OS of only 3,3 months. From (34). Reproduced with permission from Oxford University Press: Annals of Oncology © 2017.*

In stark contrast to this scenario, patients who achieve CR after initial treatment, and are free from relapse 2 years after, have a survival that is similar to the general population (35). The risk of relapse decreases as time passes, and late relapses have a better response to ASCT than early (29).

1.3.5 Prognostic factors – the IPI

Up until recently, the only prognostic tool used in clinical practice to risk-stratify DLBCL patients has been the International Prognostic Index (IPI), that is based on the presence of five easily measured clinical parameters at diagnosis (7):

- Ann Arbor Stage ≥ III
- Elevated lactate dehydrogenase (LDH) in serum
- Performance status (ECOG) ≥ 2
- Age > 60 years
- Extranodal sites ≥ 2

The IPI was proposed in 1993 (before the introduction of rituximab), and could, based on the IPI-score, assign patients to four risk groups (low (IPI 0-1), intermediate-low (IPI 2), intermediate-high (IPI 3) and high-risk (IPI 4-5)) with different five year OS rates ranging from 73% to 26% (7).

An age-adjusted IPI (aaIPI) was also developed for patients < 60 years of age, which constituted only the first three IPI factors, i.e. stage, LDH and performance status (ECOG). Based on these parameters, four different risk groups could be separated, with different 5 year OS ranging from 83% (low risk, 0 factors) to 32% (high risk, 3 factors) (7). The aaIPI also proved to be valid among older patients, and is currently in Swedish clinical routine the preferred prognostic instrument rather than the original IPI.

The IPI is also valid in the immunochemotherapy era, however with a better overall outcome in all risk groups compared to the pre-rituximab IPI-studies. In modern data, the 3-year OS ranges from 91% in low-risk patients to 59% in high-risk patients *(Figure 6, opposite page)* (36).

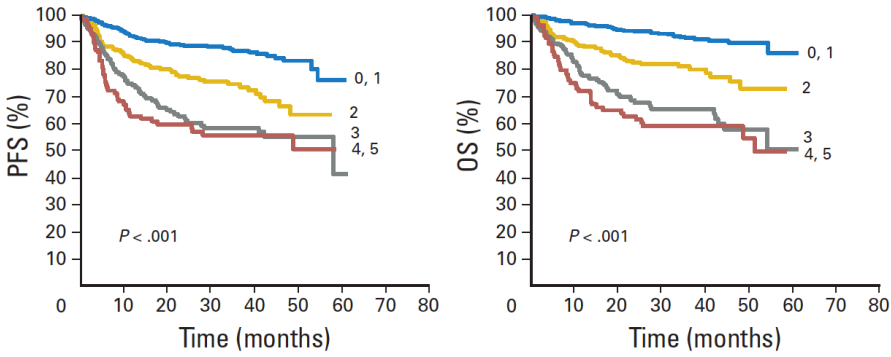

Figure 6. Progression free survival (PFS) and overall survival (OS) in relation to International Prognostic Index (IPI) in rituximab-treated patients. *Results merged from three trials: MabThera International Trial, MInT (n=380), MegaCHOEP trial (n=72) and RICOVER-60 trial (n=610). Compared to the pre-rituximab era, patients with IPI score 3 have similar outcome as those with IPI score 4-5, thus forming a new high-risk group with IPI score 3-5. From (36). Reproduced with permission from American Society Of Clinical Oncolgy: J Clin Oncol © 2010*

Despite high prognostic precision on a populational level, the IPI is however less useful in identifying the individual patients that will have an early relapse or primary progressive disease, i.e. the patients with dismal prognosis (37). Those patients can be found in all IPI risk groups, although naturally with a proportion that increases with higher IPI scores. But still – some patients with very poor prognosis are found among those with low IPI scores, and conversely, a large proportion of patients with high IPI scores, i.e. with poor prognosis as a group, still will be cured from the initial treatment. One can speculate that the clinical variables that make up the IPI are only composite surrogate markers for the biological risk profile and invasiveness of the lymphoma (as reflected in tumor stage, number of extranodal sites and serum LDH level), the patient´s response to the tumor (performance status) together with the physical state of the patient (age and performance status). More aggressive tumors and less fit patients give higher IPI-scores and thereby a worse prognosis on a group level.

There is a great need to find molecular prognostic markers to improve the outcome prediction in patients with DLBCL. Markers that better reflect the biological heterogeneity of the disease, and better explain the mechanisms behind resistance to immunochemotherapy. The following chapters describe the latest years advances in the molecular subclassification of DLBCL.

1.3.6 Cell-of-origin (COO)

An important discovery regarding the biological heterogeneity of DLBCL was made in the year 2000, when Alizadeh et al published the results from their gene expression profiling (GEP) study of frozen tumor samples from DLBCL-patients (38). They found, based on the gene expression profiles, two molecularly different forms of DLBCL, with gene expression patterns reflecting different stages of B-cell differentiation. The concept of "cell-of-origin" (COO) among DLBCL tumors was thus introduced. The first type, germinal center B-like (GCB) DLBCL, expressed genes resembling normal germinal center B-cells; whereas the second type, activated B-cell like (ABC) DLBCL, expressed genes that normally were found in activated peripheral B-cells in *in vitro* experiments. (ABC-DLBCL is now thought to have its probable natural counterpart in plasmablasts, i.e. B-cells having passed through the GC events and started to differentiate towards plasma cells, see *Figure 3)*.

As shown in *Figure 7*, the patients with GCB-DLBCL had significantly better survival than those with ABC-DLBCL. And interestingly, this was also true when patients within separate clinical risk groups was compared. Thus, the cell-of-origin concept could overcome some of the limitations of the IPI scoring discussed earlier, and with better precision find the patients in all clinical risk groups with highest risk of treatment failure.

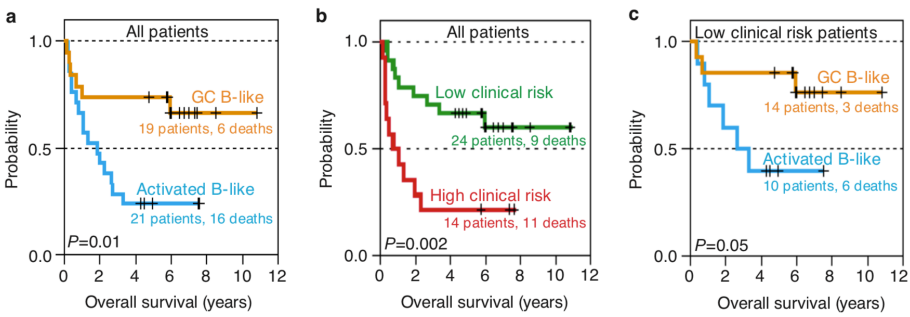

Figure 7. a) Overall survival (OS) of DLBCL patients based on gene expression profiling. *Patients with GCB-DLBCL had a significantly better OS than patients with ABC-DLBCL.*
b) OS based on clinical risk. *Low-risk patients (IPI score 0-2) had significantly better OS than high-risk patients (IPI score 3-5).*
c) OS of low clinical risk (IPI score 0-2) DLBCL patients grouped on their gene expression profiles. *Even within this low-risk group, ABC-DLBCL patients had a significantly shorter survival.*
From (38). Reproduced with permission from Macmillan Magazines Ltd: Nature © 2000

The Alizadeh study was made before the introduction of rituximab, but the differences in outcome between GCB- and ABC-DLBCL have later been confirmed in patients treated in the immunochemotherapy era, with rituximab added to the standard CHOP regimen (39, 40). However, since GEP is time consuming and not so easily standardized, attempts have been made to find an easier and more clinically accessible method to establish the COO of DLBCL tumors. In 2015, a 20-gene expression-based assay, the Lymph2Cx, could replicate the results from the Alizadeh study, but this time on formalin-fixed, paraffin embedded tissue in a standardized, much less time-consuming process (41). The Lymph2Cx assay has, with regard to its accuracy and rapid turnaround time, proven to be a useful tool in experimental settings, but has also potential to be implemented in future routine patient management.

Some researchers have found the GEP-based COO classification being a rather blunt tool, since there are several more naturally occurring subpopulations of B-cells than GCB and ABC. Centrocytes and centroblasts, cell types with major differences in function and phenotype, can for example both be found within the GCB group, and the COO classification doesn't differ between these cells. Recently, Danish researchers have used GEP to establish so called B-cell-associated gene signatures (BAGS) on immunophenotype-based flow-sorted normal B-cells, thus being able to differentiate between naïve, centroblast, centrocyte, memory and plasmablast B-cells. When applying these "BAGS" to clinical tumor samples, the different DLBCL tumors could be sorted into either of the subtypes, where the two most common, the centrocyte and centroblast subtype, showed survival differences, at least among GCB patients (42). The BAGS subtyping of DLBCL tumors have since been repeated using an easier to reproduce NanoString-based assay (43).

1.3.7 Cell-of-origin by immunohistochemistry

Since gene expression profiling is not accessible in most clinical settings, the COO of DLBCL tumors should be possible to translate to protein-based analysis with immunohistochemistry (IHC) on formalin-fixed, paraffin embedded tissue, already a routine procedure among hematopathologists involved in the diagnosing process of DLBCL. Indeed, since the year 2000, several IHC-based algorithms have been proposed for clinical use, of which the Hans algorithm is the most commonly used today (44). According to this, the tumor expression of three proteins, CD10, BCL6 and MUM, can separate the tumors into different COO-groups (44). The groups are here coined GCB- or non-GCB-DLBCL, the former group with significantly better survival than the latter, albeit not with as big differences between the groups as in the original GEP studies. Several other algorithms have been proposed, partly

based on other IHC markers (45-47). However, all IHC algorithms seem to have low concordance with GEP, show poor reproducibility and the prognostic information on individual basis is unreliable, especially in the immunochemotherapy era (40, 48). Still, IHC based algorithms to decide the COO, i.e GCB- or non-GCB-DLBCL, is for practical reasons widely used in clinical routine, simply because IHC already is an integrated part of routine hematopathology practice and GEP is not.

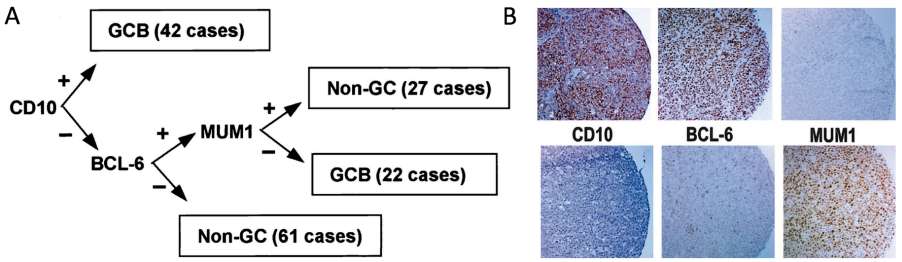

Figure 8. The Hans algorithm. *The cell of origin of DLBCL, either GCB- or non-GCB, can be established based on the expression of three proteins, CD10, BCL-6 and MUM1. B) Examples of IHC staining of a case of GCB-DLBCL (upper row), and non-GCB-DLBCL (lower row). From (44). Reproduced with permission from HighWire Press; American Society of Hematology: Blood © 2004*

1.3.8 GCB-DLBCL

DLBCL tumors with a germinal center B-like (GCB) cell-of-origin, resembles normal B-cells involved in the GC reaction, and express proteins that are normal for these cells, such as CD10 and BCL6 (see chapter 1.2.1) (38). Evident is also, just like in their normal counterparts, an ongoing somatic hypermutation driven by the enzyme AID (49).

An oncogenic pathway typical for GCB-DLBCL is caused by the t(14;18) translocation, found in around 35% of cases (see chapter 1.2.1). This translocation cause overexpression of BCL2, which has pro-proliferative and anti-apoptotic effects on the tumor cells (17, 50). 20% of GCB-DLBCL have mutations of the histone methyltransferase EZH2, which causes increased proliferation and hinders differentiation, thus forcing the tumor cells to continue in their GC mimicking state (51). Activation of the PI3K-AKT pathway is in 10% of GCB-DLBCL cases caused by deletion of the PI3K inhibitor PTEN (52, 53). Translocations involving the transcription factor and pro-proliferative protein MYC is found in 10-15% of cases (see chapter 1.2.1), and is associated with worse prognosis (54), especially if there are simultaneous translocations involving BCL2 and/or BCL6 (55), so-called double- or triple hit lymphomas, which are described in more detail later.

1.3.9 ABC-DLBCL

DLBCL tumors with an activated B-like (ABC) cell-of-origin, have their normal counterpart in plasmablasts, B-cells having passed through the GC events and started to differentiate towards plasma cells (19, 38). Typical for ABC-DLBCL is the continuous activation of the NF-κB pathway, which leads to cell proliferation and survival, and is believed to be the major causal factor behind the worse prognosis among ABC-DLBCL patients (56, 57). There are several oncogenic pathways behind this activation; like chronic B-cell receptor signaling caused by CD79a/b-mutations (20% of cases), constitutive NF-κB activation through mutations of CARD11 (10%) or MYD88 (35%), or inactivation of TNFAIP3 (30%) (16, 58, 59).

1.3.10 Double Hit and Double Expressor Lymphoma

Among DLBCL tumors overexpression of MYC and BCL2 can be detected by standard immunohistochemical methods, and tumors with elevated levels of both proteins are either so-called double hit lymphomas (DHL) (55) or double expressor lymphomas (DEL) (60). DHL is defined as a B-cell lymphoma with a MYC-rearrangement in pair with either a translocation involving BCL2 (most commonly) or BCL6. Tumors with all three rearrangements are termed triple hit lymphomas (THL). Patients with DHL/THL have an aggressive clinical presentation, often an advanced stage lymphoma, with elevated LD and frequent extranodal engagement including bone marrow or CNS (61, 62). This often results in high IPI-scores and the overall survival of this patient group is poor. DHL/THL can be found in around 10 % of DLBCL, but almost exclusively among patients with GCB-DLBCL, i.e. the cell-of-origin group with better prognosis (63). DHL/THL-status probably explains a large proportion of the treatment failures that still occurs in this patient group. FISH-analyses of those gene rearrangements now is part of the routine hematopathological work-up of DLBCL tumors, as the presence of double or triple translocations urges for more aggressive treatment. Even though no prospective studies have been made on these patient groups, retrospective studies indicate that treatments such as dose-adjusted R-EPOCH (a more intensive "R-CHOP" with added etoposide) possibly alternating with CNS-penetrating high-dose methotrexate and cytarabine combinations should be considered due to the inadequacy of R-CHOP for the DHL patients (64).

In DEL, the high levels of MYC and BCL2 are not caused by translocated genes, but instead by gene amplifications. DEL is more common than DHL, and found in about 25-30% of DLBCL cases, but in contrast to DHL almost all belonging to the ABC-DLBCL cell-of-origin subtype (60, 65). Patients

with DEL have an adverse prognosis compared to patients without DEL, but still a better prognosis than patients with DHL (60).

1.3.11 Whole exome sequencing and mutational investigations in DLBCL

With today's rapid evolution of techniques in whole exome sequencing (WES) and mutational analyses of large number of samples, progress have lately been seen in investigation of genetic drivers in DLBCL, i.e. genes essential for the survival and expansion of the malignant clone (66-68). Compared to RNA-based COO, which defines the differentiation of the cell depending on differentially expressed genes, WES can detect common mutations, which makes it possible to subclassify DLBCL patients according to clusters of mutations of functionally associated genes. In one study, almost half of 574 DLBCL patients could, according to the distribution of genetic alterations, be sorted into four subgroups, with different mutational pattern, and different survival (67). Similar and partly overlapping mutational subgroups were found in another large study on 304 DLBCL patients, however with some contradiction regarding prognosis on one of the GCB-specific subgroups (68). In the largest of these new studies, where 1001 DLBCL patients were investigated for mutations, a median of almost 8 mutations were found per case (66). The study identified 150 potential driver genes with recurrent mutations in DLBCL, and from the mutational pattern of each patient sample, a "genomic risk" was calculated, and the patients were sorted into either a genomic low risk or high risk group, the former with a significantly better survival. Prognostication according to IPI, cell-of-origin and DHL-status all proved to be valid within this large patient material; however, the genomic risk model was robust and highly significant also when tested within each of those known risk groups *(Figure 9)*.

In DLBCL cell lines, selective knock-out of the 150 identified driver genes via CRISPR screening, could identify 35 genes whose knockout resulted in decreased viability of the cell lines, thus identifying them as functional oncogenes. Those 35 genes consequently become potential drug targets in DLBCL, and indeed, therapeutic substances targeting nine of them are already under investigation in clinical trials or already approved for other indications (66).

With so-called "liquid biopsies" using serum from DLBCL patients, mutational analyses have also been possible to perform on cell-free DNA that is shed from tumor cells undergoing apoptosis (69), which make the mutational status more easily accessible, and also possible to follow longitudinally during and after treatment.

In summary, the early results from studies on whole exome sequencing and mutational analysis of DLBCL patients are very promising, both in the prognostic/predictive perspective, but also when exploring novel targets for treatment. Since high throughput sequencing in different forms are becoming more and more applied in practical clinical routine of hematopathologists, this field of research will probably expand greatly in the next years.

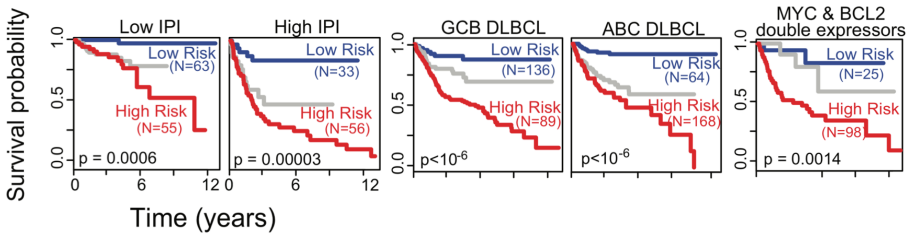

Figure 9. Genomic risk model applied within different known risk groups of DLBCL *(i.e IPI, cell-of-origin and MYC/BCL2 double expressors). Blue indicates genomic low risk, and red genomic high risk. Grey = all patients within the known risk group. In this material, the genomic risk model can significantly stratify survival within the known risk groups that was tested. From (66). Reproduced with permission from Elsevier: Cell © 2017*

1.3.12 Chemoresistance and tailored treatments

The reasons for the differences in outcome between the various molecular subgroups of DLBCL have been examined with regard to resistance to immunochemotherapy, here named *chemoresistance*. In the ABC-DLBCL, with its post germinal center phenotype, there are several postulated innate chemoresistance mechanisms that could explain the shorter survival, for example the chronic B-cell receptor (BCR) signaling and downstream activation of the transcription factor NF-κB.

Inhibition of NF-κB in vitro will kill ABC-cells, but not GCB-cells (58, 70), and in vivo direct or indirect inhibition of the pathway by drugs like ibrutinib (71), bortezomib (72) and lenalidomide (73, 74) has in relapse situations shown effects on ABC-DLBCL but not GCB-DLBCL. The drugs are currently tested in various stages of first-line treatment studies, however so far not with any positive results: a prospective randomized study of newly diagnosed DLBCL, REMoDL-B, could not show any differences in survival between patients having bortezomib or placebo added to R-CHOP, not even in the ABC subgroup (75). A placebo-controlled study of frontline ibrutinib added to R-CHOP to patients with ABC/non GCB DLBCL is currently active, (ClinicalTrials.gov Identifier: NCT01855750), and there is an ongoing study

to compare the addition of lenalidomide to frontline R-CHOP in newly diagnosed DLBCL patients (76).

