Electronic Journal of Biotechnology 52 (2021) 1–12

Contents lists available at ScienceDirect

Electronic Journal of Biotechnology

journal homepage:

Research Article

Proteomics-based identification of cancer-associated proteins in chronic lymphocytic leukaemia

Suliman A. Alsagaby [a,b,*], Ian A. Brewis [c], Rajendran Vijayakumar [d], Fahad A. Alhumaydhi [e], Ameen S. Alwashmi [e], Naif K. Alharbi [f,g], Waleed Al Abdulmonem [h], Mariappan Premanathan [a], Guy Pratt [i], Christopher Fegan [b], Christopher Pepper [b,j], Paul Brennan [b]

[a] Department of Medical Laboratories Sciences, College of Applied Medical Sciences, Majmaah University, Majmaah 11932, PO Box 1712, Saudi Arabia
[b] Division of Cancer & Genetics, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, United Kingdom
[c] TIME Institute, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, United Kingdom
[d] Department of Medical Laboratories Sciences, College of Applied Medical Sciences, Majmaah University, Majmaah 11932, PO Box 1712, Saudi Arabia
[e] Department of Medical Laboratories, College of Applied Medical Sciences, Qassim University, Buraydah, Saudi Arabia
[f] Department of Infectious Disease Research, King Abdullah International Medical Research Center, Riyadh, Saudi Arabia
[g] King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia
[h] Department of Pathology, College of Medicine, Qassim University, Buraydah, Saudi Arabia
[i] CRUK Institute for Cancer Studies, University of Birmingham, Vincent Drive, Edgbaston, Birmingham B15 2TT, United Kingdom
[j] Brighton and Sussex Medical School, University of Sussex, Brighton, United Kingdom

ARTICLE INFO

Article history:
Received 5 August 2020
Accepted 19 April 2021
Available online 27 April 2021

Keywords:
Biomarkers
Cancer-associated proteins
Chronic lymphocytic leukaemia
HP1BP3
Nucleolin
Proteomics
THRAP3

ABSTRACT

Background: Chronic lymphocytic leukaemia (CLL) is a neoplasm of B-cells characterized by variable prognosis. Exploring the proteome of CLL cells may provide insights into the disease. Therefore, eleven proteomics experiments were conducted on eleven primary CLL samples.
Results: We reported a CLL proteome consisting of 919 proteins (false discovery rate (FDR) $\leq$ 1%) whose identification was based on the sequencing of two or more distinct peptides (FDR of peptide sequencing $\leq$ 1%). Mass spectrometry-based protein identification was validated for four different proteins using Western blotting and specific antibodies in different CLL samples. Small sizes of nucleolin (~57 kDa and ~68 kDa) showed a potential association with good prognosis CLL cells (n = 8, p < 0.01). Compared with normal B-cells, CLL cells over-expressed thyroid hormone receptor-associated protein 3 (THRAP3; n = 9; p = 0.00007), which is implicated in cell proliferation; and heterochromatin protein 1-binding protein 3 (HP1BP3; n = 10; p = 0.0002), which promotes cell survival and tumourogenesis. A smaller form of HP1BP3, which may correspond to HP1BP3 isoform-2, was specifically identified in normal B-cells (n = 10; p = 0.0001). *HP1BP3* and *THRAP3* predicted poor prognosis of CLL ($p \leq 0.05$). Consistently, *THRAP3* and *HP1BP3* were found to be associated with cancer-related pathways ($p \leq 0.05$).
Conclusions: Our findings add to the known proteome of CLL and confirm the prognostic importance of two novel cancer-associated proteins in this disease.
How to cite: Alsagaby SA, Brennan P, Brewis IA, et al. Proteomics-based identification of cancer-associated proteins in chronic lymphocytic leukaemia. Electron J Biotechnol 2021;51. https://doi.org/10.1016/j.ejbt.2021.04.006

1. Introduction

Chronic lymphocytic leukaemia (CLL) is a malignant disease that affects B-cells and is characterised by a lymphocytosis of mature like B-cells that accumulate in peripheral blood, lymph node, bone marrow and spleen [1]. CLL is predominantly a disease affecting older people; the median age of diagnosis is 70 years of age [2]. The cause of CLL is still unknown, yet some hereditary factors have been found to increase the risk of developing CLL [3]. Significant advances in CLL therapy have been accomplished yet the current treatment does not eradicate the disease but reduces the tumour burden and helps prolong patients' survival [4,5].

While genetic factors and chromosomal abnormalities play key roles in CLL [6], aberrant protein expression has been shown to

Peer review under responsibility of Pontificia Universidad Católica de Valparaíso
* Corresponding author.
E-mail address: s.alsaqaby@mu.edu.sa (S.A. Alsagaby).

https://doi.org/10.1016/j.ejbt.2021.04.006
0717-3458/

heavily impact on the biology and the clinical outcomes of the disease [7]. Signalling through the B-cell receptor is a key factor for the progression of CLL [8]. Interestingly, BCR expression and signalling was shown to be strongly associated with the CLL cells carrying unmutated immunoglobulin variable heavy (*IGHV*) genes (UM-CLL) compared with those harbouring mutated *IGHV* (M-CLL) [9]. Consequently, UM-CLL is typically characterized by poor prognosis, whereas M-CLL identifies patients with a generally favourable clinical course [10,11]. Another interesting protein is zeta chain-associated protein kinase 70 (ZAP-70) that is not expressed in normal B-cells, but is highly produced in CLL cells from patients with a rapid progression of the disease, where it aberrantly participates in the signal transduction following BCR engagement [12,13,14]. Furthermore, over-expression of CD38, CXCR4, and CD49d is characteristic of enhanced migration of CLL cells to the microenvironment of lymph nodes, in which they receive proliferative and pro-survival signals, and are considered predictors of a high-risk form of the disease [15,16,17,18,19]. In contrast to normal B-cells, CLL cells exhibit an increased expression of the anti-apoptotic protein BCL-2, which among others, such as MCL-1, promote the survival of CLL cells and predict an undesirable prognosis of the disease [20,21,22]. Collectively, these findings highlight the heavy influence of aberrant protein expression on the behaviour of CLL.

Proteomics studies have been conducted on CLL cells to provide new insights into the biology of CLL by identifying new proteins with the potential to alter the pathology [23]. For example, heat shock protein 27, which inhibits apoptosis [24], was associated with high-risk CLL [25]. In addition, haematopoietic lineage cell-specific protein, which was implicated in CLL migration and tissue infiltration [26] was shown to correlate with the rapid progression of CLL [27]. Furthermore, the onco-protein T cell leukaemia/lymphoma enhances the activation of the NF-κB pathway and induces a CLL-like disease in transgenic mice, was reported to be highly expressed in CLL cells from patients with poor porgnonsis [28,29]. In a comprehensive proteomics study, increased expression of adhesion proteins, such as CD44, and decreased expression of proteins involved in the egress of CLL cells from lymph node, like adenylate cyclase-inhibiting G alpha protein and dynamin-2, were implicated in the retention of the malignant cells in the microenvironment of lymph nodes [30]. Therefore, proteomics holds great promise in being able to deepen our understanding about the aberrant protein expression associated with malignant diseases such as CLL.

In the present work we reported a CLL proteome consisting of 919 proteins with false discovery rate (FDR) = 1%. Our findings identified small sizes of nucleolin specifically in CLL cells from patients with good prognosis and detected an up-regulated expression of THRAP3 and HP1BP3 in CLL cells compared with normal B-cells. Consistently, high expression of *HP1BP3* and *THRAP3* was found to significantly predict poor prognosis in CLL patients.

2. Experimental

2.1. Isolation of CLL cells

Peripheral blood samples were drawn into EDTA tubes from CLL patients prior to differential blood count monitoring. The reminder of the samples (approximately 3 ml) was utilized to extract mononuclear peripheral blood cells (PBMCs) using lymphoprep (Stemcell Technologies) and gradient centrifugation. The project was approved by the South East Wales Research Ethics Committee and all clinical materials were obtained with the patients' written informed consent in accordance with the ethical approval (02/4806). CD19 positive population was assessed in the isolated PBMCs using an Accuri C6 cytometer and CFlow software (BD) following staining with anti-human CD19 antibody conjugated to allophycocyanin (Invitrogen). Samples with less than 95% CD19 positive cells (CLL cells) were subjected to T cell depletion using CD3-magnetic beads (Invitrogen) in order to enrich for CLL cells. Viability of the samples was measured using an Annexin V apoptosis assay kit (eBiosceience) and flow cytometry. Table 1 summarised the characteristics of CLL samples that were used in this study. Detailed description of the CLL samples and the disease prognosis are provided in Table S1.

2.2. Isolation of normal B-cells from healthy donors

Fresh buffy coat samples from healthy volunteers (n = 4) were obtained from the Welsh Blood Service. Lymphoprep and gradient centrifugation were used to isolate PBMCs from the buffy coat samples. Assessment of the CD19 positive cells in the isolated PBMCs was conducted using an Accuri C6 cytometer and CFlow software (BD) following staining with anti-human CD19 antibody conjugated to allophycocyanin (Invitrogen). Positive isolation using CD19-magnetic beads (Invitrogen) was performed to enrich for B-cells. Next, CD19-magentic beads were detached from the isolated cells (B-cells) using DETACHaBEADS CD19 (Invitrogen). The viability of B-cells was assessed using an Annexin V apoptosis assay kit (eBiosceience) and an Accuri C6 cytometer with CFlow software (BD).

2.3. Cellular fractionation

Thawed cells were treated with 0.5 ml hypotonic buffer (10 mM triethylammonium bicarbonate, 1.5 mM $MgCl_2$, and 10 mM KCl) containing 1 mM phenylmethanesulfonylfuoride and 0.1% (v/v) Nonidet P40 (NP40) detergent for 15 min on ice. Samples were then centrifuged for 5 min (10,000 ×g) at 4°C and the supernatant from each sample was transferred into another tube and labelled NP40 fractions. The remaining pellets were washed twice with 1 ml hypotonic buffer followed by centrifugation. Next, 0.2 ml 1% sodium dodecyl sulphate (SDS) solution was added and the sample was subjected to heating at 90°C for 20 min followed by sonication. This was sufficient to solubilize the washed pellets (termed SDS fractions). Protein concentrations were measured using the bicinchoninic acid assay (BCA) protein assay (Sigma).

To generate complete cell lysate, cells were resuspended in 0.2 ml 1% SDS supplemented with PMSF (final concentration 1 mM) and sonicated for 5 min. To enhance the solubilization of the cells, samples were heated at 90°C for 20 min. Following centrifugation of the samples for 5 min (10,000 ×g), cellular pellet was not visible indicating complete solubilization of the cells. Protein concentration of the complete cell lysate was determined using the bicinchoninic acid assay (BCA) protein assay (Sigma).

Table 1
Characteristics of the CLL patient cohort.

Factor	Subset	Number
Median Age	70	
Range	53–88	36
CD38	<20%	16
	≥20%	20
IGHV status	<98%	12
	≥98%	15
ZAP-70	<20%	9
	≥20%	21

IGHV status ≥ 98%, ZAP-70 ≥ 20% and CD38 ≥ 20% are characteristics of poor prognosis. However, *IGHV* status < 98%, ZAP-70 < 20% and CD38 < 20% are predictors of good prognosis.

2.4. Protein detection by gel electrophoresis and western blotting

Cellular proteins were separated by one dimensional (1D) gel electrophoresis using pre-cast NuPAGE 4–12% Bis-Tris Zoom gels (Invitrogen) under reducing conditions. For specific protein detection, proteins were transferred onto polyvinylidene fluoride membranes (GE Healthcare), probed with specific antibodies and identified using alkaline phosphatase chemiluminescent detection protocol. Antibodies against eight different proteins were applied: from Sigma anti-actin antibody (A-2066; used at 1/1000); from Abcam anti-histone H4 antibody (ab31830; used at 0.5/1000), anti-HP1BP3 antibody (ab98894; used at 1/1000) and anti-THRAP3 antibody (ab71985; used at 1/1000); from Sant Cruz Biotechnology anti-nucleolin antibody (sc-8031; used at 1/1000); from Upstate Biotechnology anti-coronin-1A antibody (#07-493; used at a 1/2000 dilution).

2.5. Protein precipitation and digestion

Protein precipitation using a 2D clean up kit (GE Healthcare) was performed on identical amounts of protein (20 μg) from NP40 fraction, SDS fraction and complete cell lysate that were generated from CLL samples. Following resuspension of the precipitated proteins of NP40 and SDS fractions in 20 μL of 20 mM triethylammonium bicarbonate (TEAB) they were denaturated using 1 μL of 2% SDS and reduced with 2 μL of 50 mM tris(2-carboxyethyl) phosphine (TCEP) at 60°C for 1 h. Protein alkylation was performed using 1 μL of 200 mM methylmethane-thiosulfonate in isopropanol for 10 min at room temperature. Finally, protein digestion was conducted using 2 μL of sequencing grade porcine trypsin (Promega) overnight at 37°C (the ration of enzyme/substrate = 1/20).

Precipitated proteins from the complete lysate were resuspended in 50mMTris HCl with 0.1% SDS. DTT (final concentration = 20 mM) was added to the protein suspension and samples were incubated for 10 min at 95°C. Next, iodoacetamide (IAA) was added to the samples (final concentration = 40 mM) which were incubated at room temperature in the dark for 30 min. DTT (final concentration = 10 mM) was added to the samples. For digestion, samples were incubated with 2 μL of sequencing grade porcine trypsin (Promega) overnight at 37°C (the ration of enzyme/substrate = 1/20).

2.6. Separation of peptides by liquid chromatography

Peptides (13.3 μg) from NP40 and SDS fractions of CLL samples were dried using a vacuume centrifuge and then resuspended in 15 μL LC loading solution (2% (v/v) acetonitrile in water with 0.05% (v/v) trifluoroacetic acid (TFA)). Next, from the 13.3 μg peptides 9 μg was injected in a 2D nano LC system (UltiMate 30000, Dionex). Peptides from the NP40 fractions and SDS fractions were independently separated on a strong cation-exchange (SCX) column (Bio-SCX, 500 μm, 15 mm, 5 μm, Dionex) using increasing concentrations of NaCl (flow-through, 100 mM, 200 mM, 400 mM, 800 mM, and 1 M). Then, the peptide fractions were cleaned using a reverse-phase desalting column. Next the peptides were subjected to additional separation using a reverse-phase (RP) column (PepMap 75 μm i.d., 30 cm, 3 μm, 100 Å, Dionex) at a flow rate equal to 300nL/min. We used two buffers for separation on the RP column; A = 2% (v/v) ACN in water with 0.05% (v/v) TFA and B = 90% (v/v) ACN in water with 0.01% (v/v) TFA. The separation on the RP column was conducted using an increasing concentration of solvent B; from 5% to 20% for 25 min and then from 20% to 50% for 25 min. To monitor the separation profile, a chromatogram was recorded at 214 nm. A robot micro-fraction collector (Dionex) was used to spot the separated peptides (1 spot/8 seconds) onto an LC/MALDI plate. MALDI matrix α-cyano-4-hydroxycinnamic acid (CHCA) (2 mg/ml CHCA in 70% (v/v) in 0.1% (v/v) TFA containing 10 fmol/ml Glu-Fib) was constantly added to the column effluent via an m-tee mixing piece at a flow rate of 1.4 ml/min.

Peptides (10 μg) from the complete cell lysate of CLL samples were dried using a vacuume centrifuge and then resuspended in 15 μL LC loading solution (2% (v/v) acetonitrile in water with 0.05% (v/v) fluoroacetic acid (FA)). Of the 10 μg peptides, 2 μg were injected into HPLC (1268 Infinity; Agilent Technologies) coupled with Chrip-Cube (Agilent Technologies). The separation was conducted on a Polaris-HR-Chip-3C18 with a 360 nl enrichment column and a 75 μm × 150 mm analytical column, which both are packed with Polaris C18-A, 180 Å, 3 μm stationary phase (Agilent Technologies). Mobile phase consisted of two puffers: A = 5% (v/v) ACN in water with 0.1% (v/v) FA and B = 95% (v/v) ACN in water with 0.1% (v/v) FA. The peptides were injected into the enrichment column using the mobile phase A at flow rate = 3 μl/min. Next, the separation of the peptide on the analytical column was performed using an increasing concentration of solvent B: 3% to 55% for 46 min and from 55% to 75% for 15 min at flow rate = 0.4 μl.