Subgroups within the GCB-DLBCL also have genetic alterations leading to chemoresistance, like the 35% of cases with constitutional activation of the anti-apoptotic protein BCL2 via t(14,18) (50), which in itself was a negative prognostic marker before the rituximab era, and still is if coupled with MYC-rearrangements, i.e. DHL (77). Activation of the PI3K-AKT pathway is in 10% of GCB-DLBCL cases caused by deletion of the PI3K inhibitor PTEN (52, 53), which make the tumors resistant to rituximab. Targeted novel agents against BCL2 and Pi3K could have theoretical advances in patients with GCB-DLBCL compared to ABC-patients (78), and indeed, venetoclax, a BCL2 inhibitor, has been positively evaluated as a single agent in the relapse situation (79), and is currently tested as an addition to R-CHOP or G-CHOP (G=Obinutuzumab, anti-CD20 monoclonal antibody) in a frontline setting for newly diagnosed DLBCL patients (ClinicalTrials.gov Identifier: NCT02055820).

Acquired chemoresistance, regardless of COO, is also a feature in DLBCL cases. MDR1, an ATP-dependent efflux pump that can transport multiple drugs, for example doxorubicin and vincristine, out of malignant cells(80), have been shown to be increased in chemoresistant non-Hodgkin lymphoma cells compared to treatment-naïve samples (81-83). Resistance to rituximab in B-cell lymphoma cell lines can be linked both to a reduction of CD20 expression, and also to a down regulation of pro-apoptotic proteins Bax and Bak (84, 85).

Other cellular functions that are found to be important contributors of drug resistance in DLBCL are histone deacetylases (HDACs), enzymes that make the chromatin more compact, and thus repressing transcription of pro-apoptotic genes (86). Novel HDAC inhibitors have, probably by restoring the expression of repressed genes leading to cell-cycle arrest, differentiation and apoptosis (87), been shown to regain DLBCL tumor sensitivity to immunochemotherapy in previously resistant cases (88-90). Also valproic acid, a classic anti-epileptic drug with recently discovered HDAC-inhibitory properties, has shown effects as a sensitizer for CHOP-induced death among DLBCL cell lines (91), and has been used in a phase I trial together with R-CHOP for untreated DLBCL patients with promising results (92).

Other proteins that have been linked to chemoresistance in B-cell lymphomas are the tumor suppressor p53 (93-95) and also the type II topoisomerase, the latter however so far contradictory regarding if its occurrence makes the tumors more (96) or less (97) sensitive to chemotherapy. Another interesting

take on chemoresistance is the proposed “chemoresistant niche”, that could promote survival of residual lymphoma cells due to microenvironmental changes caused by the release of interleukin 6 (IL-6) and tissue inhibitor of metalloproteinases 1 (TIMP1) in the thymus in response to doxorubicin treatment (98). In summary, there are several known mechanisms behind both innate and acquired chemoresistance in DLBCL, mechanisms that in studies have been targeted by novel treatment agents both in front-line and relapsed settings, but so far none of these have made their way into the standard treatment of the disease, i.e. R-CHOP.

1.4 A NEED FOR NEW PROGNOSTIC AND PREDICTIVE MARKERS

The primary treatment for patients with DLBCL have for long times, despite the great heterogeneity of the disease, been picked from a rather limited palette of regimens. With the exception of some patients with stage I disease, in which case the number of treatment courses can be reduced in favor of radiotherapy (99), variations of anthracycline-based immunochemotherapy, i.e. R-CHOP, for 6-8 courses have been the treatment chosen for almost all patients, and the guide for deciding the exact regimen have been the clinically based IPI (100-102). However, in recent times, treatment algorithms in care programs have also taken into account information about COO (GCB- or non-GCB by immunohistochemical methods) or DE-/DH-status, and opted for slightly differentiated treatments depending on those factors, for example recommending stronger prophylaxis for CNS-recurrence among patients with DE-tumors, or a choosing a Burkitt-like treatment for patients with DH-tumors (103). Several novel lymphoma agents (like lenalidomide, bortezomib and ibrutinib) (71-74), have been tested, but not yet made their way into standard treatment of DLBCL outside of clinical studies. Also, even with the great impact of the seemingly robust GEP-based COO-data from the original Alizadeh study (38), the results regarding the differences in survival between patients with GCB- and ABC-DLBCL have later been questioned in large studies on patients receiving immunochemotherapy: analyses of survival with respect to COO (established through GEP) from the RICOVER-60 and MegaCHOEP trials could not reproduce the differences in survival between GCB and ABC DLBCL patients (104), neither could data from the REMoDL-B study (75).

So, despite the progress in molecular subclassification, and the robustness and reliability of the old, clinically based IPI scoring, there is still no available system that with precision can identify the individual patients at highest risk of treatment failure, even though the newly developed genomic risk model

(66) seem to be the best effort so far. In fact, the actual failure to initial DLBCL treatment is the strongest indicator of bad prognosis we know of (5, 6, 34), but the prognostic tools we use today are not sufficient for detecting those patients early enough in the course of disease (37). Also, we lack the means of counteracting the poor prognosis of those patients, not least because the mechanisms behind treatment failure are not sufficiently explored. Consequently, there is a strong need for new prognostic markers to identify the patients at highest risk, and also a strong need for a better characterization of the underlying mechanisms behind treatment failure.

1.4.1 Proteomics

Through gene expression profiling (GEP), that measures the patterns of mRNA mirroring the differentiation of a cell, we have reached a better understanding of the background of DLBCL tumors via the cell-of-origin (COO) concept (38). However, despite the molecular basis of the concept, and despite it being based on a global assessment of gene expression, the COO status has not proven to be reliable or sufficient in neither prognostication or clinical decision-making of individual DLBCL patients. Also, GEP is not a standardized method that can be applied to routine clinical practice. More clinically applicable immunohistochemical methods to establish the COO, like the Hans algorithm (44), have also had limited clinical value, and the results have been contradictory and difficult to reproduce (105-108). Also, those methods are based on expression of only a few different proteins, just like investigation of other, non-COO based molecular biomarkers like MYC and BCL2, whereas a global investigation of protein expression should be more representative of tumor biology. Based on data from the human genome project, the number of protein coding genes in humans have been estimated to around 20 000 (109), but the number of potential proteins or variations of proteins is probably even higher, due to for example mRNA splicing and posttranslational protein-modification like phosphorylation and glycosylation (110). Thus, the gene expression profile of a cell or tissue doesn´t necessarily reflect the protein expression, as proven in a frontline study of yeast, in which simultaneous measures of mRNA-levels and protein analyses showed a low correlation (111).

Another way to search for potential protein biomarkers in DLBCL could be by a proteomic analysis. With proteomics, the total protein expression pattern, i.e. *the proteome*, of a sample is explored. Proteomics have evolved from classical methods like two-dimensional gel electrophoresis, but now, because of the high sensitivity and speed, different variations of mass spectrometry (MS) is used for modern proteomic analyses (112). In recent years different proteomic approaches have been applied, mostly on cell lines or in animal

models, to investigate the proteome in DLBCL tumor material (113-116). Neither of these studies have addressed the prognostic challenge in DLBCL, nor the mechanisms behind chemoresistance, but MS-based proteomic analyses of both DLBCL cell lines (117) and tumor material from patients (118) have been successful in differentiating between GCB- and ABC-DLBCL, proving the validity of the method.

Since treatment refractoriness and early relapse of disease are the biggest known risk factors in DLBCL, an investigation of the tumor proteome from those high-risk patients, compared with patients with low risk disease, i.e. patients that are cured from standard immunochemotherapy, could give vital information of the differences in tumor biology between those patient groups, regardless of their COO. This comparison could potentially find protein signatures typical for chemoresistant disease, and also give clues to the molecular mechanisms behind treatment failure, which in turn could help us find cellular targets of novel treatments.

1.4.2 Metabolomics

The metabolome is made up from the different metabolites found in an investigated sample, e.g. serum, and is the final down-stream product of the metabolic processes in the organism, taking into account both healthy and physiological as well as pathological processes (119). This means that in a patient with a disease, the pattern of serum metabolites not only could give information about the disease, but also about the overall physical state of the patient. The different levels of "omics" capture the biological levels in an organism from different perspectives. If the transcriptome reflect the differentiation and the proteome reveals more about the phenotype of the cell or tissue, the metabolome is more representative for the phenotype of the entire organism. Events which *might happen* are captured by genomics, events which *are happening* are captured by proteomics, and events which *have happened* are captured by metabolomics (120). Metabolomics, where low molecular weight molecules are detected in body fluids or tissues, is an emerging and promising tool for diagnostic and differentiating purposes in cancer. Serum metabolomic approaches for diagnostic and prognostic purposes in different malignancies, among them acute myeloid leukemia (121) and breast cancer (122, 123), have shown promising results. Serum from DLBCL patients is, compared to representative tumor material, easily accessed. As discussed in the previous chapter, the DLBCL patient group with worst survival rates is the patients with primary refractory disease or early relapse. Investigation of the serum metabolome from this high-risk patient group, and comparison with the serum metabolome from patients that were cured from standard immunochemotherapy, could potentially give valuable

information about DLBCL prognostication and also of mechanisms behind treatment failure.

2 AIM

The overall aim of this thesis was to search for novel prognostic and predictive biomarkers in diffuse large B-cell lymphoma (DLBCL), and also to investigate the mechanisms behind chemoresistance.

In addition, we had more specific aims in the different papers:

- To use a quantitative proteomic analysis in fresh-frozen tumor tissue from two groups of DLBCL patients who had been treated with modern immunochemotherapy with totally different clinical outcome, that is, (i) early relapse/refractory patients (REF/REL) and (ii) long-term progression-free patients (CURED), in order to explore possible differences in global protein expression (Paper I).

- To use a quantitative proteomic approach to analyze the global protein expression in formalin-fixed paraffin-embedded tumor tissues from a larger number of (i) REF/REL and (ii) CURED DLBCL patients, with the aim of revealing new mechanisms involved in immunochemotherapy resistance (Paper III).

- To further investigate the possible influence of actin organization and remodeling on drug resistance that we found in Paper I (Paper III).

- To use ^{1}H NMR spectroscopy to compare the serum metabolome from i) REF/REL and ii) CURED DLBCL patients, to determine if differences in clinical outcome could be correlated to different metabolomic profiles (Paper II).

3 PATIENTS AND METHODS

3.1 PATIENTS

All studies were performed on two DLBCL patient subgroups, that were sorted retrospectively based on their response to initial treatment: (i) patients with primary refractory disease or relapse within 1 year after diagnosis (REF/REL); and (ii) patients that were long-term progression-free with a follow-up, from diagnosis, of at least 5 years, clinically considered cured from disease (CURED). The patients were identified from the Swedish Lymphoma Registry, and patients from western Sweden, with sufficient tumor material or serum available in biobanks could be included. All patients had received modern immunochemotherapy (i.e., R-CHOP: the monoclonal CD20-antibody rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone) with curative intent. Patients with primary mediastinal large B-cell lymphoma, primary central nervous system lymphoma, HIV-related lymphoma and transformed lymphoma were excluded. Clinical information was obtained from the patient medical records including treatment and progression-free and overall survival. Ethical approval for the studies was obtained from the Regional Ethics Review Board, Göteborg.

3.1.1 Patients in paper I

All adult patients with de novo DLBCL diagnosed between January 2004 and December 2008 at the Section of Hematology of Sahlgrenska University Hospital and treated with curative intent were identified. From each of the two subgroups, i.e. REF/REL or CURED patients, five patients were selected on the basis on the availability of freshly frozen tumor tissue samples in biobanks. To avoid obvious morphologic differences between the two groups, the pathologists carefully examined the tumor tissue, and only samples with evenly distributed blasts without signs of necrosis or abundant visual stroma were selected. The clinical characteristics of the patients are shown in *Table 1*.

Table 1. Clinical characteristics of the patients from paper I.

Pat No	***Age, y***	***Sex***	***Ann Arbor Stage***	***B-Symptoms (Yes/No)***	***S-LDH***	***Performance (ECOG)***	***aaIPI***	***Outcome***
1	70	M	III	Yes	High	2	3	REF/REL
2	50	M	IV	No	Normal	0	1	REF/REL
3	60	F	IV	Yes	High	3	3	REF/REL
4	85	M	I	No	Normal	0	0	REF/REL
5	63	M	II	No	High	0	1	REF/REL
6	55	F	I	No	High	0	1	CURED
7	74	M	III	No	Normal	0	1	CURED
8	58	F	II	Yes	High	0	1	CURED
9	75	M	II	No	High	0	1	CURED
10	63	F	III	No	Normal	0	1	CURED

aaIPI: Age adjusted International Prognostic Index; CURED: Progression-free with a follow-up of at least 5 years; REF/REL: Progressive disease during treatment or early relapse, i.e. relapse within 1 year after completion of treatment; S-LDH: Serum lactate dehydrogenase.

3.1.2 Patients in paper II

All adult patients with de novo DLBCL diagnosed between January 2004 and December 2012 at the Section of Hematology of Sahlgrenska University Hospital and neighboring hospitals, and treated with curative intent were identified. Patient with biobanked serum samples available could be included. Again, in order to study patients with clearly different clinical outcome, we selected two subgroups on the basis of response to initial treatment: i)REF/REL patients (n=27), and ii) CURED patients (n=60). For metabolomic analysis, serum samples from the time of diagnosis (before start of treatment) were obtained from the biobank at the Department of Virology at Sahlgrenska University Hospital, where serum from lymphoma patients is tested for hepatitis A–C and HIV prior to the start of immunochemotherapy. Surplus serum is frozen and stored at -80° C until further use. The serum samples were not collected in a standardized fashion, but in a clinical routine setting.

Patient characteristics are given in Table 2. We found no differences in age or sex distribution between the patient groups but, as expected, REF/REL patients had in larger extent high aaIPI scores. On the other hand, the proportion of GCB versus non-GCB did not differ significantly between the groups (Table 2).

Table 2. Clinical characteristics of the patients from Paper II

	CURED patients n (%)	REF/REL patients n (%)	p-value
Total number of patients	60 (69)	27 (31)	
Male	30 (50)	17 (63)	n.s.
Female	30 (50)	10 (37)	
Age, median (range) years	61 (20-88)	67 (29-85)	n.s.
Ann Arbor Stage III/IV	27 (45)	18 (67)	n.s. (p=0.072)
Serum-LDH elevated	33 (55)	22 (81)	0.020
Performance Status (ECOG) 2-4	17 (28)	13 (48)	n.s. (p=0.091)
Cell-of-origin (GCB/Non-GCB)	32/28	16/7*	n.s.
aaIPI 2-3	23 (38)	18 (67)	0.014

aaIPI: age adjusted International Prognostic Index; CURED: Progression-free with a follow-up of at least 5 years; REF/REL: Progressive disease during treatment or early relapse, i.e. relapse within 1 year after completion of treatment; GCB: Germinal Center B-cell-like; LDH: lactate dehydrogenase.
** REF/REL n = 23.*

3.1.3 Patients in Paper III

We again selected two DLBCL patient subgroups based on their response to initial treatment: (i) REF/REL patients; and (ii) CURED patients. From the Swedish Lymphoma Registry, a total of 270 adult DLBCL patients in western Sweden, diagnosed between 1 January 2004 and 31 December 2014, belonged to one of these subgroups. All patients received immunochemotherapy (R-CHOP). Archived formalin-fixed, paraffin-embedded (FFPE) tissue sections from the time of diagnosis were re-evaluated. Only cases showing large areas of blasts and with sufficient amount of tumor tissue were included, which were found in a total of 97 DLBCL patients, 44 patients from the REF/REL group, and 53 patients from the CURED group. Patient characteristics are shown in Table 3. REF/REL patients were older and had a higher percentage of high-risk aaIPI score, MYC-positive and BCL2/MYC double expressors compared to CURED patients. There were no statistically significant differences in sex distribution, proportion of GCB versus non-GCB, or Ki67 index.

Table 3. Clinical characteristics of the patients from Paper III

	CURED patients	**REF/REL patients**	**p-value**
	n (%)	**n (%)**	
Total number of patients	53 (55)	44 (45)	
Male	25 (47)	29 (66)	n.s.
Female	28 (53)	15 (34)	
Age, median (range) years	64 (22-84)	71 (38-80)	0.03
aaIPI 2-3	19 (36)	26(59)	0.02
Cell-of-origin (GCB/non-GCB)	30/23	17/27	n.s.
Ki-67, % (median)	74	79	n.s.
BCL2 (≥50%)	37 (70)	39 (89)	0.03
MYC (≥40%)	7 (14)*	18 (42)*	0.003
BCL2/MYC double expressors	6 (12)*	18 (40)*	0.001

aaIPI: age adjusted international prognostic index; CURED: Progression-free with a follow-up of at least 5 years; REF/REL: Progressive disease during treatment or early relapse, i.e. relapse within 1 year after completion of treatment; GCB: Germinal Center B-cell-like.

**CURED n = 50 and REF/REL n = 43.*

3.2 GENERAL STUDY DESIGN

As mentioned, all studies (Paper I-III) were retrospective and investigated two distinct clinical subgroups of DLBCL patients; (i) (REF/REL); and (ii) (CURED) patients. At the time of the studies approximately 20-25 % of patients in western Sweden that were diagnosed with DLBCL during the different time intervals described in earlier chapters, could be assigned to either the REF/REL or the CURED group. The rest of the patients did not at that time meet the inclusion criteria, most often because the time of follow-up was to short (less than five years). Other reasons for not being included could be that the patient didn´t receive immunochemotherapy with curative intent but instead a treatment with palliative approach, or that the patient complied with any of the exclusion criteria described in chapter 3.1.

In all studies, freshly frozen or paraffin-embedded tumor material (Paper I and III) or patient serum (Paper II) that had been collected at the time of diagnosis, i.e. before the start of immunochemotherapy treatment, was analyzed and compared between the patient groups. The availability of this patient material in biobanks was a limiting factor regarding the final number of patients that could be included in all three studies. In Paper III for example, of the 270 potential study participants, representative FFPE tissue sections was found from 97 (36%) of the patients.

3.3 METHODS

3.3.1 Proteomics by mass spectrometry (MS)

This chapter discusses MS-based proteomics in general. More specific descriptions of the two different proteomic approaches used in Paper I and II are found in the following chapters.

The principle behind mass spectrometry is to measure the masses of different ionized, i.e. electrically charged, compounds in a sample that has been transferred to gas phase (most often from liquid phase). The mass-to-charge (*m/z*) ratios of the compounds are identified, quantified and plotted in a mass spectrum. A mass spectrum is a histogram where different compounds (with different *m/z* values) are represented by different peaks (124). The position of a peak along the x-axis in the mass spectrum corresponds to the molecular mass of the compound, and the height of the peak corresponds to the relative abundance of the substance in the sample. Proteins have too high molecular weight to be properly handled in MS, and a degradation to peptides must be done before analysis. Also, a separation of the sample is necessary before it enters the mass spectrometer, since the technique only can handle a limited

number of peptides per time unit. The separation is done with high pressure liquid chromatography (LC), but before that, some kind of pre-separation is often made on complex samples (125, 126). Many low molecular weight substances can in MS directly be identified based on their m/z ratios, but more complex compounds, like peptides, need to be fragmented and reanalyzed in a second, tandem MS phase (MS/MS) before a correct identification can be made (125). MS can measure the relative abundance of a substance in a sample, but for an exact quantification some kind of reference is needed. This is handled differently in the different proteomic approaches used in paper I and III, but a principal workflow of MS-based proteomics are showed below.

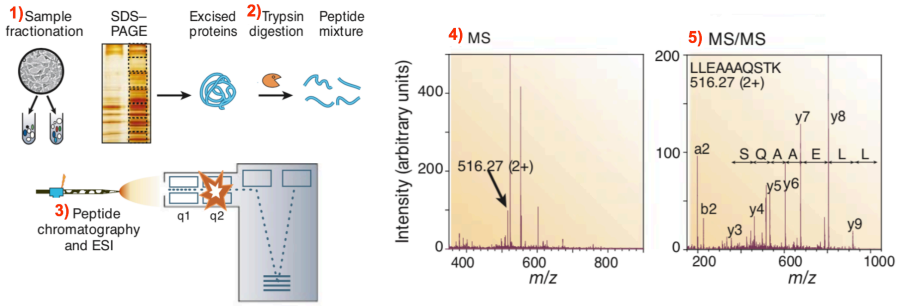

Figure 10. Principal workflow of MS-based proteomics.
1) *In complex protein samples (like homogenized tumor biopsies), the proteins are isolated from the rest of the tissue with different techniques like biochemical fractionation or affinity selection. The protein mix are then sometimes pre-separated by gel electrophoresis, to obtain smaller subsets of proteins that (for reasons of higher resolution) are analyzed in the mass spectrometer in sequence instead of simultaneously.*
2) *The proteins are digested into smaller peptides by trypsin, an enzyme that cleaves peptide chains at the carboxyl sides of the amino acids lysine and arginine.*
3) *The peptide mixture is fed into the high pressure liquid chromatograph (HPLC), which separates them according to their molecular characteristics, among them charge and hydrophobic properties. The peptides finally go through electrospray ionization (ESI), in which they are ionized and in gas form sprayed into the mass spectrometer.*
4) *The masses and intensities (relative abundances) of the peptides passing through the mass spectrometer are measured, and plotted in a histogram, the mass spectrum. The position of a peak along the x-axis in the mass spectrum corresponds to the molecular mass of the peptide, and the height of the peak corresponds to the relative abundance of the peptide in the sample.*
5) *Since many peptides have similar masses, the exact identity of the peptide cannot be decided in from the first MS, but a second step is necessary, the tandem MS (MS/MS) analysis. In this step the peptides from the first MS step are fragmented, and a new mass spectrum is made from those fragments, stored for later matching against protein sequence databases.*
From (125). Reproduced with permission from Springer Nature; Nature © 2003.

3.3.2 SILAC-based quantitative proteomic analysis on freshly frozen tumor tissue (paper I)

In paper I, a quantitative LC-MS/MS proteomic analysis with the SILAC-based technique was used, for characterization of proteins from freshly frozen tumor tissue (127). *Stable isotope labeling of amino acids in cell culture* (SILAC), is a method in which an isotope-labeled mix of proteins is analyzed alongside the protein mix from the tumor, thus creating an internal reference in each experiment, which in turn make a comparison between patients possible (128). A detailed description of the method is given in paper I; here below follows a shorter version, with references to the graphic description of the workflow given in *Figure 11*.

To enable quantification of a broad number of proteins, and for production of a reference protein mix, five DLBCL cell lines were metabolically labeled with stable isotopes. This was done by using a cell culture medium in which the amino acids lysine and arginine was replaced by $^{13}C_6$-lysine and $^{13}C_6$-arginine, which meant that on all positions, the proteins in the cultured cells had been incorporated with "heavy" versions of these two amino acids. The isotope-labeled cells were harvested, and proteins were extracted. Equal amounts of proteins from the five cell extracts were mixed to produce a SILAC-reference mix (*Figure 11a*).

From each tumor sample proteins were extracted, and aliquots with equal amounts of protein from each sample (containing proteins with "light" arginine and lysine) were mixed with aliquots with corresponding amounts of protein from the SILAC reference mix (containing proteins with "heavy" arginine and lysine), thus creating 1:1 mixtures of "wild type" proteins from tumor samples and isotope-labeled proteins from the reference mix (*Figure 11b*).

To enable a higher resolution in the MS analysis, the proteins were pre-separated with two parallel methods: SDS-PAGE gel electrophoresis (*Figure 11c*) and via the FASP protocol (129) (*Figure 11b*).

In both experimental paths, proteins were enzymatically digested into peptides with trypsin. Trypsin cleaves peptide chains at the carboxyl sides of the amino acids lysine and arginine, which rather elegantly results in peptides containing *either* one arginine *or* one lysine amino acid. This in turn means that every peptide from the tumor sample will contain one "light" arginine *or* "light" lysine, and every peptide from the SILAC reference-mix will contain one "heavy" arginine *or* "heavy" lysine. $^{13}C_6$-lysine and $^{13}C_6$-arginine each contain six ^{13}C atoms instead of the normal ^{12}C atoms, and therefore weigh

exactly 6 Da more than their "light" counterparts. Consequently, every peptide from the SILAC reference mix will weigh exactly 6 Da more than the corresponding peptide from the tumor sample.

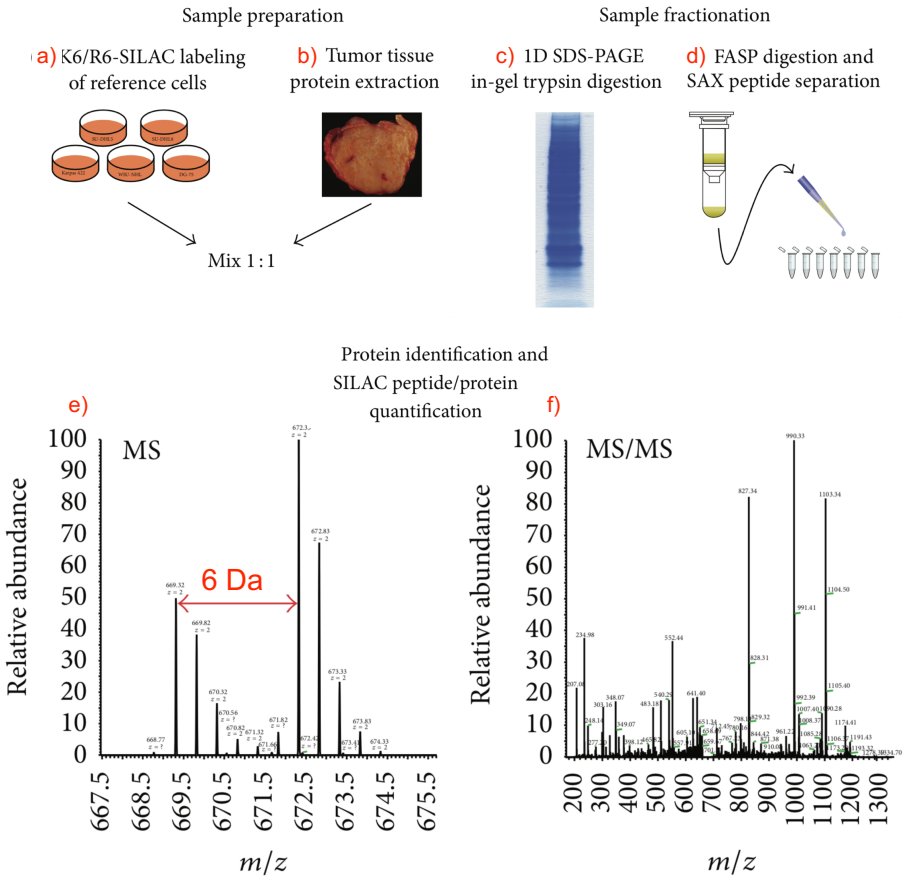

Figure 11. The experimental workflow of the SILAC proteomic analysis.
K6 : $^{13}C_6$-lysine; R6 : $^{13}C_6$-arginine. See chapter 3.3.2 for further definitions. From (115), an Open Access publication from Hindawi.

The peptides were again separated, now in high pressure liquid chromatography, before entering the MS and MS/MS as described in the previous chapter and *Figure 10*. However, with the SILAC method, every peptide in the mass spectrum generate two peaks: one to the left representing the peptide from the tumor mix, and one to the right, exactly 6 Da heavier, representing the same peptide from the SILAC reference mix (*Figure 11e)*. Since 1:1 mixtures of "wild type" proteins from tumor samples and isotope-labeled proteins were used in all experiments, the ratio of the relative

abundance of the "heavy" and "light" peptides (H/R ratio) could be used for relative quantification and comparison of protein levels in the different tumor samples. In the last phase of the analysis, the tandem MS/MS was performed on fragmented peptides for peptide identification (*Figure 11f)*.

3.3.3 Protein identification and quantification (SILAC-based proteomics, paper I)

Mass spectra from the MS/MS analyses of fragmented peptides were processed using the MaxQuant software version 1.2.0.18 (130), and peptides were identified with the Andromeda search engine (131), integrated in the MaxQuant package. Searches were then performed against the human subsection of the UniProtKB database, to couple each identified peptide to the corresponding protein. Peptide and protein false discovery rate (FDR) was set to 0,01. The ratio of the relative abundance of the "heavy" and "light" peptides (H/R ratio), was used for relative quantification and comparison of protein levels between the patients. A two sample *t*-test was performed to determine significant differences in protein ratios between the groups.

3.3.4 Western Blotting (paper I)

Expression levels of selected proteins were validated by immunoblot analysis of tumor protein extracts from all patients. Equal amounts of protein were separated on gels, transferred to nitrocellulose membranes and incubated with the primary antibody. For signal detection, membranes were then incubated with HRP (horseradish peroxidase) conjugated secondary antibody. The SILAC reference mix was used as a control.

3.3.5 Tandem mass tag-based quantitative proteomic analysis on formalin-fixed, paraffin-embedded tumor tissue (paper III)

In the first proteomic study in paper I, the number of participating patients was limited to ten. The main reason for this low number was the scarce availability of freshly frozen tumor tissue in the biobanks, but another important factor was the complexity of the workflow making the analysis of each sample very time consuming. For the second proteomic study in paper III, we selected a method that enabled us to include more patients. Tandem mass tag (TMT)-based proteomic analysis is, compared to the SILAC-based method, a principally different proteomic approach, in which several patient samples can be analyzed simultaneously (132). TMTs are small chemical labels, that in vitro are attached to the N-terminal of trypsin-digested peptides prior to the LC-MS/MS procedure. In this experiment we used 10-plex TMTs, where each TMT is isotope-labeled, but with the isotopes attached in unique

positions for each TMT. This means that the ten different tags have identical mass and chemical properties, i.e. being isobaric, thus behaving identically in high pressure LC and MS. However, after peptide fragmentation in the tandem MS/MS phase, the resulting mass spectra are different for peptides with different TMT-labels, which makes relative quantification of peptide levels possible (133).

The proteomic analysis was performed on formalin-fixed, paraffin-embedded (FFPE) DLBCL tumor tissue from time of diagnosis. A reference mix assembled from 9 random samples (5 REF/REL and 4 CURED patients), was used for ratio calculations. A thorough description of the workflow is given in paper III (134) and in its additional data.

3.3.6 Immunohistochemistry (paper III)

Tissue micro array (TMA) blocks were constructed from the original FFPE-blocks, after selection of suitable tumor areas. Deparaffinized sections (4 μm) from TMA blocks were stained with antibodies for CD10, BCL6, MUM1, BCL2, Ki67, c-MYC, RPS5, RPL17, Enah/Vasp-like protein and anti-MARCKS-like protein. GCB/non-GCB classification was determined according to the Hans algorithm (44). For BCL2 and MYC, cut-off values of ≥ 50% and ≥ 40 %, respectively, were used, and patients expressing both BCL2 and MYC were designated "double expressors" (1). Two pathologists, blinded to the clinical outcome, independently evaluated the immunoreactivity toward RPS5, RPL17, Enah/Vasp-like protein and MARCKS-like protein. As all of these four antibodies showed cytoplasmic staining, a visual approach was used and cases were divided in two categories: negative/weak/intermediate *versus* high intensity based on comparison to all cases for each TMA. Concordance was >80% and the remaining cases were examined in a double-head microscope to reach consensus.

3.3.7 Protein network analysis

The differentially expressed proteins in paper I were analyzed using three different resources for network analysis:

- DAVID (Database for Annotation, Visualization and Integrated Discovery) Bioinformatic resources version 6.7 (https://david.ncifcrf.gov/)
- PANTHER (Protein ANalysis THrough Evolutionary Relationships) system version 7 (http://www.pantherdb.org)
- STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) version 9.1 (https://string-db.org/)

See paper I for more details.

In paper III, the STRING database was used, and also, for pathway analysis, the Kyoto Encyclopedia of Genes and Genomes (KEGG) database resource (http://www.genome.jp/kegg/).

3.3.8 Metabolomics through ^{1}H NMR spectroscopy (paper II)

Two analytical techniques dominate the field of metabolomics: ^{1}H nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS). NMR is a fast method, that within a couple of minutes can measure relatively unaltered samples - the method can even be used *in vivo*. Also, the method is highly reproducible (135), possible to fully automate, and, importantly, also directly quantitative, without the need for added reference substances (136). With NMR spectroscopy around 70 metabolites can be detected and quantified in human serum (137). However, compared to MS, NMR has a low sensitivity, and is restricted to measuring the most abundant metabolites in a sample. In contrast, MS-based metabolomics have a very high sensitivity and low detection limits, but the method is destructive for the samples, not quantitative by nature, and also have interlaboratory reproducibility problems. Also, MS-based metabolomics are very time consuming, that requires a greater amount of pre-analytical sample preparation (138).

In paper III, we used ^{1}H NMR spectroscopy on patient serum for the metabolomic analysis. For general information about ^{1}H NMR please see appendix. The workflow of the ^{1}H NMR spectroscopy is in detail described in paper II. Identification of metabolites was done with a combination of the Chenomx 8.1 NMR software (Chenomx, Edmonton, Canada), and annotation from databases such as HMDB (139) and the Birmingham metabolite library (140). On selected samples, for identification of metabolites with overlapping peaks in one dimensional ^{1}H NMR; additional two dimensional ^{1}H-^{1}H and ^{1}H-^{13}C correlated NMR experiments were performed. The metabolites were annotated according to the four levels of identification suggested by the Metabolomics Standard Initiative (MSI) (141).

3.4 STATISTICAL ANALYSES

3.4.1 Univariate statistics

In paper I, the protein levels in the different samples, measured as ratios between "heavy" and "light" peptides (H/L ratios), were Log2-transformed to obtain a normal distribution and were centered to zero. A two sample Student's t-test was performed to determine significant differences in protein ratios between the groups, using Perseus module (version 1.2.0.17) available

in the MaxQuant environment. A p-value < 0.05 was considered to be statistically significant.

In paper II, differences in patient proportions were tested with Pearson's Chi-Square and differences in age distribution were tested with Mann–Whitney U-test. Mann-Whitney was also used for comparison between metabolite levels, after Benjamini-Hochberg (BH) correction (142) for multiple testing. All univariate calculations were performed in MATLAB (MathWorks, Natick, MA).

In paper III, Pearson's chi-squared test and Mann–Whitney U-test were used to compare the different clinical characteristics and immunohistochemical biomarkers between the two patient groups. A two-tailed Welch's t-test was used to compare log-transformed average peptide expression (with/without the Benjamini-Hochberg (BH) procedure), and Pearson's chi-squared test to analyze the proportions of expressed peptides. Statistical analyses were performed with SPSS, version 22 (IBM Corp., Armonk, NY, USA) or R, version 3.3.2 (R Foundation, Vienna, Austria) software.

3.4.2 Multivariate statistics

In paper I and II, Principal Components Analysis (PCA) (143), an un-supervised multivariate method, was used for data overview, detection of trends, and outliers. PCA allows a simple visualization by reducing data dimensionality and by separating information from random variation. In a PCA-plot, the different patients are plotted according to the values they get in the 1st PC (principal component), which is the PC that captures the most variation in the data, the x-axis, and the 2nd PC, that captures the second most variation in the data, the y-axis. Patients with similar proteomic- or metabolomic patterns will cluster together in the graph.

In paper I, a supervised partial least-squares regression analysis (PLS-DA) (144) was performed. In paper II, instead an orthogonal projection to latent structures, (OPLS-DA) (145, 146), also a supervised multivariate regression method, was used. The methods analyze and compare pre-defined groups, (hence "supervised"), and allow the understanding of which variables (i.e. proteins or metabolites) that are most correlated to the differentiation between the groups, making it possible to make predictions for new samples. In paper II, discriminant metabolites were considered significant if their OPLS-DA p loadings absolute magnitude was larger than the correspondent confidence interval, while also being significantly different in a Mann–Whitney test (at 95% confidence, after Benjamini–Hochberg (BH) correction for multiple testing). Validation of the model was performed by evaluating the cross-

validated scores, permutation test, and crossvalidation-ANOVA (CV-ANOVA). All multivariate analyses were done in in SIMCA (UMETRICS AB, Umeå, Sweden).

4 RESULTS

4.1 PROTEOMIC ANALYSIS (PAPER I)

4.1.1 Identification of differentially expressed tumor proteins using LC-MS/MS.

Proteins were extracted from freshly frozen pretreatment tumor samples from five patients belonging to either of the two subgroups of DLBCL patients, (i) REF/REL; and (ii) CURED.

In total, 3,588 unique protein groups were identified at 1% FDR (false detection rate – determined through an algorithm in the peptide identification software), among which B-cell lineage specific markers (e.g., CD20, CD22, CD40, and CD79a) were present as well as proteins involved in B-cell receptor mediated signaling (e.g., mitogen-activated protein kinase 3 (MAPK3), spleen tyrosine kinase (SYK), Bruton's tyrosine kinase (BTK), and protein kinase C (PKC)).

We successfully quantified 3,027 (84%) of the identified proteins in at least one of the samples. Identification and quantification in all samples were obtained for 1305 proteins and 87 of these proteins were significantly (Student's t-test, $P < 0.05$) differentially expressed between the two patient groups. 66 proteins were overexpressed in the CURED group of patients; 21 proteins were instead overexpressed in the REF/REL group. The most functionally relevant differentially expressed proteins are described in table 1.

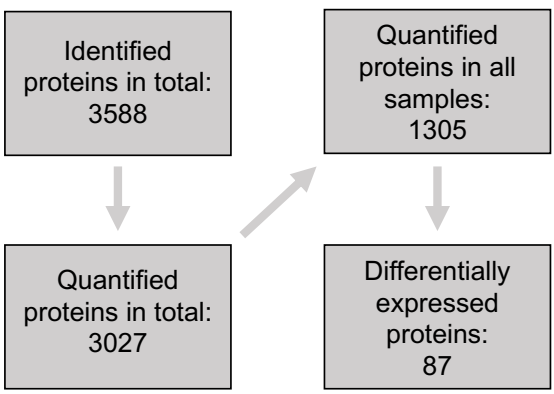

Figure 12. Numbers of identified and quantified proteins in paper I.