2.7. Mass spectrometry analysis

Resolved peptides from the NP40 fractions and SDS fractions were analysed using an Applied Biosystems 4800 MALDI TOF/TOF mass spectrometer with a 200 Hz solid-state laser operating at a wavelength of 355 nm. The precursor masses in all LC-MALDI sample positions were measured in an MS positive reflector mode using 800 laser shots (mass range 700–3000 Da; focus mass 1400). Glu-Fib that was added to the MALDI matrix was used as an internal standard in each spot for the MS analysis. For peptide sequencing, precursor fragmentation was conducted for the most abundant six precursors with signal to noise higher than 30 in each spot position. Common trypsin autolysis peaks and matrix ion signals and precursors within 300 resolution of each other were excluded from the selection. In MS/MS positive ion mode 4000 spectra were averaged with 1 kV collision energy (collision gas was air at a pressure of 1.661026 Torr) and default calibration.

The resolved peptides from the complete cell lysate were eluted into Agilent 6540 UHD Accurate-Mass quadrupole time-of-flight (Q-TOF) mass spectrometer through electro-spray ionization (ESI) source using Agilent HPLC Chrip-Cube. The ESI was performed with a drying gas (nitrogen) at temperature of 355°C and flow rate 4 L/min. The mass spectrometry analysis was conducted at positive ion mode with a mass/charge (m/z) range of 250–3000 and five spectra/sec. Auto MS/MS of precursor ions was done with the following settings: maximum precursor per cycle was five, absolute threshold was 2000 counts and relative threshold percentage was 0.01. Active exclusion was enabled at 2 spectra and released after 1 min. Precursor charge state selection and preference were set to 2+ and then 3+.

2.8. Data analysis

To perform protein identification based on peptide sequencing, the MS and MS/MS spectra from the NP40 fractions and SDS fractions were loaded into ProteinPilot 2.0.1 software (Applied Biosystems) with the Paragon and ProGroup algorithms and searched against the Human SwissProt database (release 2018(04)) that had 20,385 entries. The search was performed only against human proteins with trypsin as an enzyme used for digestion and cysteine alkylation with methyl methane-thiosulfonate. One missed cleavage was allowed. Two different scores are reported by ProteinPilot for each protein identification. First, the total protein score, which

gives a measurement of all peptides that contribute to an identification of a protein. Second, the unused protein score, which provides an assessment of only peptides that are specifically used to identify a protein but not other proteins with a higher rank. To determine the rate of incorrect protein identification, ProteinPilot 2.0.1 software was also used to run false discovery rate (FDR) analysis on the MS and MS/MS spectra. Protein identification was restricted to those that were reported with an unused protein score ≥ 2 with confidence threshold $\geq 99\%$ and FDR = 1%.

The MS and MS/MS spectra generated from the complete cell lysate were loaded into Spectrum Mill Protein Identification software (Agilent Technologies) and searched against the Human SwissProt database (release 2018(04)) that had 20,385 entries. The search parameters were set as follows: protein digestion was by trypsin, the maximum number of missed cleavages was two as recommended by the user manual provided by Agilent Technologies. Carbamidomethylation of cysteine was selected for fixed modification, and for variable modifications (oxidized methionine, pyroglutamic acid (N-termQ), phosphorylated serine, phosphorylated threonine and phosphorylated tyrosine) were selected. The precursor mass tolerance was set at +/− 20 ppm and the product mass tolerance was +/− 50 ppm. To ensure reporting high quality data set of CLL proteome, auto-validation of peptide sequencing and protein identification was set at FDR $\leq 1\%$. Only protein identifications (with protein score > 20; FDR $\leq 1\%$) based on two or more district peptides (with peptide score > 6; FDR $\leq 1\%$) were reported in this study.

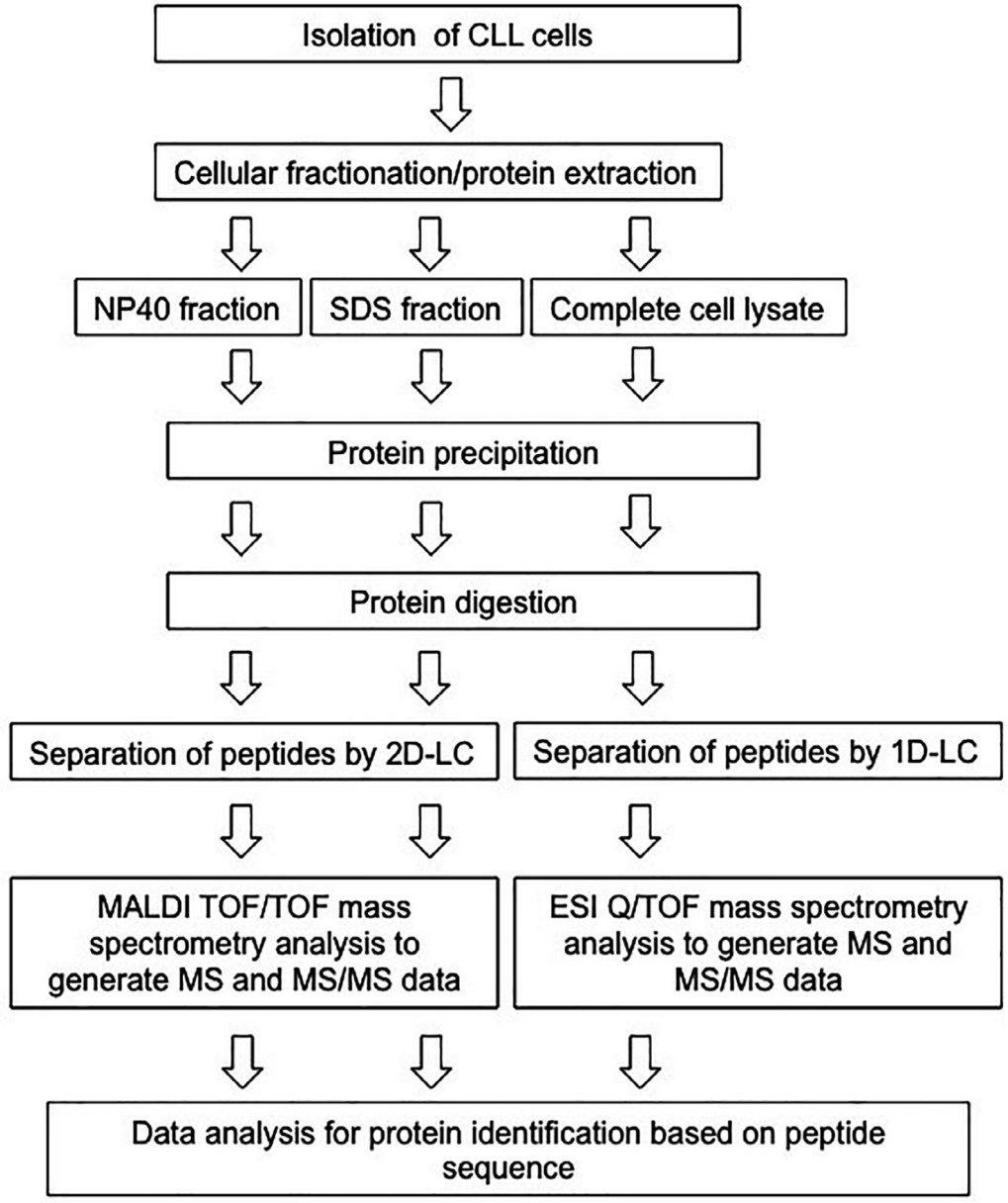


Fig. 1. Schematic illustration of the proteomics workflow that was used in the present study. Following the isolation of CLL cells, they were subjected to cellular fractionation. The protein extract in the NP40 fraction, SDS fraction and complete cell lysate was independently precipitated and digested using trypsin. The resultant peptides were separately resolved by two-dimensional liquid chromatography (2D-LC; NP40 and SDS fractions) or one-dimensional liquid chromatography (1D-LC; complete cell lysate). Next, peptides were subjected to mass spectrometry analysis to determine the peptide masses (MS) and the mass of peptide fragments (MS/MS); MLADI TOF/TOF MS for NP40 and SDS fractions and ESI Q/TOF MS for complete cell lysate. The MS and MS/MS data were used to perform peptide sequencing using ProteinPilot software or Spectrum Mill Protein Identification software with the Human SwissProt (protein) database in order to identify proteins.