Table 4. The most functionally relevant differentially expressed proteins from paper I

Protein name	ID	Fold change (cur/rel)	*P* value	Function
WAFL	Q5T1M5	6.50	0.021	Regulation of actin and microtubule dynamics
STAT1	P42224	5.87	0.034	Inducer of apoptosis; negative regulation of NF-kappaB signalling
Tetraspanin CD82	P27701	5.15	0.014	Attenuation of plasma membrane-dependent actin organization
EHD4	Q9H223	4.02	0.037	Regulation of endocytic transport
Integrin beta-2; CD18	P05107	3.31	0.048	Transmembrane cell adhesion molecule
Drebrin-like protein	Q9UJU6	3.19	0.019	Actin-binding
ARP2/3 subunit 16	O15511	2.94	0.006	Regulation of actin cytoskeleton
Rac-associated protein-1	Q7L576	2.81	0.041	Regulation of actin cytoskeleton
Saposin C	P07602	2.50	0.020	Antiapoptotic effect via PI3K pathway
SHIP-1	Q92835	2.44	0.038	Involved in B cell receptor signaling pathway
CD11c; integrin alpha	P20702	2.31	0.005	Transmembrane cell adhesion molecule
ARP2/3 subunit 18	O15143	2.15	0.030	Regulation of actin cytoskeleton
Annexin A6	A6NN80	2.08	0.029	Stabilizing cortical actin cytoskeleton
Kindlin-3	Q86UX7	2.03	0.043	Activation and binding partner of integrins
Protein flightless-1 homolog	Q13045	2.01	0.007	Actin binding
CAP1	Q01518	2.00	0.025	Regulation of actin cytoskeleton
MAPK1	P28482	1.92	0.032	Involved in B-cell receptor signaling pathway
ARP2	P61160	1.90	0.035	Regulation of actin cytoskeleton
Syndapin-2	Q9UNF0	1.87	0.014	Linkage of membrane trafficking with the cytoskeleton
Moesin	P26038	1.82	0.027	Actin-binding; stabilizing microtubules at cell cortex
Proteasome MECI-1	P40306	1.77	0.041	Involved in activation of NF-kappaβ in B cells
JNK/SAPK-inhibitory kinase	Q9H2K8	1.49	0.046	Involved in B-cell receptor signaling pathway via MAPK
Caspase 3	P42574	1.42	0.047	Induction of cell apoptosis
STAG2	Q8N3U4	1.30	0.043	Tumor suppressor
eIF-2A protein kinase	P19525	0.14	0.021	Conserving protein synthesis under environmental stress
CNOT1	A5YKK6	0.34	0.032	Counteracts ER-induced stress apoptosis
NOC3	Q8WTT2	0.35	0.003	Ribosomal; essential for cell division
SKAR	Q9BY77	0.37	0.048	Promotion of cell growth via mTOR and PI3K signaling pathway
eRF3a	P15170	0.39	0.038	Inhibition of apoptosis via survivin
PDCD4	Q53EL6	0.51	0.003	Tumor suppressor via mTOR signaling pathway
MUM-1	Q15306	0.54	0.044	Transcription factor; poor prognostic marker in DLBCL
TAF15	Q92804	0.58	0.001	DNA-binding; induces rapid cell proliferation
RCC1	P18754	0.58	0.015	Chromatin regulator; involved in C-myc transcriptional activation
IKZF1	Q13422	0.61	0.044	Transcription factor; poor prognostic marker in acute lymphoblastic leukemia
SKI protein	Q13573	0.77	0.022	Protooncoprotein

Fold change: ratio between the relative abundance of the protein in CURED patients relative to REF/REL patients. Fold change > 1 indicates overexpression of the protein in the CURED group, and fold change < 1 indicates overexpression in the REF/REL group.

4.1.2 Protein network and functional analysis

The 87 differentially expressed proteins were subjected to functional characterization using the bioinformatics software DAVID, PANTHER and STRING. With the DAVID database system, 5 functional annotation clusters was specified. The cluster with the highest enrichment score consisted of 11 proteins involved in regulation of the actin cytoskeleton. Other clusters were formed by proteins involved in mitochondrial or transmembrane protein networks, antigen processing, and membrane and intracellular transport.

With PANTHER, the differentially expressed proteins were classified according to *molecular function* (the function of the protein by itself or with directly interacting proteins at a biochemical level) and *biological process* (the function of the protein in the context of a larger network of proteins that interact in a process at the level of the cell or organism). For molecular function, the main areas were binding (33.3%), catalytic activity (31.4%), and transcription regulator activity (7.8%), and for biological process, proteins involving metabolic process (24.8%), cellular process (15.2%), and transport (10.3%) were the three main groups.

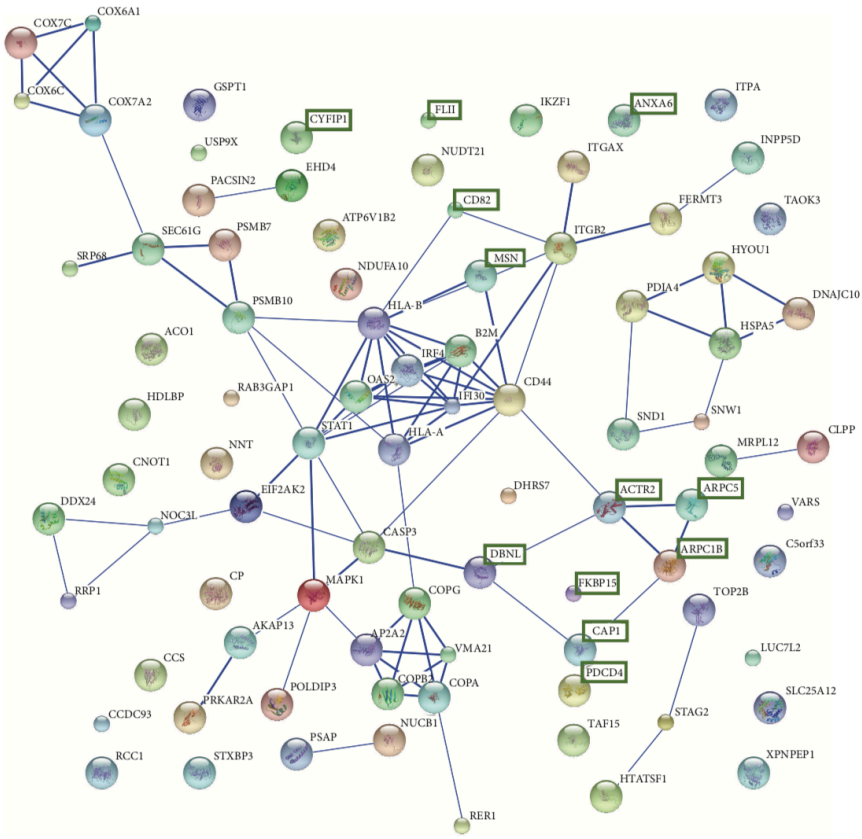

Figure 13. Visualization of network interactions between the 87 differentially expressed proteins by using the STRING database. *Stronger associations are represented by thicker lines. The total number of interactions between the proteins were highly enriched ($P < 0.00001$), as were interactions in the regulation of the actin cytoskeleton network ($P = 0.0043$). Proteins involved in the actin network and actin modulation, which were all overexpressed in the progression-free group, are highlighted with green boxes. From (127), an Open Access publication from Hindawi.*

With STRING, protein-protein interactions and protein networks could be visualized graphically *(Figure 13)*. Three tightly connected protein clusters could be suggested: a) HLA-A/ HLA-B/ B2M/ IRF4/ IFI30/ CD44, b)COPA/ COPB2/ COPG/ AP2A2 and c)ACTR2/ ARPC1B/ ARPC5/ CAP1/ DNBL. The total number of interactions between the proteins was highly enriched ($p < 0.00001$), as were interactions in the regulation of the actin cytoskeleton network ($P = 0.0043$).

4.1.3 Multivariate data analysis

An unsupervised principal component analysis (PCA) was performed in order to evaluate the quality of the data and detect possible outliers. For this we used the Log2 transformed H/L ratios for the 1305 proteins for which quantitative values from all 10 patient samples were obtained. The PCA found that all samples were within the 95 % confidence interval of the model, with no outliers. Among the proteins overexpressed in the CURED group we found a high proportion of proteins associated with the actin cytoskeleton, as analyses in the DAVID and STRING databases could confirm. We used Log2-transformed H/L ratios for five differently expressed proteins involved in regulation of actin cytoskeleton dynamics; moesin, CAP1, actin regulatory protein-G (CAP-G), annexin A6, and programmed cell death protein 4 (PDCD4), as input variables in a supervised partial least-squares regression analysis (PLS-DA). The group variable was CURED patients versus REF/REL patients. The PLS-DA model *(Figure 14)* separated the two groups, indicating that in our material, the expression of these actin-related proteins could be used as a discriminator between patients with diametrically different clinical outcome.

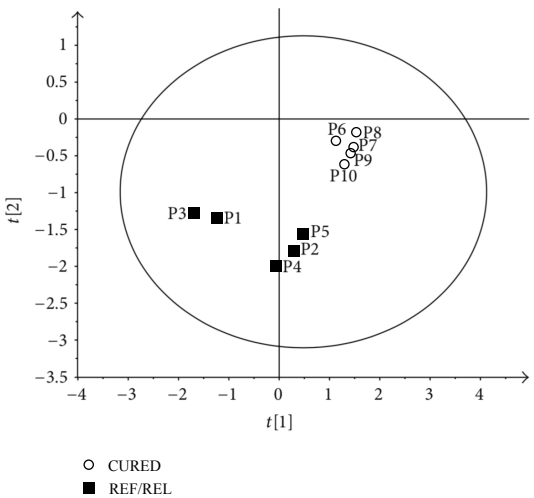

Figure 14. A supervised partial least-squares regression analysis (PLS-DA) including five selected differentially expressed proteins*, involved in regulation of actin cytoskeleton dynamics (moesin, CAP1, CAPG, annexin A6 and PDCD4) discriminates the two patient groups. Filled squares are patients 1–5 (REF/REL patients) while open circles are patients 6–10 (CURED patients). From (127), an Open Access publication from Hindawi.*

4.1.4 Validation by Western blotting

On three of the proteins associated with the actin cytoskeleton (moesin, annexin A6 and CAP1), a Western blotting was performed to validate the results from the LC-MS/MS proteomic analysis. As depicted in *Figure 15*, the data could be confirmed, as the protein levels of the three proteins were higher among the CURED patients compared to the REF/REL patients.

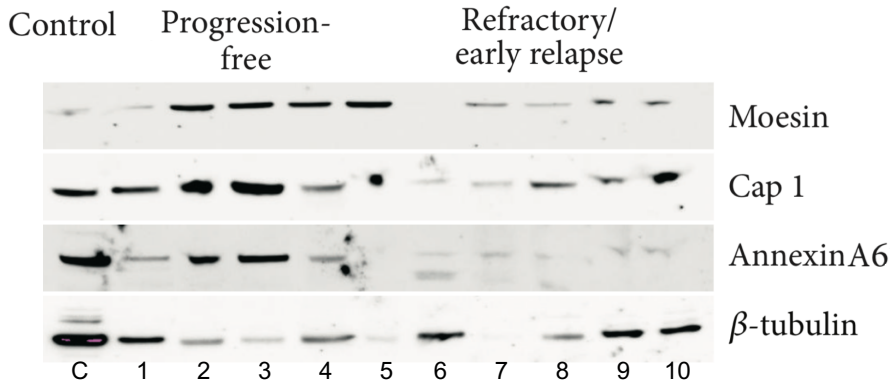

Figure 15. Western blot validation of differences found in the proteomic analysis for moesin, annexin A6, and CAP1*. "Progression- free" represents the CURED patient group (lane 1-5), and "refractory/early relapse" represents the REF/REL patients (lane 6-10). The SILAC-reference mix was used as a control (lane C) and normalization was performed by loading of equal amounts of protein into each lane of the gel. β-tubulin = loading control. From (127), an Open Access publication from Hindawi.*

4.2 METABOLOMIC ANALYSIS (PAPER II)

4.2.1 ^{1}H NMR spectroscopy of serum samples

Serum samples from 87 DLBCL patients, 27 from the REF/REL group and 60 from the CURED group, were analyzed in ^{1}H NMR. 205 peaks were obtained in the initial analysis (as explained in the appendix, each peak represents an NMR signal from single or groups of protons, that depending on their chemical surroundings have different chemical shifts, which places their peaks at different positions along the x-axis). A typical ^{1}H NMR high-field spectrum, that was recorded for each of the serum samples, is illustrated in *Figure 16.*

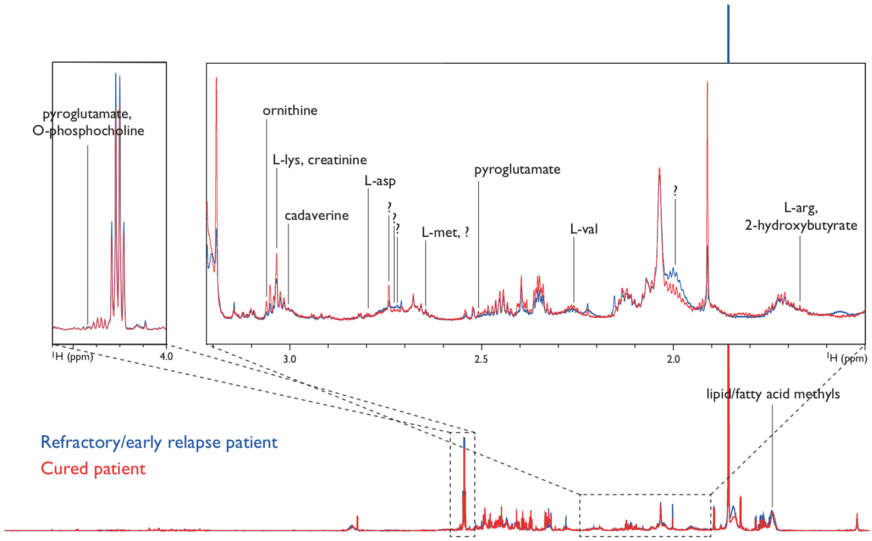

Figure 16. A merged example of two 1H NMR mean spectra (800 MHz) for DLBCL patients *(blue denotes a REF/REL patient and red denotes a CURED patient). Discriminating metabolite signals, some of which are not identified, are indicated. From (147). Reproduced with permission from Taylor & Francis; LEUKEMIA & LYMPHOMA © 2016.*

4.2.2 Multivariate data analysis

An unsupervised principal component analysis (PCA) was performed in order to evaluate the quality of the data and detect possible outliers. With this method six samples were excluded from further analysis, four from the CURED group, and two from the REF/REL group. With further PCA the number of peaks could be reduced from 205 to 92, after having removed probable multiple peaks from the same metabolite. The 92 remaining peaks were run through supervised OPLS-DA that was able to discriminate between the REF/REL and CURED patient groups, however not completely *(Figure 17a)*. The separation between the groups was validated in a permutation test and the statistical significance of the model was tested in a cross validation-ANOVA ($p = 2{,}7 \times 10^{-6}$). The aaIPI scores were superimposed into the OPLS-DA discrimination model and, even though more patients with high aaIPI scores were in the REF/REL group, no clear separation between the groups could be made according to low or high aaIPI risk score. This indicates that in this patient-material aaIPI, compared to metabolomic pattern, were of relatively lower value for the prediction of which of the patients responded unsatisfactory to immunochemotherapy *(Figure 17b)*.

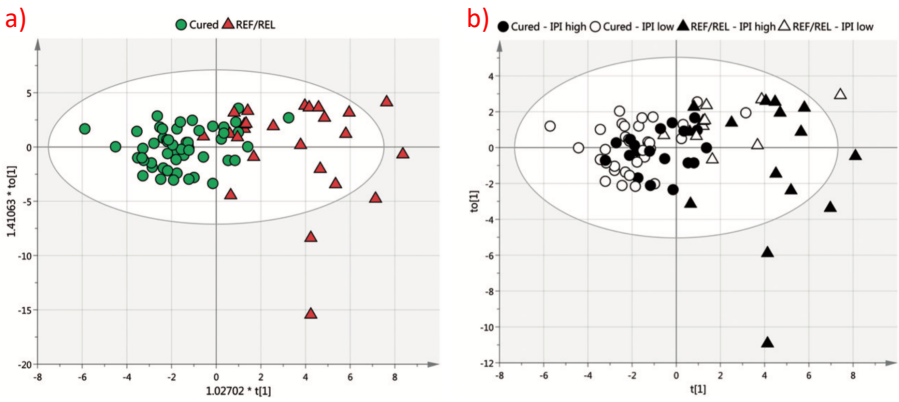

Figure 17. a) Score plot from supervised OPLS-DA *applied to discriminate between REF/REL and CURED patient groups regarding the 92 peaks in the* ^{1}H *NMR spectra. Green circles indicate CURED patients, and red triangles indicate REF/REL patients. A separation, although not complete, can be observed.*
b) The same OPLS-DA score plot, but colored according to aaIPI score*: low aaIPI score (open symbols) versus high (filled symbols); circles = CURED patients; triangles = REF/REL patients. Open and filled symbols are evenly distributed in the plot, which indicates that the OPLS-DA in this patient material is unable to discriminate between patients with regard to low or high aaIPI score. From (147). Reproduced with permission from Taylor & Francis; LEUKEMIA & LYMPHOMA © 2016.*

4.2.3 Identification of the most relevant individual metabolites

Of the 92 NMR peaks used in the OPLS-DA model, 14 peaks were found to be significantly different between the groups.

The 14 NMR peaks discriminating between REF/REL and CURED patients were subject to identification and annotation *(Table 5)*. Five of the NMR signals could not be identified from available databases, but among the rest some amino acids (lysine and arginine) and the degradation product from lysine, cadaverine, were found in significantly higher concentrations in serum in REF/REL patients. Also, a compound involved in oxidative stress conditions (2-hydroxybutyrate) and the byproduct of muscle metabolism, creatinine, were found to be higher in the REF/REL group. In contrast, the amino acids aspartate, valine and ornithine, and a metabolite in the glutathione cycle, pyroglutamate, were found to be higher in the CURED patient group.

Table 5. Discriminant metabolites

Annotation of discriminant metabolite	Fold change (REF-REL/CURED)*	p value**
Unknown	2.66	0.0030
Lipids, fatty acid methyls	2.25	0.0170
Unknown	2.09	0.0030
2-hydroxybutyrate, L-arginine	2.07	0.0030
Cadaverine	1.55	0.0414
Unknown	1.53	0.0074
L-lysine, Creatinine	1.35	0.0414
L-valine	−1.09	0.0236
Pyroglutamate, O-phosphocholine	−1.20	0.0054
Unknown	−1.28	0.0322
Pyroglutamate	−1.41	0.0030
L-aspartate	−1.57	0.0030
Unknown	−1.58	0.0236
Ornithine	−1.61	0.0362

**Fold change is positive if the concentration was higher in the REF/REL group, and negative if the concentration was higher in the CURED group. **Discriminant metabolites were considered significant if their OPLS-DA p-loadings absolute magnitude was larger than the correspondent confidence interval, while also being significantly different in a Mann–Whitney test (a p value <0.05 after BH correction for multiple testing).*

4.3 PROTEOMIC ANALYSIS (PAPER III)

4.3.1 Proteomic analysis: identification of differentially expressed proteins

In total, 2127 proteins were identified, of which 442 were found in all patient samples. Of these 442 proteins, 102 (t-test, $p < 0.05$) were differentially expressed between the two patient groups. Sixty-five proteins were overexpressed in the REF/REL group and among them were two proteins that previously have been reported as a negative prognostic marker or associated with multidrug resistance in DLBCL: Y-box protein 1 (148) and pontin/RuvBL1 protein (149). Furthermore, 46 of the 65 overexpressed proteins in the REF/REL group were ribosomal proteins (RPs). In contrast, only 2 of the total of 37 overexpressed proteins in the CURED group were RPs ($p = 7.6 \times 10^{-10}$).

Interestingly, and in accordance with the results from Paper I (127), we found that 20 of 37 overexpressed proteins in the CURED group and only 1 of 65 of the overexpressed proteins in the REF/REL group ($p = 1.4 \times 10^{-9}$), were

associated with actin organization, regulation or remodeling *(Figure 18)*. Some of the differentially expressed proteins in paper I were confirmed as overexpressed in the CURED group: actin-related protein 2 ($p = 0.0032$) and actin-related protein 2/3 complex ($p = 0.044$), while others were not (CAP1, $p = 0.27$; moesin, $p = 0.70$). All differentially expressed RPs and actin-related proteins are listed in Paper III.

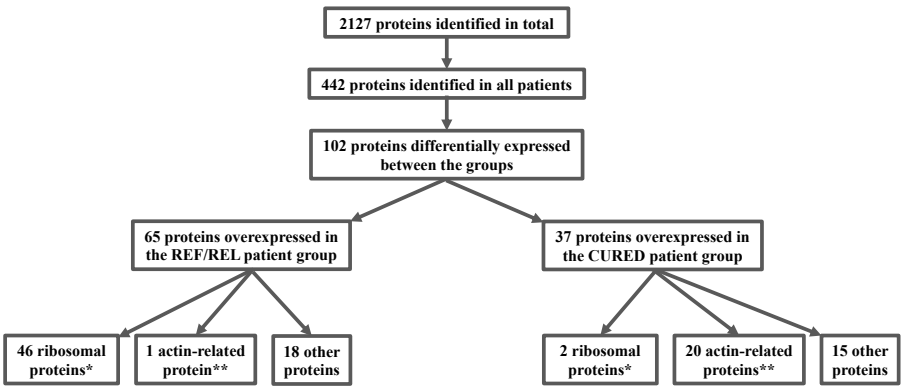

Figure 18. Numbers of identified and quantified proteins in paper III. *$*p = 7.6 \times 10^{-10}$; $**p = 1.4 \times 10^{-9}$.*

4.3.2 Ribosomal proteins and known prognostic biomarkers

Subgroup analyses were performed, in which the RP expression in REF/REL patients was analyzed relative the immunohistochemical expression of MYC and BCL2:

- 13 of 46 RPs were more highly expressed in MYC+ than in MYC- REF/REL patients.
- 22 of 46 RPs had significantly higher expression in BCL2+ than in BCL2- REF/REL-patients (where a cut-off of 70% BCL2- positive cells was used to capture the possible influence of higher BCL2 expression).
- 9 of 46 RPs were overexpressed in BCL2/MYC double expressor REF/REL patients compared to non-double expressor patients.
- Conversely, none of the RPs were more highly expressed among the MYC-, BCL2- or non-double expressor REF/REL patients compared to the positive groups.

4.3.3 Protein interaction and network analysis

When the 102 differentially expressed proteins were subjected to functional characterization using the STRING database, two tightly connected protein clusters appeared: ribosomal proteins and actin cytoskeleton regulators (*Figure 2 in Paper III*). Also when analyzing functional enrichments according to the KEGG pathway database interactions between a) ribosomal proteins and (b) proteins involved in the regulation of actin cytoskeleton, were determined to be the most significant pathways.

4.3.4 Immunohistochemical staining

Immunohistochemical staining was performed with antibodies against two ribosomal proteins (RPS5 and RPL17) and two actin-related proteins (Enah/Vasp-like protein and MARCKS-like protein). For RPS5, 39% of REF/REL patients showed strong intensity compared to 12% of CURED patients (P = 0.003) *(Fig 19A–C),* and 37% of REF/REL patients had strong intensity for RPL17 compared to 12% of CURED patients (P = 0.004) *(Fig 19D–F).* For MARCKS-like protein, 20% of CURED patients had strong intensity compared to 5% of REF/REL patients (P = 0.03) *(Fig 19G–I).* For Enah/Vasp-like protein, no differences were seen between the groups. In summary, we found statistically significant differences in immunohistochemical staining between the groups in three of the four tested proteins. However, the proportion of samples with high intensity staining was relative low among all proteins in all patient groups; the highest proportion, 39%, was seen with RPS5 among REF/REL patients.

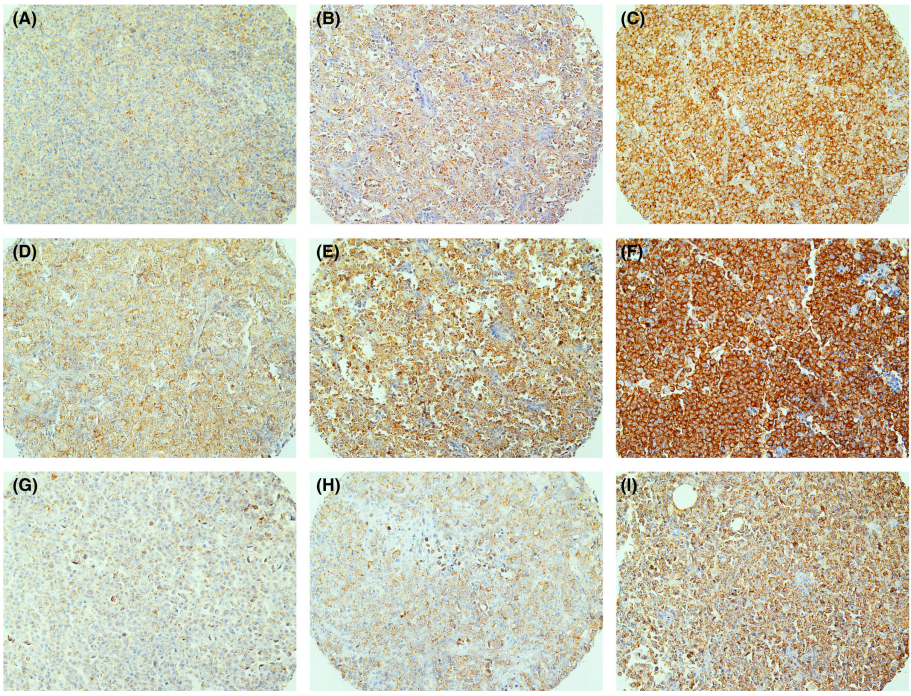

Figure 19. Immunohistochemical staining of formalin-fixed, paraffin-embedded tumor tissue (magnification 20x).
1st Row: *Ribosomal protein S5 (RPS5). (A) weak, (B) intermediate, and (C) strong intensity.*
2nd Row: *Ribosomal protein L17 (RPL17). (D) weak, (E) intermediate, and (F) strong intensity.*
3rd Row: *MARCKS-related protein. (G) weak, (H) intermediate, and (I) strong intensity.*
From (134). Reproduced with permission from John Wiley and Sons; British Journal of Haematology © 2018.

5 DISCUSSION

5.1 LC-MS/MS PROTEOMICS IN DLBCL

5.1.1 SILAC-based proteomics on freshly frozen tumor samples

A major aim with the study behind paper I was to evaluate if LC-MS/MS based proteomics on patient tumor samples could be used as a tool to reveal differences in protein expression between two diametrically different subgroups of DLBCL patients, i.e. REF/REL and CURED patients. SILAC-based quantitative proteomics on different DLBCL cell lines had earlier successfully shown differences in the protein expression between the GCB and ABC subtypes (117). The pilot study in paper I is the first that use the SILAC-based approach on actual DLBCL patient samples, and also the first that compare patients with regard to their response to immunochemotherapy instead of cell-of-origin status. In freshly frozen tumor tissue we could identify over 3500 proteins, of which more than 3000 was possible to quantify in at least one patient sample. This can be compared to the range of quantified proteins from 2103 (114) to 6223 (117) in SILAC-based studies on DLBCL cell lines. Cultured cells allow, compared to patient tissue samples, more control in the experimental situation, not least regarding the capacity to metabolically label the entire proteome. The resulting number of quantified proteins in proteomic studies on cell lines are therefore probably bound to be higher. Yet, analyses of tumor cell lines do not necessarily reflect the tumor biology seen in vivo where, for example, the tumor cells also interact with their microenvironment. In that respect, studies on tumor cell lines, even though hypothesis-generating, could perhaps have a more limited role when searching for proteins associated with treatment response, refractoriness, and clinical outcome. In our study we identified proteins with B-cell specificity, both membrane markers (CD20, CD22, CD40, CD79a) and proteins involved in B-cell receptor signaling (MAPK3, SYK, BTK and PKC). Another indicator of relevance regarding the findings, was the 72 % overlap of the proteins identified in our study with the proteins found in the study on DLBCL cell lines referred to above (117). Thus, the high number of proteins identified and quantified in our study, together with their relevance, indicate that the SILAC-based proteomic technique could be used on patient tumor tissue samples as a potential platform for finding proteins of differentiating and prognostic relevance for DLBCL.

5.1.2 TMT-based proteomics on formalin-fixed paraffin embedded (FFPE) tumor samples

Major limitations of the SILAC-based proteomics used in paper I were the scarce availability of freshly frozen tumor tissue in biobanks, and also the great amount of pre-analytical sample preparation making the method very time consuming. Due to rapid advances in proteomic techniques, we could for the follow-up study in paper III use a technique that didn't require metabolic labeling, and was possible to perform on small amounts of tissue from FFPE tumor samples, that were more accessible in biobanks. With TMT-based proteomics on almost 100 DLBCL patient samples, we could identify 2127 proteins in total. This is less than we found in paper I, but one must bear in mind that in paper III, very small tumor sample volumes were used. In a German proteomic study on 40 FFPE DLBCL tumor samples, almost 9000 proteins were identified, but here they combined two different techniques, both a SILAC-based and a label free method, on much larger tumor samples (118). The small amount of tumor tissue necessary for the TMT-based proteomic approach used in paper III, could still identify a sufficient number of proteins to open up possibilities for larger studies on the easily accessible FFPE DLBCL tumor tissue.

5.2 PROTEOMIC PATTERNS VERSUS SINGLE BIOMARKERS

The prognostic risk-models used in clinical practice are still based on either clinical characteristics of the patient (the IPI (7)) or on a few biomarkers (the Hans algorithm (44) and BCL2/MYC-status (55, 60)). In contrast, patterns of gene expression can assign DLBCL tumors to their cell-of-origin (38), and lately DLBCL patients have been possible to risk stratify according to the pattern of their mutational status (66). These patterns, based on multiple biomarkers, have so far not been possible to apply in routine clinical practice, but has been important as explanatory models for the pathogenesis of DLBCL, and also basis for continuous hypothesis-driven research. In paper I, we identified over 3500 proteins. However, merely identifying a large number of proteins is obviously not enough to pursue possible biomarkers, as it seems unlikely that single or few biomarkers could mirror the complexity in this disease. Instead, from a biological point of view, it is more interesting to study the patterns or networks of proteins to further understand underlying mechanisms of disease progression and drug resistance. Indeed, proteins related to the actin cytoskeleton were in our studies (paper I and III) enriched in the CURED patient group, and ribosomal proteins (RPs) enriched in the REF/REL group (paper III). However, the expression of the actin-related

proteins or RPs were variable within the groups, and the differences in protein levels between the groups were not numerically large, despite them being statistically significant. In addition, the immunohistochemical staining of both RPs and actin-related proteins in paper III didn´t show any clinically useful differences between the groups. In summary, although we found differences in the expression of both RPs and actin-related proteins in our studies, they can't as yet be used as single biomarkers in a clinical setting, but the findings can instead be seen as hypothesis-generating, more discussed in the following chapters.

5.3 THE ACTIN CYTOSKELETON AND DLBCL

In paper I, we found that among the 66 proteins overexpressed in the CURED patient group, a high proportion of proteins involved in the regulation and organization of the actin cytoskeleton, for example, annexin A6, members from ARP2/3 complex, drebrin-like protein, CAP1, moesin, and WAFL. Also, by inserting the differentially expressed proteins into the DAVID database, the most enriched annotated cluster was actin-binding proteins or proteins involved in regulation of actin cytoskeleton. Similarly, the STRING database found that interactions in the regulation of the actin cytoskeleton network were highly enriched ($p = 0.0043$). When five proteins involved in regulation of actin cytoskeleton dynamics (annexin A6, CAP1, CAP-G, moesin and PCDP4) were applied in a supervised regression analysis, it allowed a discrimination of the two patients groups. In addition, by using immunoblotting, we could confirm three of these proteins (annexin A6, CAP1 and moesin) to be overexpressed in the CURED patient group.

In paper III, where a different proteomic approach was used on a larger selection of patients, the results from paper I seemed to be confirmed as we found that several actin-related proteins again were overexpressed among the CURED patients. Furthermore, STRING/KEGG pathway analysis showed that proteins involved in actin cytoskeleton regulation were significantly enriched in this patient group.

In concurrence with our results, another proteomic study on DLBCL patients using 2-DE and MALDI-TOF/TOF-MS technique, concluded that several actin related proteins (ezrin, pleckstrin and annexin 5) were overexpressed among DLBCL patients sensitive to CHOP treatment (150).

The actin cytoskeleton is essential for many cellular processes, including cell migration, cytokinesis, vesicular trafficking, endocytosis, and morphogenesis (151), and dysregulation of the actin cytoskeleton is noted in a variety of

diseases, such as autoimmune disorders, neurodegenerative diseases, and cancer metastasis (152, 153). Also specific for B-cells, the actin network remodeling have been demonstrated to be a crucial factor for both up- and downregulation of B-cell receptor (BCR) signaling (154). As BCR signaling is implicated as a vital pathway for lymphoma development (155), it is tempting to speculate that the levels of actin-modulating proteins not only could influence BCR signaling, but by extension also the effects of immunochemotherapy.

Below, all actin-related proteins found to be overexpressed in the CURED group in either of paper I or III are presented in *italics*.

Different members of the actin cytoskeleton protein network have earlier been reported to be involved in the mechanisms behind drug resistance against oncovin, a drug included in the standard R-CHOP-regimen used for treatment in DLBCL. Low expressions of CAP-G, *HSP27* and *L-plastin* were reported in oncovin-resistant ALL (156, 157). A proteomic analysis of a mouse xenograft model of acute B-cell lymphoblastic leukemia has shown that downregulation of several actin-related proteins (among them *moesin*, CAP-G, HSP70 and ezrin), were involved in *in-vivo* oncovine resistance (158). Even though the exact mechanism by which a disrupted actin cytoskeleton can induce cellular resistance to antimicrotubule drugs remains to be determined, it seems that a normal actin cytoskeleton is required for the antimicrotubule cellular action of oncovin.

The differently expressed actin-related proteins in our studies have been investigated in other malignancies. For example, a low expression of *HSP27* was observed in doxorubicin-resistant uterine cancer cells (159). Homoharringtonine (HHT), a drug that can be used in late stages of acute myeloic leukemia, upregulated the levels of *myosin-9* in acute myeloid leukemia cells, which also was necessary for the apoptotic effects of HHT (160). *Profilin-1* potentiated several chemotherapeutic agents (including doxorubicin) to mediate cell death via the suppression of NF-κB and upregulation of p53 (161).

An *in-vitro* study of the actions of the monoclonal CD20-antibody rituximab, showed that rituximab polarizes B-cells, causing CD20 to cluster at the B-cell surface. This polarization of CD20 augments the therapeutic function of rituximab in NK-cell-mediated antibody-dependent cellular cytotoxicity (ADCC) (162). The study also, unexpectedly, showed that the capping of CD20 required an intact microtubule network, and that the actin-related proteins *moesin* and ICAM-1 also selectively were recruited to the CD20 cap. Thus, a higher expression of proteins involved in the modulation of the actin

cytoskeleton network in DLBCL cells could imply functional importance for an efficient rituximab-mediated cell killing by ADCC.

RhoA, a member of the Rho GTPase family, was upregulated in the CURED patient group in paper III. *RhoA* has an important role in the regulation of actin polymerization/nucleation, actin stabilization and the formation of stress fibres (163). This effect is mediated by a number of downstream partners, such as *profilin*, *cofilin*, and *the ARP2/3 complex* (all found overexpressed in the CURED group) and Rho kinase 1 (ROCK1). Inactivating mutations in the *RhoA* pathway promoted B-cell lymphoma development and mutations of *RhoA* itself or the G-protein coupled receptor Gα13 were reported in Burkitt lymphoma (164) and GCB-type DLBCL (165), supporting a tumor suppressive role for the *RhoA* pathway.

The *RhoA* pathway is also involved in the cytotoxic effects of the drug doxorubicin, which affects the actin cytoskeleton by inhibition of actin polymerization (166). This action seems to be ROCK1-dependent since depletion of ROCK1 was associated with resistance to doxorubicin (167). Thus, activation of the *RhoA* pathway might be a potential mechanism for a sustained response to doxorubicin-containing immunochemotherapy in DLBCL.

The immunomodulating drug (IMiD) lenalidomide has well-known anti-tumoral effects in ABC DLBCL cells associated with the downregulation of IRF4 (MUM1) and B-cell receptor-dependent NF-κB activity (74). Another important mechanism of its action appears to be the upregulation of the actin cytoskeletal protein network including Rho-GTPases (168). Supporting this, it has been shown that pomalidomide (another drug in the IMiD family), via the activation of *RhoA*, enhanced F-actin formation, stabilized microtubules and increased cell migration, actions which were blocked by the inhibition of ROCK1 (169).

Taken together, the results from paper I and III imply that the expression of proteins involved in the modulation of the actin cytoskeleton network in DLBCL cells could be of functional importance both for an adequate response to the CHOP regimen but also for an efficient rituximab-mediated target cell killing by ADCC. This could possibly explain a better clinical outcome for DLBCL patients with high expression of actin-related proteins when treated with immunochemotherapy. Conversely, the absence of such a pattern could instead indicate resistance or refractoriness to the R-CHOP regimen, as seen in the REF/REL patients.

5.4 RIBOSOMAL PROTEINS AND DLBCL

In paper III, we found that a number of ribosomal proteins (RPs) were overexpressed, both on an individual level and enriched in a network analysis, in DLBCL patients belonging to the REF/REL group, compared to the CURED patients. Apart from a small study showing the overexpression of RPS6 in six DLBCL lymph nodes compared with reactive nodes (170), and a recent study on RP mutations (171), there is no previous available data on RP dysregulation in DLBCL.

Being a complex molecular machine, the ribosome is responsible for protein synthesis from mRNA and consists of four ribosomal RNA species and 80 RPs that form two subunits, the small (S) that read the RNA, and the large (L) that join the amino acids in the polypeptide chain making up the protein. RPs are hence termed RPS or RPL, depending on which subunit they belong to.

Even though RPs have a primary role in ribosome biogenesis and protein translation, most RPs also have extra-ribosomal functions such as cell proliferation and regulation of apoptosis, tumorigenesis, tumor progression and metastasis (172). As most cancers have an elevated protein synthesis rate, RP overexpression in the REF/REL group could be interpreted as a just an indication of increased proliferation. However, we found no differences in proliferation index (by the standard Ki67+ cells) between our patient groups. Another possible explanation for RP overexpression in the REF-REL group could be that it correlated with MYC expression. Indeed, MYC is a general inducer of protein synthesis by activating the transcription of RNA polymerase (Pol) I, II and III (173), and specifically induces the expression of several genes encoding RPs, including RPS3, RPL6, RPL23, RPL35 and RPL44 (174). However, we found that a minority (28%) of RPs were more highly expressed in MYC-positive REF/REL patients compared with MYC-negative REF/REL patients. Together, this indicates that RP overexpression could be a more general finding in DLBCL patients resistant to R-CHOP rather than attributed to proliferation rate or MYC expression.