2.9. Transcriptomics data sets

Two independent transcriptomics data sets of CLL (accession number: GSE22762 [31] and GSE39671 [32]) from GEO http://www.ncbi.nlm.nih.gov/geo/ were used for the analysis that aimed to investigate the prognostic potential of *THRAP3, HP1BP3* in CLL patients. Three reasons led to the selection of these two data sets. First, they included information about the clinical outcomes of the disease; GSE22762 contained data about overall survival (OS) and GSE39671 included information about time to first treatment (TTFT). Second, the two data sets were based on independent cohorts containing more than 100 patients each. Third, both data sets were generated using the same oligonucleotide microarray platform (Affymetrix Human Genome U133 Plus 2.0 Array). The DataSet SOFT files of the transcriptomics data sets were downloaded from GEO and used.

2.10. Statistical analysis

Protein scores (total and unused), FDR and confidence threshold of protein identification (in NP40 and SDS fractions) were calculated using ProteinPilot 2.0.1 software (Applied Biosystems) with the Paragon and ProGroup algorithms. Spectrum Mill Protein Identification software (Agilent Technologies) was used to calculate the score and FDR of peptides and proteins identified in the complete cell lysate. Prism software (graphpad) was also employed to construct Kaplan-Meier curves and to calculate *p*-values with hazard ratios (HRs) based on the long-rank test. Column graphs were prepared using Excel software and the associated *p*-values were calculated using the unpaired Student's *t*-test. Heatmap was constructed using the heatmapper tool (http://heatmapper.ca/) [33]. Pathway enrichment analysis and the associated *p* values were calculated using "Reactome Pathway Database" (https://reactome.org/) [34].

3. Results

3.1. Proteomics analysis

In the current study, we performed 11 proteomics experiments on 11 CLL samples with variable prognosis. Information about the prognostic markers of the 11 CLL samples is shown in Table S1. Six proteomics experiments (2D nano-LC with MALDI TOF/TOF MS) were conducted on the NP40 and SDS fractions from six CLL samples. Next, five proteomics experiments (1D nano-LC with ESI Q/TOF MS) were done on complete cell lysate generated from 5 CLL samples. The utilized proteomics workflow is illustrated and summarised in Fig. 1. We identified 359 proteins in the NP40 fractions and 142 proteins in the SDS fractions (73 proteins were common in both fractions; total number of proteins = 428). Table S2, Table S3, Table S4, Table S5 and Table S6 show the full list of the reported proteins and the peptides used for protein identifications. The 428 proteins were identified using two or more distinct peptides with a minimum unused protein score = 2 (confidence threshold $\geq$ 99% and FDR = 1%). From the complete cell lysate of five CLL samples we reported 718 proteins (Table S6 and Table S7). The protein identification was based on two or more distinct peptide with peptide score > 6 (FDR $\leq$ 1%) and protein score > 20 (FDR $\leq$ 1%). Collectively, our study reported a CLL proteome consisting of 919 proteins with FDR $\leq$ 1% (Table 2).

Our study identified 359 proteins in the NP40 fractions and 142 proteins in the SDS fractions; 73 proteins were common in the two fractions (Fig. 2A). To explore the type of protein enrichment in the NP40 fractions and in the SDS fractions, we used Gene Ontology (GO) data coupled with Quick GO tool (https://www.ebi.ac.uk/QuickGO/) [35]. Proteins specific to the NP40 fractions and to the SDS fractions were annotated to one or more of the following GO terms "cytosol, membrane, mitochondrion or nucleus". The analysis revealed that cytosolic proteins were more predominant in the NP40 fractions, while nuclear proteins dominated the SDS fractions (Fig. 2B).

3.2. Validation of the proteomics findings

Our proteomics work reported a CLL proteome consisting of 919 proteins (FDR $\leq$ 1%). To confirm the mass spectrometry-based protein identification, the detection of four proteins was validated in CLL samples using Western blotting and specific antibodies. The proteomics data showed that HP1BP3, histone H4, coronin-1A and nucleolin were identified in CLL cells with multiple peptides ($n \geq 3$), unused protein score $\geq$ 5.8 and sequence coverage $\geq$ 9.4%.

Table 2
Summary of the protein identifications in the three types of protein extract using different proteomic instruments.

Type of protein extract	Instruments	Size of identified proteome
NP40 fraction	2D nano-LC with MALDI TOF/TOF MS	359 proteins
SDS fraction	2D nano-LC with MALDI TOF/TOF MS	142 proteins
Complete cell lysate	1D nano-LC with ESI Q/TOF MS	718 proteins
All the three types of protein extract (NP40 fraction, SDS fraction and complete cell lysate)	2D nano-LC with MALDI TOF/TOF MS and 1D nano-LC with ESI Q/TOF MS	919 proteins

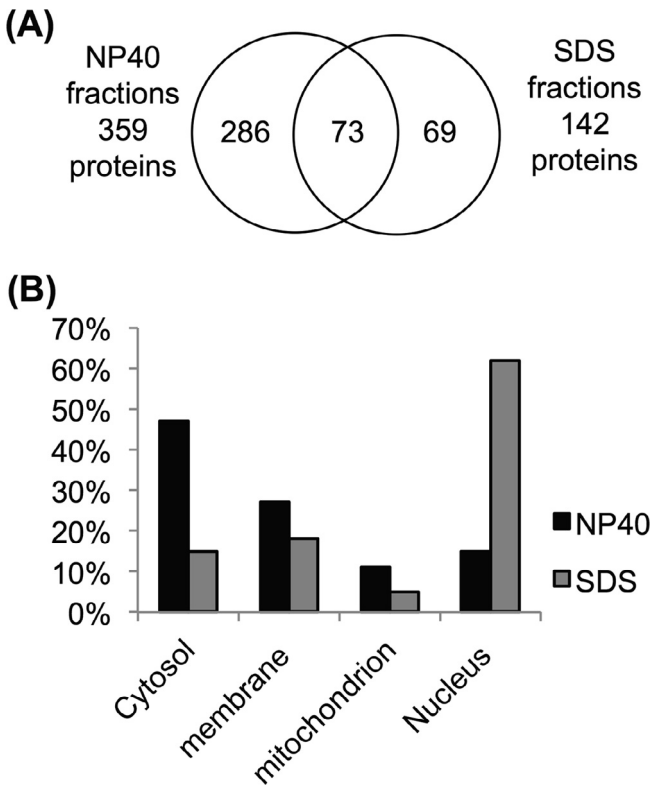
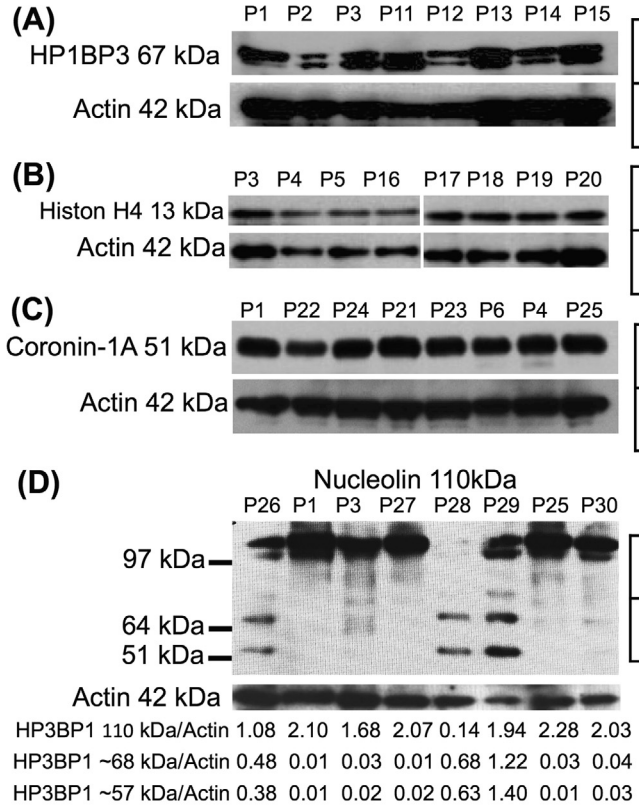


Fig. 2. Venn diagram and protein localization analyses of the reported proteins. Venn diagram shows common and specific proteins in the NP40 fractions and SDS fractions (A). Gene Ontology, coupled with Quick GO tool, was used to determine the type of protein enrichment in the NP40 fractions and SDS fractions (B).