Although many RPs are upregulated in human cancers, several of the overexpressed RPs in our study are also associated with chemotherapy resistance in other cancer forms. RPL4 and RPL5 are overexpressed in a doxorubicin-resistant human colon cancer cell line (175), and RPS13, RPL6 and RPL23 are associated with multidrug resistance (doxorubicin and vincristine) by suppressing drug-induced apoptosis (176, 177). In addition, doxorubicin-induced apoptosis in different hepatocellular carcinoma cell lines was inhibited by the overexpression of the ribosomal protein receptor for activated C kinase (RACK1), a component of the small ribosomal subunit,

and in contrast, the silencing of the same protein instead increased the cytotoxic effects of doxorubicin (178). Furthermore, RPL34 is overexpressed in pancreatic cancer patients and the knock-down of RPL34 in pancreatic cancer cells sensitized the tumor cells to gemcitabine and 5-flourouracil (179). Interestingly, even if ribosome biogenesis, i.e. formation of ribosomes, occurs in both proliferating and resting normal cells as well as cancer cells, there is evidence that the inhibition of ribosome biogenesis may cause selective damage to malignant cells (180). RNA polymerase I is the initiator of ribosome formation. The inhibition of RNA polymerase I with a novel small molecule drug, CX-5461, could in a promising study, via p53 induced apoptosis, selectively kill B-cell lymphoma cells *in-vivo* while maintaining a viable wild-type B-cell population (181). Furthermore, the combined inhibition of ribosome biogenesis and function, using CX-5461 and the mTOR inhibitor everolimus, enhanced the induction of apoptosis and improved the survival of mice transplanted with MYC B-cell lymphoma (182). Most recently, the selective targeting of ribosomal biogenesis with CX-5461 in myeloma cell lines showed that the inhibition overcame both adhesion-mediated drug resistance and resistance to conventional and novel agents (183).

Most recent data from a study of mutations in RPs in DLBCL patients, show that patients with non-mutated p53, have a statistically higher death rate if they have RP-mutations compared to those without RP-mutations. These data suggest that mutations of RPs could provide lymphoma cells with an alternative mechanism to inactivate p53-mediated apoptosis (171).

Taken together, the data from paper III indicate that RP overexpression contributes to immunochemotherapy resistance in DLBCL patients.

5.5 SERUM METABOLOMICS AND DLBCL

In paper II, the serum metabolomic signature at diagnosis (before start of treatment), analyzed by ^{1}H NMR spectroscopy, seemed to differ between DLBCL patients from the REF/REL and CURED groups. In addition, several individual metabolites discriminated between the two patient groups.

The reason why some patients fail to respond to standard immunochemotherapy is most certainly very complex, where both tumor biology and microenvironment play their role, but also global host factors could be of importance. Indeed, despite the advances in sub-classifying DLBCL into different prognostic subtypes (GCB and ABC), the IPI remains the major prognostic tool used in clinical practice to risk-stratify DLBCL

patients. One reason for this could be that when compared to gene expression profiling or algorithms for immunohistochemical staining, techniques that only reflect tumor biology, the IPI assesses both tumor- and patient-related factors. Likewise, one appeal of metabolomics is the simultaneous assessment of tumor and host, which offers a global image of the balance between the tumor metabolism and the physiological condition of the patient.

Paper II is the first, and so far the only, study of serum metabolomics in DLBCL. The results show that the two patient groups can be separated in multivariate statistical analyses according to the pattern of their metabolome, which indicates that serum might provide information on general metabolic changes that appears to be correlated to clinical outcome in DLBCL.

When studying individual metabolites that could be identified and that differed between the groups, the concentrations of lysine, arginine, cadaverine, 2-hydroxybutyrat, and creatinine were significantly higher among the REF/REL patients, whereas aspartate, valine, and ornithine were instead higher among the cured patients. There were also metabolites that were not possible to identify, but still differed significantly between the groups. Previous studies, using NMR- or MS-based metabolomics, have described the above-mentioned metabolites in different cancer forms, although there is no clear consistency in the direction of the results.

In this study, we have not compared patients with healthy controls, but instead compared two DLBCL populations with totally different clinical outcome, in order to find possible biological diversities. Yet, even though we found several metabolites to be significantly differentially expressed between the two patient groups, their expression was highly variable, which probably would make them less useful as single biomarkers. In concordance with the discussion about the proteomic analyses in paper I and III, in a hypothetical future clinical application of serum metabolomics, the detection and quantitation of single or few metabolites will probably have inferior significance compared to the global pattern of metabolites, or the metabolomic fingerprint. NMR-based metabolomics is a fast method that is highly reproducible, possible to fully automate, and, importantly, also directly quantitative, all being qualities that give the technique a great future potential of being incorporated into routine clinical practice. Indeed, some studies have already shown that clinical decision-making could benefit from metabolic phenotyping of patients during both surgery and critical care (184, 185). The results from paper II suggest that in DLBCL-patients, a certain metabolomic pattern could reflect a more aggressive or resistant disease, possibly combined with different host factors important for the response to treatment. Further studies on larger, unselected DLBCL patient cohorts are necessary to create a

metabolomic reference material. ^{1}H NMR-based metabolomics could then have possible potential to be used as a tool to refine the present prognostic models in DLBCL.

One weakness with the study can be attributed to the ^{1}H NMR spectroscopy method in itself, because of its sensitivity to variations in both pre- and post-sampling conditions. The results are affected by the dietary status of the patient, the sampling time point and the time span from patient sampling to centrifugation and freezing of the serum. None of these factors could be controlled in the study, since it was made retrospectively. On the other hand, the fact that the results turned out to be significant despite the non-standardized sampling procedure is actually a strength of the study. Still, in future prospective studies of serum metabolomics in DLBCL, the serum sampling should be performed in a standardized fashion.

6 CONCLUSIONS

6.1.1 Paper I

In conclusion, with SILAC-based quantitative proteomic analysis a large number of proteins could be identified and quantified in freshly frozen tumor tissue from patients with diffuse large B-cell lymphoma. This could broaden the possibilities of using a proteomic platform for finding proteins of differentiating and, hopefully, prognostic significance in this disease. The data also indicate that a higher expression of proteins involved in the regulation of the actin cytoskeleton could be of functional importance for a sustained response to immunochemotherapy

6.1.2 Paper II

In conclusion, the serum metabolomic profile at diagnosis, analyzed by ^{1}H NMR spectroscopy, seemed to differ between DLBCL patients with refractory disease/early relapse and long-term progression-free patients. Certainly, there are several issues regarding NMR-based metabolomics that has to be solved before it can be implemented in a broad clinical setting. Yet, the technique is evolving and even though the results need to be validated in a larger prospective cohort, the method appears promising and might become a supportive tool for more precise prediction of response to immunochemotherapy.

6.1.3 Paper III

In conclusion, the TMT-based proteomic approach found a sufficient number of proteins which could open up possibilities for this technique to be used in larger studies on more easily accessible formalin fixed, paraffin-embedded DLBCL tumor samples. The finding from paper I of overexpression of actin-related proteins among long-term progression-free DLBCL patients appeared to be confirmed in this study. In addition, a novel discovery regarding overexpression of multiple ribosomal proteins in DLBCL patients with early relapse or refractory disease was made. The findings suggest previously undescribed mechanisms for immunochemotherapy resistance in DLBCL patients.

7 FUTURE PERSPECTIVES

Based on the results from the two proteomic studies in paper I and III, a larger study on 220 unselected DLBCL patients treated with R-CHOP is under way, in which again TMT-based explorative proteomic analysis on FFPE tumor samples will be used. Compared to the study in paper III however, a considerably larger volume of tumor tissue will be used, to enable more proteins to be identified, but also to include the important tumor microenvironment in the analysis.

A study using Parallel Reaction Monitoring (PRM) is being discussed, in which a limited number, 12-15 mainly actin-related and ribosomal proteins, are analyzed with a targeted proteomic method. The analyses are made on FFPE tumor tissue from the same large, unselected DLBCL patient cohort as above, with the aim to be able to use the proteins as predictive biomarkers. The PRM technique is, compared to explorative proteomics, easier to reproduce, and have future potential to be used in a more standardized way even in a clinical setting.

A study of RNA-sequencing of FFPE tumor tissue from the same patient cohort used in paper III have shown preliminary results, where differential expression of genes connected to the actin cytoskeleton and ribosome biogenesis was seen between CURED and REF/REL-patients. The plan is to expand this study to include a larger, unselected patient group, and to integrate the RNA-sequencing results with the results from the large proteomic study above, thus obtaining a proteogenomic approach to DLBCL prognostics and investigations on treatment resistance and possible future therapeutic targets.

Functional studies are planned, in which both older immunochemotherapy drugs like R-CHOP and etoposide, and newer drugs like B-cell receptor inhibitors (ibrutinib), BCL2-inhibitors (venetoclax), IMiDs (lenalidomide) and inhibitors of ribosomal biogenesis (CX-5461), will be tested on CRISPR-modified cultured DLBCL cell lines, in which the expression of different actin-related and ribosomal genes/proteins have been either inhibited or amplified. These studies will investigate the direct impact of these proteins on the resistance to chemotherapy, and will have a potential to find proteins that could be targeted in future treatment studies.

A prospective study on serum metabolomics of newly diagnosed DLBCL patients is planned. In this study some of the limitations of the study in paper II are addressed, and the collection of serum samples will be organized in a highly standardized way, not only at time of diagnosis, but also during and

after treatment. In addition to the normal treatment evaluation, certain metabolic factors will be measured, such as BMI, muscle strength and loss of muscle volume.

ACKNOWLEDGEMENTS

This thesis would not have been possible to complete without the help and support from a number of people. I would especially like to express my deepest gratitude to:

My main supervisor P-O Andersson, who back in the days convinced me to start my research on DLBCL, and soon after, (unwillingly…?) also convinced me to take up my old career as a provincial hospital doctor. Which I, by the way, have never regretted! Thank you for great supervision, and also for the interesting philosophical, political and cultural discussions we often tend to end up in, when we really should concentrate on p-values.

My co-supervisor Herman Nilsson-Ehle, for your never-ending enthusiasm and involvement, and also for being a link back to the eighties, when I met you under very different circumstances.

My other co-supervisor, Ulla Rüetschi, for great support during the work with paper I.

The rest of the members of the DLBLC research team: Susanne Bram-Ednersson, Mimmie Stern, Henrik Fagman and Sverker Hasselblom. Thank you for great cooperation and great friendship! This thesis could of course not have been possible without your tremendous work. Susanne and Mimmie: I hope I can help you on your way to dissertation the same way you have helped me. Henrik: Thank you for letting me be your friend, and that you now and then rather convincingly pretend that you share some of my sometimes unnecessarily dry sense of humor.

All my former colleagues on the Department of Hematology at Sahlgrenska University Hospital. For training me to become a decent hematologist, and for continuing to be such good support when I call you from Kungälv with difficult questions. A special big thank you to Lars Möllgård, who by lending me a room at Sahlgrenska made it possible for me to sit and write in peace, something that proved to be invaluable! Thank you again, and I admit I also drank some of your coffee… Big thanks also to Alexander Thörnland, who was the rightful inhabitant of the room I borrowed. Sorry for the inconvenience!

The team at the NMR-center at Göteborgs Universitet, especially Anders Pedersen and Göran Karlsson, for great cooperation on paper II. Anders also for valuable input on the metabolomics chapter in this book.

Ann Törnell, my boss, for great friendship, for believing in me and for having supported me during the logistically rather complicated process of finishing this dissertation. And thank you for all the apples!

My colleagues at the Department of Medicine at Kungälvs Sjukhus, for continuously keeping the spirit of Kungälv alive.

My close colleagues at the Section of Hemathology in Kungälv: Khadija Abdulkharim, Lars Bohlin and Valdemar Erling. Thank you for being such great friends! And thank you for all the extra work you have done for me when I have been writing.

The Hema-team at Dagvården, Kungälv: Susanne Lysell, Anders Ekelind, Susanna Larsson, Maria Jönsson, Ragnhild Sand, Jannika Rolander, Birgitta Lindell and Susanne Ingård. You are simply the best! I can't imagine a better place to work at.

My brothers, Johan and Kalle, with families. Thank you for being who you are, and for always being supportive and helpful regardless of the situation.

My parents-in-law, Sven-Olof and Elisabeth. Thank you for your continuing support, and everlasting enthusiasm of being grandparents, and for helping our family so much all the time, not least during the writing of this book!

My mother and father, Kerstin and Jan Stenson. Thank you for everything! I think you would have been proud today. You are missed!

Finally, my beloved family. Magdalena, Alfred, Kajsa and Ellen. You are the most important parts of my life. Thank you for your patience, encouragement and persistent support during the preparation of this thesis.

REFERENCES

1. Swerdlow SH, Campo E, Pileri SA, Harris NL, Stein H, Siebert R, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. Blood. 2016;127(20):2375-90.
2. Sant M, Allemani C, Tereanu C, De Angelis R, Capocaccia R, Visser O, et al. Incidence of hematologic malignancies in Europe by morphologic subtype: results of the HAEMACARE project. Blood. 2010;116(19):3724-34.
3. O'Brien ME, Easterbrook P, Powell J, Blackledge GR, Jones L, MacLennan IC, et al. The natural history of low grade non-Hodgkin's lymphoma and the impact of a no initial treatment policy on survival. Q J Med. 1991;80(292):651-60.
4. Report 2016. Swedish Lymphoma Group; 2016.
5. Hitz F, Connors JM, Gascoyne RD, Hoskins P, Moccia A, Savage KJ, et al. Outcome of patients with primary refractory diffuse large B cell lymphoma after R-CHOP treatment. Ann Hematol. 2015;94(11):1839-43.
6. Crump M, Neelapu SS, Farooq U, Van Den Neste E, Kuruvilla J, Westin J, et al. Outcomes in refractory diffuse large B-cell lymphoma: results from the international SCHOLAR-1 study. Blood. 2017;130(16):1800-8.
7. A predictive model for aggressive non-Hodgkin's lymphoma. The International Non-Hodgkin's Lymphoma Prognostic Factors Project. The New England journal of medicine. 1993;329(14):987-94.
8. Nussenzweig A, Nussenzweig MC. Origin of chromosomal translocations in lymphoid cancer. Cell. 2010;141(1):27-38.
9. Shaffer AL, 3rd, Young RM, Staudt LM. Pathogenesis of human B cell lymphomas. Annu Rev Immunol. 2012;30:565-610.
10. Eibel H, Kraus H, Sic H, Kienzler AK, Rizzi M. B cell biology: an overview. Curr Allergy Asthma Rep. 2014;14(5):434.
11. Kuppers R, Klein U, Hansmann ML, Rajewsky K. Cellular origin of human B-cell lymphomas. N Engl J Med. 1999;341(20):1520-9.
12. LeBien TW, Tedder TF. B lymphocytes: how they develop and function. Blood. 2008;112(5):1570-80.
13. Pieper K, Grimbacher B, Eibel H. B-cell biology and development. J Allergy Clin Immunol. 2013;131(4):959-71.
14. Kuppers R. Mechanisms of B-cell lymphoma pathogenesis. Nat Rev Cancer. 2005;5(4):251-62.
15. Yuseff MI, Pierobon P, Reversat A, Lennon-Dumenil AM. How B cells capture, process and present antigens: a crucial role for cell polarity. Nat Rev Immunol. 2013;13(7):475-86.
16. Basso K, Dalla-Favera R. Germinal centres and B cell lymphomagenesis. Nat Rev Immunol. 2015;15(3):172-84.

17. Kuppers R, Dalla-Favera R. Mechanisms of chromosomal translocations in B cell lymphomas. Oncogene. 2001;20(40):5580-94.
18. Calado DP, Sasaki Y, Godinho SA, Pellerin A, Kochert K, Sleckman BP, et al. The cell-cycle regulator c-Myc is essential for the formation and maintenance of germinal centers. Nat Immunol. 2012;13(11):1092-100.
19. Seifert M, Scholtysik R, Kuppers R. Origin and pathogenesis of B cell lymphomas. Methods Mol Biol. 2013;971:1-25.
20. van Leeuwen MT, Turner JJ, Joske DJ, Falster MO, Srasuebkul P, Meagher NS, et al. Lymphoid neoplasm incidence by WHO subtype in Australia 1982-2006. Int J Cancer. 2014;135(9):2146-56.
21. Szekely E, Hagberg O, Arnljots K, Jerkeman M. Improvement in survival of diffuse large B-cell lymphoma in relation to age, gender, International Prognostic Index and extranodal presentation: a population based Swedish Lymphoma Registry study. Leuk Lymphoma. 2014;55(8):1838-43.
22. Smedby KE, Baecklund E, Askling J. Malignant lymphomas in autoimmunity and inflammation: a review of risks, risk factors, and lymphoma characteristics. Cancer Epidemiol Biomarkers Prev. 2006;15(11):2069-77.
23. Goldin LR, Bjorkholm M, Kristinsson SY, Turesson I, Landgren O. Highly increased familial risks for specific lymphoma subtypes. Br J Haematol. 2009;146(1):91-4.
24. Carbone PP, Kaplan HS, Musshoff K, Smithers DW, Tubiana M. Report of the Committee on Hodgkin's Disease Staging Classification. Cancer Res. 1971;31(11):1860-1.
25. Coiffier B, Thieblemont C, Van Den Neste E, Lepeu G, Plantier I, Castaigne S, et al. Long-term outcome of patients in the LNH-98.5 trial, the first randomized study comparing rituximab-CHOP to standard CHOP chemotherapy in DLBCL patients: a study by the Groupe d'Etudes des Lymphomes de l'Adulte. Blood. 2010;116(12):2040-5.
26. Fisher RI, Gaynor ER, Dahlberg S, Oken MM, Grogan TM, Mize EM, et al. Comparison of a standard regimen (CHOP) with three intensive chemotherapy regimens for advanced non-Hodgkin's lymphoma. The New England journal of medicine. 1993;328(14):1002-6.
27. Wilson WH, Dunleavy K, Pittaluga S, Hegde U, Grant N, Steinberg SM, et al. Phase II study of dose-adjusted EPOCH and rituximab in untreated diffuse large B-cell lymphoma with analysis of germinal center and post-germinal center biomarkers. J Clin Oncol. 2008;26(16):2717-24.
28. Philip T, Guglielmi C, Hagenbeek A, Somers R, Van der Lelie H, Bron D, et al. Autologous bone marrow transplantation as compared

with salvage chemotherapy in relapses of chemotherapy-sensitive non-Hodgkin's lymphoma. N Engl J Med. 1995;333(23):1540-5.
29. Gisselbrecht C, Glass B, Mounier N, Singh Gill D, Linch DC, Trneny M, et al. Salvage regimens with autologous transplantation for relapsed large B-cell lymphoma in the rituximab era. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2010;28(27):4184-90.
30. Crump M, Kuruvilla J, Couban S, MacDonald DA, Kukreti V, Kouroukis CT, et al. Randomized comparison of gemcitabine, dexamethasone, and cisplatin versus dexamethasone, cytarabine, and cisplatin chemotherapy before autologous stem-cell transplantation for relapsed and refractory aggressive lymphomas: NCIC-CTG LY.12. J Clin Oncol. 2014;32(31):3490-6.
31. van Kampen RJW, Canals C, Schouten HC, Nagler A, Thomson KJ, Vernant J-P, et al. Allogeneic Stem-Cell Transplantation As Salvage Therapy for Patients With Diffuse Large B-Cell Non-Hodgkin's Lymphoma Relapsing After an Autologous Stem-Cell Transplantation: An Analysis of the European Group for Blood and Marrow Transplantation Registry. J Clin Oncol. 2011;29(10):1342-8.
32. Gisselbrecht C, Van Den Neste E. How I manage patients with relapsed/refractory diffuse large B cell lymphoma. Br J Haematol. 2018;182(5):633-43.
33. Neelapu SS, Locke FL, Bartlett NL, Lekakis LJ, Miklos DB, Jacobson CA, et al. Axicabtagene Ciloleucel CAR T-Cell Therapy in Refractory Large B-Cell Lymphoma. N Engl J Med. 2017;377(26):2531-44.
34. Glass B, Dohm AJ, Truemper LH, Pfreundschuh M, Bleckmann A, Wulf GG, et al. Refractory or relapsed aggressive B-cell lymphoma failing (R)-CHOP: an analysis of patients treated on the RICOVER-60 trial. Ann Oncol. 2017;28(12):3058-64.
35. Maurer MJ, Ghesquieres H, Jais JP, Witzig TE, Haioun C, Thompson CA, et al. Event-free survival at 24 months is a robust end point for disease-related outcome in diffuse large B-cell lymphoma treated with immunochemotherapy. J Clin Oncol. 2014;32(10):1066-73.
36. Ziepert M, Hasenclever D, Kuhnt E, Glass B, Schmitz N, Pfreundschuh M, et al. Standard International prognostic index remains a valid predictor of outcome for patients with aggressive CD20+ B-cell lymphoma in the rituximab era. J Clin Oncol. 2010;28(14):2373-80.
37. Ninan MJ, Wadhwa PD, Gupta P. Prognostication of diffuse large B-cell lymphoma in the rituximab era. Leuk Lymphoma. 2011;52(3):360-73.
38. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature. 2000;403(6769):503-11.