Protein	Pep. #	score	%Seq.	Frac.
HP1BP3	3	5.8	9.4	SDS

Protein	Pep. #	score	%Seq.	Frac.
H4	38	29.75	56.3	SDS

Protein	Pep. #	score	%Seq.	Frac.
Coronin-1A	5	10	13.5	NP40 and SDS

Protein	Pep. #	score	%Seq.	Frac.
Nucleolin	6	9.8	11.6	NP40 and SDS

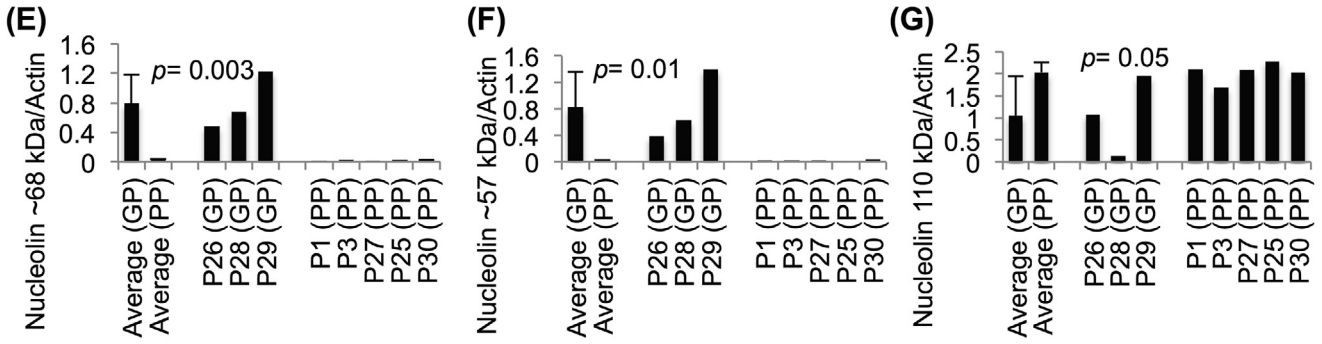


Fig. 3. Validation of the proteomics findings. Western blotting with specific antibodies was used to confirm the mass spectrometry-based protein identification. HP1BP3 (A) and Histone H4 (B) were investigated in SDS fractions of CLL samples. Coronin-1A (C) and nucleolin (D-G) were analysed in NP40 fractions of CLL samples. N: NP40 fraction; S: SDS fraction; Pep. #: peptide count that was used for protein identification by mass spectrometry; score: unused protein score; %Seq.: percentage of protein sequence coverage; Frac: fraction in which a protein was identified by mass spectrometry; GP: good prognosis (for the patients: P26, P28 and P29); PP: poor prognosis (for the patients: P1, P3, P27, P25 and P30).

The validation analyses of these proteins in CLL samples (n = 8 for each protein) using specific antibodies agreed with the mass spectrometry findings (Figs. 3A-4D). Lower molecular weight bands of nucleolin (two bands: ~57 kDa and ~ 68 kDa) were predominantly detected in CLL samples from patients at the early stage of the disease (stage A) and with good prognostic markers (M-CLL and ZAP-70 negative; n = 8, $p \leq 0.008$, Fig. 3D, Fig. 3E and Fig. 3F respectively). These data may propose a potential association of the small size nucluleolin with good prognosis of CLL. This finding was also associated with a decreased ratio of the full-length nucleolin (110 kDa) to actin in CLL cells from patients with good prognosis (n = 8, p = 0.05, Fig. 3D and 3G).

3.3. CLL-associated proteins

CLL cells were previously shown to upregulate the expression of RNA processing proteins. Therefore, we studied the expression of the RNA processing proteins (THRAP3 and HP1BP3) in CLL cells versus normal B-cells. Our proteomics data set showed that THRAP3 was detected specifically in SDS fractions with 3 different peptides, unused protein score = 6 and sequence coverage = 6.8. The identification details of HP1BP3 by mass spectrometry were mentioned earlier.

It has been controversial what defines the best normal counterpart to CLL cells, with B-cells expressing CD19, CD5, CD23 and TLR9 being suggested as good normal match of CLL cells [36]. CD5+ normal B-cells with or without CD23 and TLR9 represent a small proportion of normal B-cells population (approximately 15%) [37], thus larger volume of samples and extensive cell sorting are required to isolate sufficient number of CD5 B-cells for western blotting analysis. As a result, this restricted us to the use of CD19+ B-cells from the peripheral blood of healthy donors as a normal control of CLL cells.

The isolation of normal B-cells from four healthy donors began with the isolation of peripheral blood mononuclear cells (PBMCs)

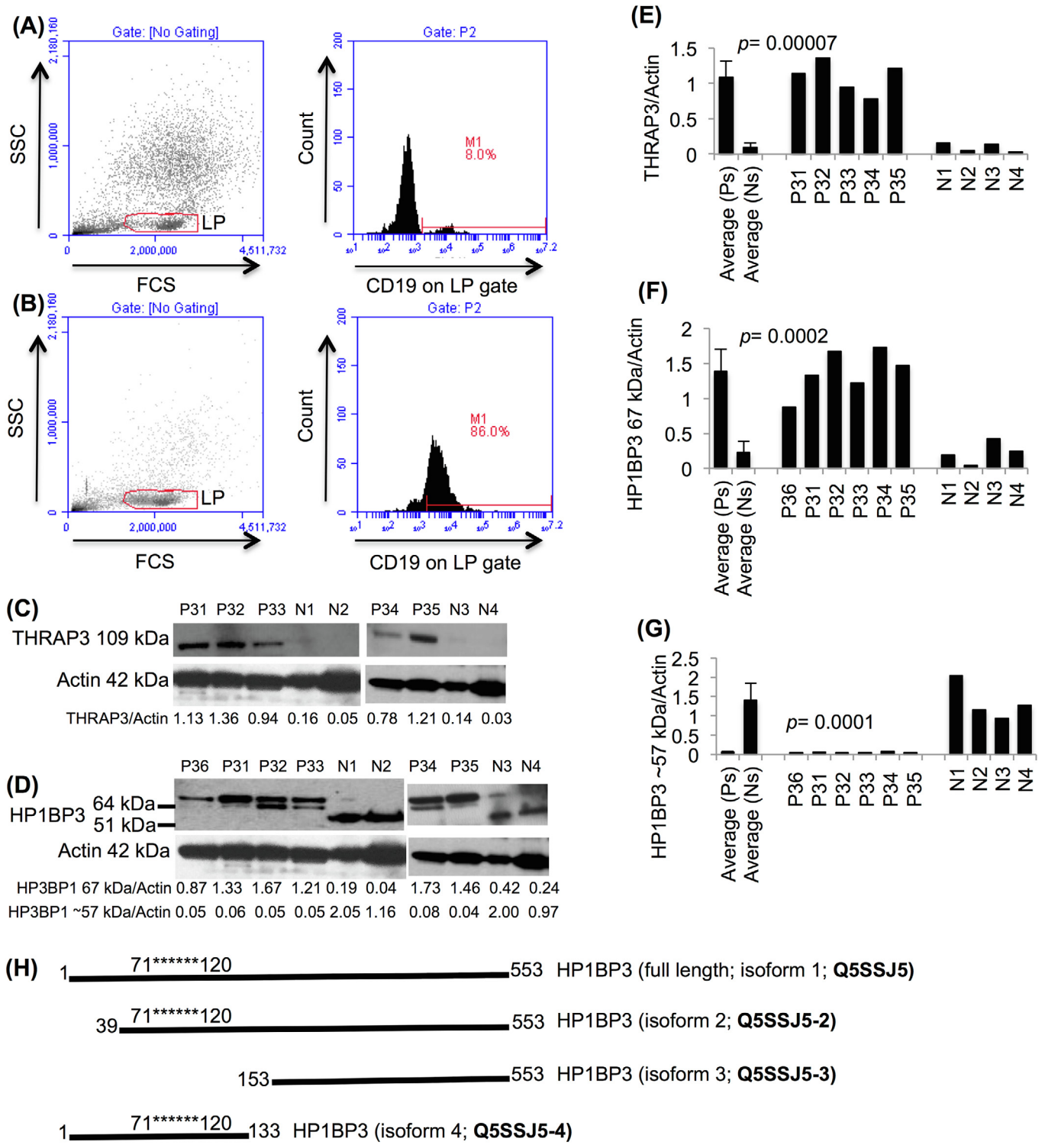


Fig. 4. CLL-associated proteins. Four normal B-cells samples were isolated from buffy coat of four healthy donors (A and B). Western blotting and specific antibodies were used to investigate the expression of THRAP3 and HP1BP3 in CLL samples (P31, P32, P33, P34, P35 and P36) and normal B-cells (N1-N4; C-G). Alternative splicing is known to produce four isoforms of HP1BP3; the amino acid sequence (71******120) is the region of the protein that the anti-HP1BP3 antibody (ab98894) identifies (H). SSC: side-scattered light; FSC: forward-scattered light; LP: lymphocyte population.

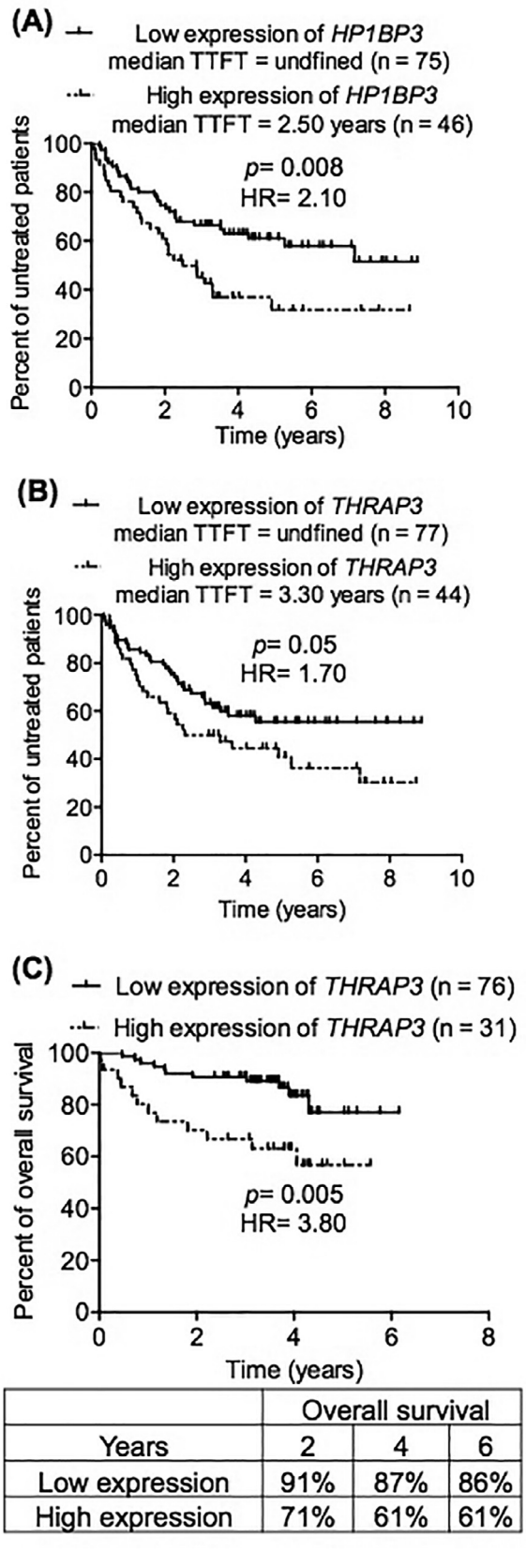
from buffy coat samples. Subsequent flow cytometric analysis showed that further enrichment of B-cells was required (average CD19+ cells = 5%; standard deviation = 2%; Fig. 4A). Next, positive purification of B-cells was conducted using CD19-magentic beads followed by DETACHaBEADS CD19 reagents to remove the beads from the isolated B-cells. This step resulted in an enrichment of the CD19+ population (CD19+ cells ≥ 85%; Fig. 4B).

The expression of THRAP3 and HP1BP3 was evaluated in CLL cells and normal B-cells using Western blotting and specific antibodies. The analyses showed that these two proteins were over-expressed in CLL cells compared with normal B-cells. The average of THRAP3/actin was 1.10 in CLL cells compared with 0.10 in normal B-cells (n = 9; p = 0.00007; Fig. 4C and Fig. 4E). Similarly, the average of HP1BP3/actin was 1.38 in CLL cells compared with

0.22 in normal B-cells (n = 10; p = 0.0002; Fig. 4D and Fig. 4F). Interestingly, the anti-HP1BP3 antibody (ab98894) detected a smaller size of HP1BP3 (~57 kDa) specifically in the normal B-cells, but not in the CLL cells (n = 10, p = 0.0001; Fig. 4D and Fig. 4G). Alternative splicing was reported to produce 4 different isoforms of HP1BP3, which are documented at the level of protein in the UniProt knowledgebase (https://www.uniprot.org/). Fig. 4H shows a representation of these isoforms and the region that is targeted by the anti-HP1BP3 antibody (ab98894) that was used in this study. Therefore, the small size of HP1BP3 (~57 kDa) detected spcifically in normal B-cells is lieky to be the isoform 2 of the protein.

Public repositories like GEO represent rich sources of omics data sets that can be used to investigate the relation between a gene of interest and a disease [38,39]. Therefore, two independent transcriptomics data sets of CLL (GSE22762 and GSE39671) from GEO with clinical details about overall survival (OS) or time-to-first-treatment (TTFT) were used to investigate whether the transcript expression of THRAP3 and HP1BP3 could be relevant to CLL prognosis. Initially the median expression of the transcripts corresponding to HP1BP3 and THRAP3 in each data set was used to divide the patients into two groups: a low expression group, where the expression of *THRAP3* and *HP1BP3* was below the median, and a high expression group, where the expression of *HP1BP3* and *THRAP3* was above the median. Next, the OS and TTFT data were compared in the low and high expression groups of the two data sets using Kaplan-Meier curve. The analysis showed that the median expression of *THRAP3* and *HP1BP3* did not significantly segregate CLL patients into two groups with different prognosis. Therefore, an effort was made using the "Cutoff Finder" tool (http://molpath.charite.de/cutoff/) [40] to search for a cutoff value (z-score) of the expression of *THRAP3* and *HP1BP3* that significantly predicts the prognosis of CLL patients. As a result, an increased expression of *HP1BP3* (above a cutoff value: z-score = 0.41) was found to predict early need of therapy (n = 121, p = 0.008, HR = 2.10, Fig. 5A). Likewise, high expression of *THRAP3* identified patients with early treatment (z-score = 0.21, n = 121, p = 0.05, HR = 1.70, Fig. 5B) and short overall survival (z-score = 0.57, n = 107, p = 0.005, HR = 3.80, Fig. 5C).

Next, we attempted to explain the association of *THRAP3* with poor prognosis of CLL by employing Pearson score (PS) to identify genes that correlated with the expression of THRAP3 in CLL cells from 130 patients (data set GSE39671). Correlation with the expression of *THRAP3* was detected for 210 genes (PS $\geq$ 0.40; p < 0.00001). Interestingly, some of these genes play roles in cell division and tumor growth, such as cell division cycle 14A (*CDC14A*), cell division cycle 40 (*CDC40*), cell cycle associated protein 1 (*CAPRIN1*), microtubule-associated protein 9 (*MAP9*), fibroblast growth factor receptor 1 (*FGFR1*), SRC proto-oncogene tyrosine kinase (*SRC*) and Src homology 2 domain containing adaptor protein B (*SHB)* and Kirsten rat sarcoma viral oncogene homolog (*KRAS*) (Fig. 6A). To gain insights into the functional roles of the 210 genes, we performed pathway enrichment analysis using "Reactome Pathway Database". The analysis found a significant enrichment by the 210 genes for cancer-related pathways like FGFR1 signalling, RAS signalling and VEGFR2 mediated cell proliferation (Table 3). Following the same concept, we employed PS to identify genes that exhibited correlation with the expression of *HP1BP3* in CLL cells using the data set GSE39671. The analysis found 2223 genes to be correlated with the expression of *HP1BP3* in 130 CLL patients (PS $\geq$ 0.40; p < 0.00001). Fig. 6B shows the correlation between *HP1BP3* and known CLL-related genes, such as *CD79B*, B-cell linker (*BLNK*), bruton tyrosine kinase (*BTK*), T-cell leukemia/lymphoma 1A (*TCL1A*) and *CD40*. Utilizing "Reactome Pathway Database" we conducted pathway enrichment analysis of the 2223 genes (Table 3).