39. Lenz G, Wright G, Dave SS, Xiao W, Powell J, Zhao H, et al. Stromal gene signatures in large-B-cell lymphomas. The New England journal of medicine. 2008;359(22):2313-23.
40. Gutierrez-Garcia G, Cardesa-Salzmann T, Climent F, Gonzalez-Barca E, Mercadal S, Mate JL, et al. Gene-expression profiling and not immunophenotypic algorithms predicts prognosis in patients with diffuse large B-cell lymphoma treated with immunochemotherapy. Blood. 2011;117(18):4836-43.
41. Scott DW, Wright GW, Williams PM, Lih CJ, Walsh W, Jaffe ES, et al. Determining cell-of-origin subtypes of diffuse large B-cell lymphoma using gene expression in formalin-fixed paraffin-embedded tissue. Blood. 2014;123(8):1214-7.
42. Dybkaer K, Bogsted M, Falgreen S, Bodker JS, Kjeldsen MK, Schmitz A, et al. Diffuse large B-cell lymphoma classification system that associates normal B-cell subset phenotypes with prognosis. J Clin Oncol. 2015;33(12):1379-88.
43. Michaelsen TY, Richter J, Brondum RF, Klapper W, Johnsen HE, Albertsen M, et al. A B-cell-associated gene signature classification of diffuse large B-cell lymphoma by NanoString technology. Blood Adv. 2018;2(13):1542-6.
44. Hans CP, Weisenburger DD, Greiner TC, Gascoyne RD, Delabie J, Ott G, et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. Blood. 2004;103(1):275-82.
45. Muris JJ, Meijer CJ, Vos W, van Krieken JH, Jiwa NM, Ossenkoppele GJ, et al. Immunohistochemical profiling based on Bcl-2, CD10 and MUM1 expression improves risk stratification in patients with primary nodal diffuse large B cell lymphoma. The Journal of pathology. 2006;208(5):714-23.
46. Choi WWL, Weisenburger DD, Greiner TC, Piris MA, Banham AH, Delabie J, et al. A New Immunostain Algorithm Classifies Diffuse Large B-Cell Lymphoma into Molecular Subtypes with High Accuracy. Clin Cancer Res. 2009;15(17):5494-502.
47. Natkunam Y, Farinha P, Hsi ED, Hans CP, Tibshirani R, Sehn LH, et al. LMO2 protein expression predicts survival in patients with diffuse large B-cell lymphoma treated with anthracycline-based chemotherapy with and without rituximab. J Clin Oncol. 2008;26(3):447-54.
48. Benesova K, Forsterova K, Votavova H, Campr V, Stritesky J, Velenska Z, et al. The Hans algorithm failed to predict outcome in patients with diffuse large B-cell lymphoma treated with rituximab. Neoplasma. 2013;60(1):68-73.
49. Lossos IS, Alizadeh AA, Eisen MB, Chan WC, Brown PO, Botstein D, et al. Ongoing immunoglobulin somatic mutation in germinal center B cell-like but not in activated B cell-like diffuse large cell lymphomas. Proc Natl Acad Sci U S A. 2000;97(18):10209-13.

50. Iqbal J, Sanger WG, Horsman DE, Rosenwald A, Pickering DL, Dave B, et al. BCL2 translocation defines a unique tumor subset within the germinal center B-cell-like diffuse large B-cell lymphoma. Am J Pathol. 2004;165(1):159-66.
51. Morin RD, Johnson NA, Severson TM, Mungall AJ, An J, Goya R, et al. Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. Nat Genet. 2010;42(2):181-5.
52. Chalhoub N, Baker SJ. PTEN and the PI3-kinase pathway in cancer. Annu Rev Pathol. 2009;4:127-50.
53. Ma Y, Zhang P, Gao Y, Fan H, Zhang M, Wu J. Evaluation of AKT phosphorylation and PTEN loss and their correlation with the resistance of rituximab in DLBCL. Int J Clin Exp Pathol. 2015;8(11):14875-84.
54. Barrans S, Crouch S, Smith A, Turner K, Owen R, Patmore R, et al. Rearrangement of MYC is associated with poor prognosis in patients with diffuse large B-cell lymphoma treated in the era of rituximab. J Clin Oncol. 2010;28(20):3360-5.
55. Horn H, Ziepert M, Becher C, Barth TF, Bernd HW, Feller AC, et al. MYC status in concert with BCL2 and BCL6 expression predicts outcome in diffuse large B-cell lymphoma. Blood. 2013.
56. Davis RE, Brown KD, Siebenlist U, Staudt LM. Constitutive nuclear factor kappaB activity is required for survival of activated B cell-like diffuse large B cell lymphoma cells. The Journal of experimental medicine. 2001;194(12):1861-74.
57. Jost PJ, Ruland J. Aberrant NF-kappaB signaling in lymphoma: mechanisms, consequences, and therapeutic implications. Blood. 2007;109(7):2700-7.
58. Davis RE, Ngo VN, Lenz G, Tolar P, Young RM, Romesser PB, et al. Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. Nature. 2010;463(7277):88-92.
59. Pasqualucci L, Trifonov V, Fabbri G, Ma J, Rossi D, Chiarenza A, et al. Analysis of the coding genome of diffuse large B-cell lymphoma. Nat Genet. 2011;43(9):830-7.
60. Hu S, Xu-Monette ZY, Tzankov A, Green T, Wu L, Balasubramanyam A, et al. MYC/BCL2 protein coexpression contributes to the inferior survival of activated B-cell subtype of diffuse large B-cell lymphoma and demonstrates high-risk gene expression signatures: a report from The International DLBCL Rituximab-CHOP Consortium Program. Blood. 2013;121(20):4021-31; quiz 250.
61. Oki Y, Noorani M, Lin P, Davis RE, Neelapu SS, Ma L, et al. Double hit lymphoma: the MD Anderson Cancer Center clinical experience. Br J Haematol. 2014;166(6):891-901.

62. Niitsu N, Okamoto M, Miura I, Hirano M. Clinical features and prognosis of de novo diffuse large B-cell lymphoma with t(14;18) and 8q24/c-MYC translocations. Leukemia. 2009;23(4):777-83.
63. Lin P, Medeiros LJ. High-grade B-cell lymphoma/leukemia associated with t(14;18) and 8q24/MYC rearrangement: a neoplasm of germinal center immunophenotype with poor prognosis. Haematologica. 2007;92(10):1297-301.
64. Friedberg JW. How I treat double-hit lymphoma. Blood. 2017;130(5):590.
65. Johnson NA, Slack GW, Savage KJ, Connors JM, Ben-Neriah S, Rogic S, et al. Concurrent expression of MYC and BCL2 in diffuse large B-cell lymphoma treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. J Clin Oncol. 2012;30(28):3452-9.
66. Reddy A, Zhang J, Davis NS, Moffitt AB, Love CL, Waldrop A, et al. Genetic and Functional Drivers of Diffuse Large B Cell Lymphoma. Cell. 2017;171(2):481-94 e15.
67. Schmitz R, Wright GW, Huang DW, Johnson CA, Phelan JD, Wang JQ, et al. Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma. N Engl J Med. 2018;378(15):1396-407.
68. Chapuy B, Stewart C, Dunford AJ, Kim J, Kamburov A, Redd RA, et al. Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. Nat Med. 2018;24(5):679-90.
69. Rossi D, Diop F, Spaccarotella E, Monti S, Zanni M, Rasi S, et al. Diffuse large B-cell lymphoma genotyping on the liquid biopsy. Blood. 2017;129(14):1947-57.
70. Davis RE, Brown KD, Siebenlist U, Staudt LM. Constitutive nuclear factor kappaB activity is required for survival of activated B cell-like diffuse large B cell lymphoma cells. J Exp Med. 2001;194(12):1861-74.
71. Wilson WH, Young RM, Schmitz R, Yang Y, Pittaluga S, Wright G, et al. Targeting B cell receptor signaling with ibrutinib in diffuse large B cell lymphoma. Nat Med. 2015;21(8):922-6.
72. Dunleavy K, Pittaluga S, Czuczman MS, Dave SS, Wright G, Grant N, et al. Differential efficacy of bortezomib plus chemotherapy within molecular subtypes of diffuse large B-cell lymphoma. Blood. 2009;113(24):6069-76.
73. Hernandez-Ilizaliturri FJ, Deeb G, Zinzani PL, Pileri SA, Malik F, Macon WR, et al. Higher response to lenalidomide in relapsed/refractory diffuse large B-cell lymphoma in nongerminal center B-cell-like than in germinal center B-cell-like phenotype. Cancer. 2011;117(22):5058-66.
74. Yang Y, Shaffer AL, 3rd, Emre NC, Ceribelli M, Zhang M, Wright G, et al. Exploiting synthetic lethality for the therapy of ABC diffuse large B cell lymphoma. Cancer Cell. 2012;21(6):723-37.

75. Davies AJ, Barrans S, Maishman T, Cummin TE, Bentley M, Mamot C, et al. Differential efficacy of bortezomib in subtypes of diffuse large b-cell lymphoma (DLBCL): a prospective randomised study stratified by transcriptome profiling: REMoDL-B. Hematol Oncol. 2017;35(S2):130-1.
76. Nowakowski GS, Chiappella A, Witzig TE, Spina M, Gascoyne RD, Zhang L, et al. ROBUST: Lenalidomide-R-CHOP versus placebo-R-CHOP in previously untreated ABC-type diffuse large B-cell lymphoma. Future Oncol. 2016;12(13):1553-63.
77. Snuderl M, Kolman OK, Chen YB, Hsu JJ, Ackerman AM, Dal Cin P, et al. B-cell lymphomas with concurrent IGH-BCL2 and MYC rearrangements are aggressive neoplasms with clinical and pathologic features distinct from Burkitt lymphoma and diffuse large B-cell lymphoma. Am J Surg Pathol. 2010;34(3):327-40.
78. Sehn LH, Gascoyne RD. Diffuse large B-cell lymphoma: optimizing outcome in the context of clinical and biologic heterogeneity. Blood. 2015;125(1):22-32.
79. Davids MS, Roberts AW, Seymour JF, Pagel JM, Kahl BS, Wierda WG, et al. Phase I First-in-Human Study of Venetoclax in Patients With Relapsed or Refractory Non-Hodgkin Lymphoma. J Clin Oncol. 2017;35(8):826-33.
80. Gottesman MM, Fojo T, Bates SE. Multidrug resistance in cancer: role of ATP-dependent transporters. Nat Rev Cancer. 2002;2(1):48-58.
81. Ambudkar SV, Kimchi-Sarfaty C, Sauna ZE, Gottesman MM. P-glycoprotein: from genomics to mechanism. Oncogene. 2003;22(47):7468-85.
82. Goldstein LJ, Galski H, Fojo A, Willingham M, Lai SL, Gazdar A, et al. Expression of a multidrug resistance gene in human cancers. J Natl Cancer Inst. 1989;81(2):116-24.
83. Webb M, Brun M, McNiven M, Le Couteur D, Craft P. MDR1 and MRP expression in chronic B-cell lymphoproliferative disorders. Br J Haematol. 1998;102(3):710-7.
84. Smith MR. Rituximab (monoclonal anti-CD20 antibody): mechanisms of action and resistance. Oncogene. 2003;22(47):7359-68.
85. Olejniczak SH, Hernandez-Ilizaliturri FJ, Clements JL, Czuczman MS. Acquired resistance to rituximab is associated with chemotherapy resistance resulting from decreased Bax and Bak expression. Clin Cancer Res. 2008;14(5):1550-60.
86. Kikuchi J, Wada T, Shimizu R, Izumi T, Akutsu M, Mitsunaga K, et al. Histone deacetylases are critical targets of bortezomib-induced cytotoxicity in multiple myeloma. Blood. 2010;116(3):406-17.
87. Xu WS, Parmigiani RB, Marks PA. Histone deacetylase inhibitors: molecular mechanisms of action. Oncogene. 2007;26(37):5541-52.

88. Shimizu R, Kikuchi J, Wada T, Ozawa K, Kano Y, Furukawa Y. HDAC inhibitors augment cytotoxic activity of rituximab by upregulating CD20 expression on lymphoma cells. Leukemia. 2010;24(10):1760-8.
89. Frys S, Simons Z, Hu Q, Barth MJ, Gu JJ, Mavis C, et al. Entinostat, a novel histone deacetylase inhibitor is active in B-cell lymphoma and enhances the anti-tumour activity of rituximab and chemotherapy agents. Br J Haematol. 2015;169(4):506-19.
90. Xue K, Gu JJ, Zhang Q, Mavis C, Hernandez-Ilizaliturri FJ, Czuczman MS, et al. Vorinostat, a histone deacetylase (HDAC) inhibitor, promotes cell cycle arrest and re-sensitizes rituximab- and chemo-resistant lymphoma cells to chemotherapy agents. J Cancer Res Clin Oncol. 2016;142(2):379-87.
91. Ageberg M, Rydstrom K, Relander T, Drott K. The histone deacetylase inhibitor valproic acid sensitizes diffuse large B-cell lymphoma cell lines to CHOP-induced cell death. Am J Transl Res. 2013;5(2):170-83.
92. Drott K, Hagberg H, Papworth K, Relander T, Jerkeman M. Valproate in combination with rituximab and CHOP as first-line therapy in diffuse large B-cell lymphoma (VALFRID). Blood Adv. 2018;2(12):1386-92.
93. Wilson WH, Teruya-Feldstein J, Fest T, Harris C, Steinberg SM, Jaffe ES, et al. Relationship of p53, bcl-2, and tumor proliferation to clinical drug resistance in non-Hodgkin's lymphomas. Blood. 1997;89(2):601-9.
94. Jardin F, Jais JP, Molina TJ, Parmentier F, Picquenot JM, Ruminy P, et al. Diffuse large B-cell lymphomas with CDKN2A deletion have a distinct gene expression signature and a poor prognosis under R-CHOP treatment: a GELA study. Blood. 2010;116(7):1092-104.
95. Levine AJ. p53, the cellular gatekeeper for growth and division. Cell. 1997;88(3):323-31.
96. Pentheroudakis G, Goussia A, Voulgaris E, Nikolaidis K, Ioannidou E, Papoudou-Bai A, et al. High levels of topoisomerase IIalpha protein expression in diffuse large B-cell lymphoma are associated with high proliferation, germinal center immunophenotype, and response to treatment. Leuk Lymphoma. 2010;51(7):1260-8.
97. Nobili S, Napoli C, Puccini B, Landini I, Perrone G, Brugia M, et al. Identification of pharmacogenomic markers of clinical efficacy in a dose-dense therapy regimen (R-CHOP14) in diffuse large B-cell lymphoma. Leuk Lymphoma. 2014;55(9):2071-8.
98. Gilbert LA, Hemann MT. DNA damage-mediated induction of a chemoresistant niche. Cell. 2010;143(3):355-66.
99. Miller TP, Dahlberg S, Cassady JR, Adelstein DJ, Spier CM, Grogan TM, et al. Chemotherapy alone compared with chemotherapy plus radiotherapy for localized intermediate- and high-grade non-Hodgkin's lymphoma. N Engl J Med. 1998;339(1):21-6.

100. Coiffier B, Lepage E, Briere J, Herbrecht R, Tilly H, Bouabdallah R, et al. CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. The New England journal of medicine. 2002;346(4):235-42.
101. Habermann TM, Weller EA, Morrison VA, Gascoyne RD, Cassileth PA, Cohn JB, et al. Rituximab-CHOP versus CHOP alone or with maintenance rituximab in older patients with diffuse large B-cell lymphoma. J Clin Oncol. 2006;24(19):3121-7.
102. Pfreundschuh M, Schubert J, Ziepert M, Schmits R, Mohren M, Lengfelder E, et al. Six versus eight cycles of bi-weekly CHOP-14 with or without rituximab in elderly patients with aggressive CD20+ B-cell lymphomas: a randomised controlled trial (RICOVER-60). The lancet oncology. 2008;9(2):105-16.
103. Jerkeman M, Amini RM, Andersson PO, Drott K, Ekström Smedby K, Erlanson M, et al. Aggressive B-cell lymphomas - Swedish national care program. Home page of regional cancer centres (RCC): National Board of Health and Welfare.; 2018 2018-02-27.
104. Staiger AM, Ziepert M, Horn H, Scott DW, Barth TFE, Bernd HW, et al. Clinical Impact of the Cell-of-Origin Classification and the MYC/ BCL2 Dual Expresser Status in Diffuse Large B-Cell Lymphoma Treated Within Prospective Clinical Trials of the German High-Grade Non-Hodgkin's Lymphoma Study Group. J Clin Oncol. 2017;35(22):2515-26.
105. Colomo L, Lopez-Guillermo A, Perales M, Rives S, Martinez A, Bosch F, et al. Clinical impact of the differentiation profile assessed by immunophenotyping in patients with diffuse large B-cell lymphoma. Blood. 2003;101(1):78-84.
106. Berglund M, Thunberg U, Amini RM, Book M, Roos G, Erlanson M, et al. Evaluation of immunophenotype in diffuse large B-cell lymphoma and its impact on prognosis. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc. 2005;18(8):1113-20.
107. Culpin RE, Sieniawski M, Angus B, Menon GK, Proctor SJ, Milne P, et al. Prognostic significance of immunohistochemistry-based markers and algorithms in immunochemotherapy-treated diffuse large B cell lymphoma patients. Histopathology. 2013;63(6):788-801.
108. de Jong D, Rosenwald A, Chhanabhai M, Gaulard P, Klapper W, Lee A, et al. Immunohistochemical prognostic markers in diffuse large B-cell lymphoma: validation of tissue microarray as a prerequisite for broad clinical applications--a study from the Lunenburg Lymphoma Biomarker Consortium. J Clin Oncol. 2007;25(7):805-12.
109. Moraes F, Goes A. A decade of human genome project conclusion: Scientific diffusion about our genome knowledge. Biochem Mol Biol Educ. 2016;44(3):215-23.