	Overall survival		
Years	2	4	6
Low expression	91%	87%	86%
High expression	71%	61%	61%

Fig. 5. Investigations of the prognostic potential of *THRAP3* and *HP1BP3* in CLL. Two data sets of CLL from GEO were used (GSE22762 and GSE39671). Low expression was defined as patients with expression of *THRAP3* or *HP1BP3* < a z-score; high expression of *THRAP3* or *HP1BP3* was defined as > a z-score. The z-scores 0.41, 0.20, 0.57 were used for the analyses of *HP1BP3* with TTFT (A), *THRAP3* with TTFT (B) and *THRAP3* with OS (C), respectively. UD: undefined; HR: hazard ratio of high expression versus low expression.

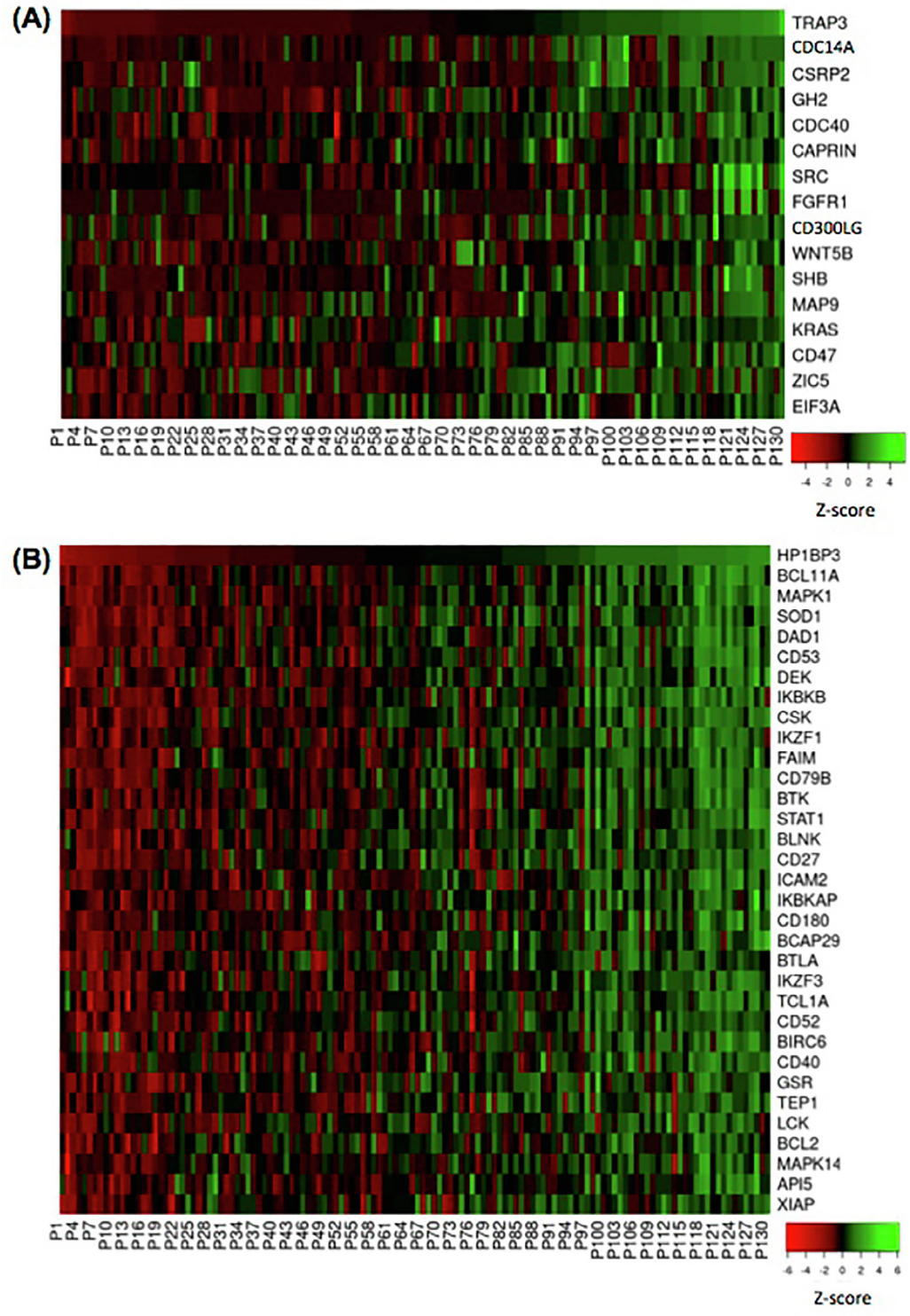


Fig. 6. Heatmap presentation of the correlation of *THRAP3* and *HP1BP3* with other genes in CLL cells from 130 patients. The analysis was conducted on the data set GSE39671; the Pearson score (PS) ranged from 0.40 to 0.58 for *THRAP3* (A) and 0.51 to 0.76 for *HP1BP3* (B) with *p* value less than 0.00001. Green is high expression and red is low expression.

4. Discussion

Proteomics is a useful approach to study aberrant protein expression in CLL in order to provide novel insights into the pathology of the disease [23]. The dynamic range of cellular proteins was reported to reach six-fold [41]. Therefore, cellular fractionation was previously described as an effective method to increase the chance of identifying larger number of proteins using gel-free based proteomic approaches [42,43,44,45]. Following the same concept, we performed cellular fractionation on CLL cells (NP40 fractions and SDS fractions). The NP40 fractions appeared to be enriched with cytosolic proteins, whereas the SDS fractions were enriched with nuclear proteins. For example, only one histone protein (nuclear protein) was detected in the NP40 fraction. In

Table 3
Pathway enrichment analysis of genes that correlated with *THRAP3* or *HP1BP3*.

Pathway name	No of genes found	No of genes in total	*p* value	Analysis
SHC-mediated cascade:FGFR1	4	30	0.002	*THRAP3*
FRS-mediated FGFR1 signaling	4	31	0.002	*THRAP3*
Signaling by FGFR1 in disease	5	58	0.003	*THRAP3*
Signaling by FGFR1	5	62	0.004	*THRAP3*
Signalling to RAS	3	26	0.010	*THRAP3*
Signaling by RAS mutants	4	54	0.014	*THRAP3*
Signaling by moderate kinase activity BRAF mutants	4	54	0.014	*THRAP3*
Signaling downstream of RAS mutants	4	54	0.014	*THRAP3*
Activation of RAS in B cells	2	11	0.015	*THRAP3*
VEGFR2 mediated cell proliferation	3	31	0.016	*THRAP3*
Signaling by EGFR	4	61	0.020	*THRAP3*
Signaling by Insulin receptor	5	97	0.025	*THRAP3*
RAF activation	3	37	0.025	*THRAP3*
Signalling to ERKs	3	42	0.035	*THRAP3*
Synthesis of DNA	54	132	0.000002	*HP1BP3*
DNA Replication	55	141	0.000006	*HP1BP3*
S Phase	64	179	0.00002	*HP1BP3*
mRNA Splicing	64	196	0.0002	*HP1BP3*
Oxygen-dependent proline hydroxylation of Hypoxia-inducible Factor Alpha	26	72	0.004	*HP1BP3*
Downstream signaling events of B Cell Receptor (BCR)	31	93	0.006	*HP1BP3*
Extension of Telomeres	21	62	0.02	*HP1BP3*
Activation of NF-kappaB in B cells	23	72	0.02	*HP1BP3*
Cell Cycle	167	715	0.03	*HP1BP3*

"Reactome Pathway Database" was employed to perform pathway enrichment analysis. "No of genes found" indicates the number of genes that correlated with *THRAP3* or *HP1BP3* and were assigned to a pathway. "No of genes in total" is the number of all known genes that function in a pathway.

contrast, 13 different histone proteins were identified in the SDS fractions. In the complete cells lysate preparation, only 4 histone proteins were identified, although the proteomics experiments were conducted using Agilent 6540 UHD Accurate-Mass Q-TOF mass spectrometer, which yielded larger proteome identifications compared with Applied Biosystems 4800 MALDI TOF/TOF mass spectrometer. Although the size of the proteome detected in the SDS fractions was small (142 proteins), working on SDS fractions seemed justifiable when DNA-associated proteins are of interest.

The good quality of our CLL proteome was evident at the validation analyses. The specific antibody-based protein identification of all examined proteins (n = 4) in CLL samples was consistent with the proteomics findings, adding more validity to the reported CLL proteome.