110. Mann M, Jensen ON. Proteomic analysis of post-translational modifications. Nat Biotechnol. 2003;21(3):255-61.
111. de Godoy LM, Olsen JV, Cox J, Nielsen ML, Hubner NC, Frohlich F, et al. Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. Nature. 2008;455(7217):1251-4.
112. Cox J, Mann M. Is proteomics the new genomics? Cell. 2007;130(3):395-8.
113. Boyd RS, Dyer MJS, Cain K. Proteomic analysis of B-cell malignancies. J Proteomics. 2010;73(10):1804-22.
114. Huang X, Shen Y, Liu M, Bi C, Jiang C, Iqbal J, et al. Quantitative proteomics reveals that miR-155 regulates the PI3K-AKT pathway in diffuse large B-cell lymphoma. Am J Pathol. 2012;181(1):26-33.
115. Li J, Okamoto H, Yin C, Jagannathan J, Takizawa J, Aoki S, et al. Proteomic characterization of primary diffuse large B-cell lymphomas in the central nervous system. J Neurosurg. 2008;109(3):536-46.
116. Romesser PB, Perlman DH, Faller DV, Costello CE, McComb ME, Denis GV. Development of a malignancy-associated proteomic signature for diffuse large B-cell lymphoma. The American journal of pathology. 2009;175(1):25-35.
117. Deeb SJ, D'Souza RC, Cox J, Schmidt-Supprian M, Mann M. Super-SILAC Allows Classification of Diffuse Large B-cell Lymphoma Subtypes by Their Protein Expression Profiles. Molecular & cellular proteomics : MCP. 2012;11(5):77-89.
118. Deeb SJ, Tyanova S, Hummel M, Schmidt-Supprian M, Cox J, Mann M. Machine Learning Based Classification of Diffuse Large B-cell Lymphoma Patients by their Protein Expression Profiles. Mol Cell Proteomics. 2015.
119. Psychogios N, Hau DD, Peng J, Guo AC, Mandal R, Bouatra S, et al. The human serum metabolome. PLoS One. 2011;6(2):e16957.
120. Amberg A, Riefke B, Schlotterbeck G, Ross A, Senn H, Dieterle F, et al. NMR and MS Methods for Metabolomics. Methods Mol Biol. 2017;1641:229-58.
121. Chen WL, Wang JH, Zhao AH, Xu X, Wang YH, Chen TL, et al. A distinct glucose metabolism signature of acute myeloid leukemia with prognostic value. Blood. 2014;124(10):1645-54.
122. Jobard E, Pontoizeau C, Blaise BJ, Bachelot T, Elena-Herrmann B, Tredan O. A serum nuclear magnetic resonance-based metabolomic signature of advanced metastatic human breast cancer. Cancer Lett. 2014;343(1):33-41.
123. Tenori L, Oakman C, Claudino WM, Bernini P, Cappadona S, Nepi S, et al. Exploration of serum metabolomic profiles and outcomes in women with metastatic breast cancer: A pilot study. Mol Oncol. 2012;6(4):437-44.

124. Urban PL. Quantitative mass spectrometry: an overview. Philos Trans A Math Phys Eng Sci. 2016;374(2079).
125. Aebersold R, Mann M. Mass spectrometry-based proteomics. Nature. 2003;422(6928):198-207.
126. Vermeulen M, Selbach M. Quantitative proteomics: a tool to assess cell differentiation. Curr Opin Cell Biol. 2009;21(6):761-6.
127. Ruetschi U, Stenson M, Hasselblom S, Nilsson-Ehle H, Hansson U, Fagman H, et al. SILAC-Based Quantitative Proteomic Analysis of Diffuse Large B-Cell Lymphoma Patients. International journal of proteomics. 2015;2015:841769.
128. Ong SE, Mann M. Stable isotope labeling by amino acids in cell culture for quantitative proteomics. Methods Mol Biol. 2007;359:37-52.
129. Wisniewski JR, Zougman A, Mann M. Combination of FASP and StageTip-based fractionation allows in-depth analysis of the hippocampal membrane proteome. J Proteome Res. 2009;8(12):5674-8.
130. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotech. 2008;26(12):1367-72.
131. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. J Proteome Res. 2011;10(4):1794-805.
132. Thompson A, Schafer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, et al. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal Chem. 2003;75(8):1895-904.
133. Werner T, Sweetman G, Savitski MF, Mathieson T, Bantscheff M, Savitski MM. Ion coalescence of neutron encoded TMT 10-plex reporter ions. Anal Chem. 2014;86(7):3594-601.
134. Bram Ednersson S, Stenson M, Stern M, Enblad G, Fagman H, Nilsson-Ehle H, et al. Expression of ribosomal and actin network proteins and immunochemotherapy resistance in diffuse large B cell lymphoma patients. Br J Haematol. 2018.
135. Keun HC, Ebbels TMD, Antti H, Bollard ME, Beckonert O, Schlotterbeck G, et al. Analytical Reproducibility in 1H NMR-Based Metabonomic Urinalysis. Chem Res Toxicol. 2002;15(11):1380-6.
136. Markley JL, Bruschweiler R, Edison AS, Eghbalnia HR, Powers R, Raftery D, et al. The future of NMR-based metabolomics. Curr Opin Biotechnol. 2017;43:34-40.
137. Nagana Gowda GA, Gowda YN, Raftery D. Expanding the limits of human blood metabolite quantitation using NMR spectroscopy. Anal Chem. 2015;87(1):706-15.
138. Riekeberg E, Powers R. New frontiers in metabolomics: from measurement to insight. F1000Res. 2017;6:1148.

139. Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, Liu Y, et al. HMDB 3.0--The Human Metabolome Database in 2013. Nucleic Acids Res. 2013;41(Database issue):D801-7.
140. Ludwig C, Easton J, Lodi A, Tiziani S, Manzoor S, Southam A, et al. Birmingham Metabolite Library: A publicly accessible database of 1-D 1H and 2-D 1H J-resolved NMR spectra of authentic metabolite standards (BML-NMR)2011. 8-18 p.
141. Salek RM, Steinbeck C, Viant MR, Goodacre R, Dunn WB. The role of reporting standards for metabolite annotation and identification in metabolomic studies. Gigascience. 2013;2(1):13.
142. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological). 1995;57(1):289-300.
143. Wold S, Esbensen K, Geladi P. Proceedings of the Multivariate Statistical Workshop for Geologists and GeochemistsPrincipal component analysis. Chemometrics and Intelligent Laboratory Systems. 1987;2(1):37-52.
144. Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems. 2001;58(2):109-30.
145. Trygg J, Wold S. Orthogonal projections to latent structures (O-PLS). Journal of Chemometrics. 2002;16(3):119-28.
146. Bylesjö M, Rantalainen M, Cloarec O, Nicholson JK, Holmes E, Trygg J. OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. Journal of Chemometrics. 2006;20(8-10):341-51.
147. Stenson M, Pedersen A, Hasselblom S, Nilsson-Ehle H, Karlsson BG, Pinto R, et al. Serum nuclear magnetic resonance-based metabolomics and outcome in diffuse large B-cell lymphoma patients - a pilot study. Leuk Lymphoma. 2016:1-9.
148. Miao X, Wu Y, Wang Y, Zhu X, Yin H, He Y, et al. Y-box-binding protein-1 (YB-1) promotes cell proliferation, adhesion and drug resistance in diffuse large B-cell lymphoma. Exp Cell Res. 2016;346(2):157-66.
149. Nishiu M, Yanagawa R, Nakatsuka S-i, Yao M, Tsunoda T, Nakamura Y, et al. Microarray Analysis of Gene-expression Profiles in Diffuse Large B-cell Lymphoma: Identification of Genes Related to Disease Progression. Jpn J Cancer Res. 2002;93(8):894-901.
150. Liu Y, Zeng L, Zhang S, Zeng S, Huang J, Tang Y, et al. Identification of differentially expressed proteins in chemotherapy-sensitive and chemotherapy-resistant diffuse large B cell lymphoma by proteomic methods. Med Oncol. 2013;30(2):528.
151. Lee SH, Dominguez R. Regulation of actin cytoskeleton dynamics in cells. Mol Cells. 2010;29(4):311-25.

152. Hall A. The cytoskeleton and cancer. Cancer Metastasis Rev. 2009;28(1-2):5-14.
153. Bamburg JR, Wiggan OP. ADF/cofilin and actin dynamics in disease. Trends Cell Biol. 2002;12(12):598-605.
154. Song W, Liu C, Seeley-Fallen MK, Miller H, Ketchum C, Upadhyaya A. Actin-mediated feedback loops in B-cell receptor signaling. Immunol Rev. 2013;256(1):177-89.
155. Niemann CU, Wiestner A. B-cell receptor signaling as a driver of lymphoma development and evolution. Semin Cancer Biol. 2013;23(6):410-21.
156. Verrills NM, Po'uha ST, Liu ML, Liaw TY, Larsen MR, Ivery MT, et al. Alterations in gamma-actin and tubulin-targeted drug resistance in childhood leukemia. J Natl Cancer Inst. 2006;98(19):1363-74.
157. Verrills NM, Walsh BJ, Cobon GS, Hains PG, Kavallaris M. Proteome analysis of vinca alkaloid response and resistance in acute lymphoblastic leukemia reveals novel cytoskeletal alterations. J Biol Chem. 2003;278(46):45082-93.
158. Verrills NM, Liem NL, Liaw TY, Hood BD, Lock RB, Kavallaris M. Proteomic analysis reveals a novel role for the actin cytoskeleton in vincristine resistant childhood leukemia--an in vivo study. Proteomics. 2006;6(5):1681-94.
159. May EW, Lin ST, Lin CC, Chang JF, Hung E, Lo YW, et al. Identification of up- and down-regulated proteins in doxorubicin-resistant uterine cancer cells: reticulocalbin-1 plays a key role in the development of doxorubicin-associated resistance. Pharmacol Res. 2014;90:1-17.
160. Zhang T, Shen S, Zhu Z, Lu S, Yin X, Zheng J, et al. Homoharringtonine binds to and increases myosin-9 in myeloid leukaemia. Br J Pharmacol. 2016;173(1):212-21.
161. Zaidi AH, Raviprakash N, Mokhamatam RB, Gupta P, Manna SK. Profilin potentiates chemotherapeutic agents mediated cell death via suppression of NF-κB and upregulation of p53. Apoptosis. 2016;21(4):502-13.
162. Rudnicka D, Oszmiana A, Finch DK, Strickland I, Schofield DJ, Lowe DC, et al. Rituximab causes a polarization of B cells that augments its therapeutic function in NK-cell-mediated antibody-dependent cellular cytotoxicity. Blood. 2013;121(23):4694-702.
163. Thumkeo D, Watanabe S, Narumiya S. Physiological roles of Rho and Rho effectors in mammals. Eur J Cell Biol. 2013;92(10–11):303-15.
164. Rohde M, Richter J, Schlesner M, Betts MJ, Claviez A, Bonn BR, et al. Recurrent RHOA mutations in pediatric Burkitt lymphoma treated according to the NHL-BFM protocols. Genes Chromosomes Cancer. 2014;53(11):911-6.

165. Muppidi JR, Schmitz R, Green JA, Xiao W, Larsen AB, Braun SE, et al. Loss of signaling via Gα13 in germinal center B cell-derived lymphoma. Nature. 2014;516(7530):254-8.
166. Colombo R, Necco A, Vailati G, Milzani A. Dose-dependence of doxorubicin effect on actin assembly in vitro. Exp Mol Pathol. 1988;49(3):297-304.
167. Wei L, Surma M, Gough G, Shi S, Lambert-Cheatham N, Chang J, et al. Dissecting the Mechanisms of Doxorubicin and Oxidative Stress-Induced Cytotoxicity: The Involvement of Actin Cytoskeleton and ROCK1. PLoS One. 2015;10(7):e0131763.
168. Riches JC, Sangaralingam A, Kiaii S, Chaplin T, Cekdemir D, Iqbal S, et al. Impact of Lenalidomide on Gene Expression Profiles of Malignant and Immune Cells in Patients with Chronic Lymphocytic Leukemia. Blood. 2011;118(21):976-.
169. Xu Y, Li J, Ferguson GD, Mercurio F, Khambatta G, Morrison L, et al. Immunomodulatory drugs reorganize cytoskeleton by modulating Rho GTPases. Blood. 2009;114(2):338-45.
170. Hagner PR, Mazan-Mamczarz K, Dai B, Balzer EM, Corl S, Martin SS, et al. Ribosomal protein S6 is highly expressed in non-Hodgkin lymphoma and associates with mRNA containing a 5' terminal oligopyrimidine tract. Oncogene. 2011;30(13):1531-41.
171. Derenzini E, Agostinelli C, Rossi A, Rossi M, Scellato F, Melle F, et al. Genomic alterations of ribosomal protein genes in diffuse large B cell lymphoma. Br J Haematol. 2018.
172. Xu X, Xiong X, Sun Y. The role of ribosomal proteins in the regulation of cell proliferation, tumorigenesis, and genomic integrity. Sci China Life Sci. 2016;59(7):656-72.
173. Gomez-Roman N, Felton-Edkins ZA, Kenneth NS, Goodfellow SJ, Athineos D, Zhang J, et al. Activation by c-Myc of transcription by RNA polymerases I, II and III. Biochemical Society Symposium. 2006;73:141-54.
174. de las Heras-Rubio A, Perucho L, Paciucci R, Vilardell J, Lleonart ME. Ribosomal proteins as novel players in tumorigenesis. Cancer Metastasis Rev. 2014;33(1):115-41.
175. Bertram J, Palfner K, Hiddemann W, Kneba M. Overexpression of ribosomal proteins L4 and L5 and the putative alternative elongation factor PTI-1 in the doxorubicin resistant human colon cancer cell line LoVoDxR. Eur J Cancer. 1998;34(5):731-6.
176. Shi Y, Zhai H, Wang X, Han Z, Liu C, Lan M, et al. Ribosomal proteins S13 and L23 promote multidrug resistance in gastric cancer cells by suppressing drug-induced apoptosis. Exp Cell Res. 2004;296(2):337-46.
177. Du J, Shi Y, Pan Y, Jin X, Liu C, Liu N, et al. Regulation of multidrug resistance by ribosomal protein l6 in gastric cancer cells. Cancer Biol Ther. 2005;4(2):242-7.

178. Ruan Y, Sun L, Hao Y, Wang L, Xu J, Zhang W, et al. Ribosomal RACK1 promotes chemoresistance and growth in human hepatocellular carcinoma. J Clin Invest. 2012;122(7):2554-66.
179. Wei F, Ding L, Wei Z, Zhang Y, Li Y, Qinghua L, et al. Ribosomal protein L34 promotes the proliferation, invasion and metastasis of pancreatic cancer cells. Oncotarget. 2016;7(51):85259-72.
180. Brighenti E, Trere D, Derenzini M. Targeted cancer therapy with ribosome biogenesis inhibitors: a real possibility? Oncotarget. 2015;6(36):38617-27.
181. Bywater MJ, Poortinga G, Sanij E, Hein N, Peck A, Cullinane C, et al. Inhibition of RNA Polymerase I as a Therapeutic Strategy to Promote Cancer-Specific Activation of p53. Cancer Cell. 2012;22(1):51-65.
182. Devlin JR, Hannan KM, Hein N, Cullinane C, Kusnadi E, Ng PY, et al. Combination Therapy Targeting Ribosome Biogenesis and mRNA Translation Synergistically Extends Survival in MYC-Driven Lymphoma. Cancer Discov. 2016;6(1):59-70.
183. Lee HC, Wang H, Baladandayuthapani V, Lin H, He J, Jones RJ, et al. RNA Polymerase I Inhibition with CX-5461 as a Novel Therapeutic Strategy to Target MYC in Multiple Myeloma. Br J Haematol. 2017;177(1):80-94.
184. Villasenor A, Kinross JM, Li JV, Penney N, Barton RH, Nicholson JK, et al. 1H NMR global metabolic phenotyping of acute pancreatitis in the emergency unit. J Proteome Res. 2014;13(12):5362-75.
185. Nicholson JK, Holmes E, Kinross JM, Darzi AW, Takats Z, Lindon JC. Metabolic phenotyping in clinical and surgical environments. Nature. 2012;491(7424):384-92.
186. Hornak J. The Basics Of NMR 1997 [Available from: http://www.cis.rit.edu/htbooks/nmr/bnmr.htm.]

APPENDIX

^{1}H NMR SPECTROSCOPY

NMR is dependent on the physical property of spin, that is possessed by protons, neutrons and electrons. The nucleus of a hydrogen atom, i.e. a proton, have spin, and can therefore be used in NMR experiments. A proton, because of its spin, behaves like a small magnet with a north and a south pole. When placed in a strong magnetic field, the proton can, due to the spin, align to the magnetic field in two different ways; either with the magnetic poles aligned (low energy state), or in the opposite direction, i.e. against the magnetic field (high energy state). A transition from the low energy spin state to the high energy spin state require input of energy in the form of electromagnetic radiation. Nuclear magnetic resonance (NMR) occurs when a radiofrequency pulse is resonantly absorbed and consequently emitted by certain atom nuclei, in this case protons, by transitions between the different energy levels. The energy that is emitted is measured as an oscillating induced current in coils surrounding the sample in the NMR spectrometer. The time-domain detected signal can be Fourier-transformed and plotted as an NMR spectrum with signals occurring according to their resonance frequency. The exact frequency of a given nucleus is dependent on the atoms surrounding the proton/hydrogen atom in the investigated molecule. This phenomenon is called chemical shift, measured in ppm to be independent of the employed magnetic field strength of the instrument, which in effect causes different molecules to give different ^{1}H NMR signals, making it possible to identify different substances in a sample (186). An example of an NMR spectrum is depicted below.

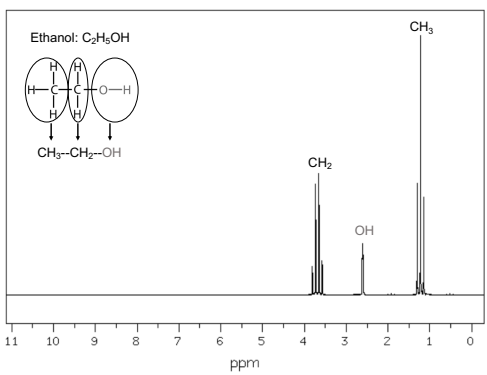

Figure 20. *^{1}H NMR spectrum of ethanol dissolved in chloroform-d. The structural formula of the ethanol molecule, with three hydrogen-containing groups (CH_3, CH_2 and OH) with different chemical shifts, give three peaks along the x-axis, on which the chemical shifts are plotted as ppm. Each hydrogen-containing group will, depending on the spin of their protons, influence the chemical shifts of the other two groups, thus forming the unique pattern of peaks upon which the substance, i.e. ethanol, can be identified. Spectrum downloaded from SDBSWeb : https://sdbs.db.aist.go.jp (National Institute of Advanced Industrial Science and Technology).*