Nucleolin (110 kDa) is a pro-survival protein that stabilizes the transcript of BCL-2 in CLL cells and inhibits FAS-induced apoptosis in B-cell lymphomas [46,47]. In CLL cells, the expression of nucleolin positively correlates with the level of BCL-2, supporting the conception that nucleolin is a stabilizing protein of BCL-2 mRNA [47]. The same study also showed that in contrast to normal B-cells, CLL cells over-express both nucleolin and BCL-2. As an anti-apoptotic protein BCL-2 is one of the main drivers of the survival of CLL cells and is considered a marker of poor prognosis in CLL patients [20]. Our preliminary data (n = 8) showed lower abundance of the intact nucleolin (110 kDa) that was associated with the detection of small sizes of the protein (~68 kDa and ~ 57 kDa) in CLL cells from patients with good prognosis. These findings might indicate an active degradation of nucleolin (110 kDa) in CLL cells from patients with good prognosis compared with those from patients with poor prognosis. The stability of nucleolin (110 kDa) has been shown to be proliferation-dependent [48]. In proliferating peripheral blood lymphocytes (PBLs) nucleolin (110 kDa) was the main detected band, whereas smaller sizes of nucleolin (68 kDa, ~57 kDa and 43 kDa) were predominantly found in less proliferative or non-dividing PBLs [48]. Given the roles of nucleolin (110 kDa) in the survival and proliferation, our findings are in consistence with the phenotype of good prognosis of CLL. Furthermore, our data may provide an explanation, at least partially, for the lower capacity of good prognosis CLL cells to undergo proliferation and resist apoptosis compared with poor prognosis CLL cells [49]. The sample size we used to study nucleolin was small (n = 8). Therefore, studying the expression of nucleolin in a larger number of CLL samples is required to confirm our finding.

In the present study, THRAP3, which is an RNA processing protein, was found with high expression in CLL cells compared with normal B-cells. This finding supports a recently published work, where CLL cells were shown to over-express proteins implicated in mRNA processing [50]. THRAP3 plays role in the cell cycle progression, where it is required for the production of mature transcript of cyclin D1 [51]. CLL cells that express cyclin D1 were shown to reside in the proliferation centre in lymph nodes, indicating a role of cyclin D1 in the proliferation of CLL cells [52]. In contrast to the historical view, that described CLL as a dysfunctional apoptosis-driven disease, CLL is currently viewed as a dynamic cancer composed of malignant cells with variable rates of proliferation and apoptosis [49]. In fact, CLL cells have a prolonged proliferation history compared with normal B-cells as evidenced by the shorter telomeres of the malignant cells [53]. Given the role of THRAP3 on the progression of cell cycle, its increased expression in CLL cells may support their expansion. Interestingly, CLL cells harboring unmutated *IGHV* (poor prognostic marker) were found to up-regulate the expression of THRAP3 compared with those carrying mutated *IGHV* (good prognostic marker) [30]. In line with this view, using data sets of CLL transcriptomics from GEO, we showed that high expression of THRAP3 mRNA is predictive of early treatment and short overall survival in CLL patients in two independent transcriptomics data sets (total number of CLL patients = 228). In agreement with this finding, we also found genes that have been implicated in cell division and tumour growth, such as *CDC14A* [54], *CDC40* [55], *CAPRIN1* [56], *MAP9* (also known as ASAP) [57], *FGFR1* [58], *SRC* [59], *KRAS* [60], and *SHB* [61] to significantly correlate with the expression of *THRAP3* in CLL cells. These results may provide rationale for the prediction of poor prognosis of CLL by the high expression of *THARP3*.

Our study identified a specific expression of a smaller form (~57 kDa) of HP1BP3 in normal B-cells compared with CLL cells. HP1BP3 has four isoforms that result from alternative splicing that were documented at the level of protein in UniProt knowledge base (https://www.uniprot.org/). The anti-HP1BP3 antibody

(ab98894) used in this study targets a peptide sequence within the N terminal (amino acids: 71–120) of the protein. This sequence is present in HP1BP3 isoform 1 (size = 67 kDa), isoform 2 (size = 57 kDa) and isoform 4 (size = 15 kDa). It is possible that the HP1BP3 (~57 kDa), specifically detected in normal B-cells, corresponds to isoform 2. Therefore, this finding proposes HP1BP3 (~57 kDa) as a marker to aid in the differentiation between normal B-cells and CLL cells. Examining HP1BP3 (~57 kDa) in a larger number of CLL and normal B-cell samples would seem worthwhile in order to draw a definitive conclusion about its possible utility as a diagnostic marker.

The expression of the intact form of HP1BP3 (67 kDa) was higher in CLL cells compared with normal B-cells. HP1BP3 is implicated in cellular proliferation, where it maintains heterochromatin integrity during G1 to S phase of the cell cycle affecting the duration of G1 phase [62]. In addition, HP1BP3 is a pro-survival protein; it mediates chromatin condensation that protects cells from hypoxia-induced apoptosis [63]. Furthermore, HP1BP3 increases the tumourogenesis of malignant cells. For example, HP1BP3-deficient cancer cells (A341) produced tumours in 50% of injected mice, whereas A341 cells expressing HP1BP3 induced tumor growth in all tested mice and the tumor was bigger and more resistant to chemotherapy in mice that received HP1BP3 expressing A341 cells [63]. Consistent with the known roles of HP1BP3, our data showed that high expression of HP1BP3 transcript was predictive of shorter time to first treatment in CLL patients. In line with this finding, our correlation analysis identified an association between the expression of *HP1BP3* and others genes that enriched for CLL-related pathways such as cell cycling [49], BCR signaling [9], activation of NF-κB [64] and extension of telomeres [53]. Taken together, these findings suggest a role of HP1BP3 in the growth and survival of CLL cells.

5. Conclusions

Overall, our study adds to the known proteome of CLL. Nucleolin, THRAP3 and HP1BP3 are proteins that were implicated in cellular proliferation and survival; and appeared in our study to be CLL-associated proteins and/or predict the prognosis of the neoplasm. Therefore, our findings provide a rationale for further investigations on these proteins using a bigger number of samples to determine their roles in CLL with the view of novel biomarkers and targeted therapy.

Ethical approval

The present study was approved by the South East Wales Research Ethics Committee and all clinical materials were obtained with the patients' written informed consent in accordance with the ethical approval (02/4806).

Conflicts of interest

The authors declare no conflict of interest.

Financial support

The authors would like to thank the Deanship of Scientific Research at Majmaah University in Saudi Arabia for supporting this work under Project Number [R-2021-1].

Acknowledgments

The authors would like to thank the Cardiff University Central Biotechnology Service (CBS), Proteomic Facility.

Supplementary material

https://doi.org/10.1016/j.ejbt.2021.04.006.

References

[1] Dighiero G. CLL biology and prognosis. Hematology Am Soc Hematol Educ Program 2005;1:278–84. https://doi.org/10.1182/asheducation-2005.1.278 PMid: 16304392.
[2] Montserrat E, Bosch F, Rozman C. B-cell chronic lymphocytic leukemia: Recent progress in biology, diagnosis, and therapy. Ann Oncol 1997;8(1):S93–S101. https://doi.org/10.1093/annonc/8.suppl_1.S93 PMid: 9187440.
[3] Fabbri G, Dalla-Favera R. The molecular pathogenesis of chronic lymphocytic leukaemia. Nat Rev Cancer 2016;16(3):145–62. https://doi.org/10.1038/nrc.2016.8 PMid: 26911189.
[4] Shustik C, Bence-Bruckler I, Delage R, Owen CJ, Toze CL, Coutre S. Advances in the treatment of relapsed/refractory chronic lymphocytic leukemia. Ann Hematol 2017;96(7):1185–96. https://doi.org/10.1007/s00277-017-2982-1 PMid: 28389687.
[5] Bose P, Gandhi V. Recent therapeutic advances in chronic lymphocytic leukemia. F1000Research 2017;6:1924. https://doi.org/10.12688/f1000research.11618.1 PMid: 29152232.
[6] Döhner H, Stilgenbauer S, Benner A, et al. Genomic aberrations and survival in chronic lymphocytic leukemia. N Engl J Med 2000;343(26):1910–6. https://doi.org/10.1056/NEJM200012283432602 PMid: 11136261.
[7] Alsagaby SA, Brennan P, Pepper C. Key Molecular drivers of chronic lymphocytic leukemia. Clinical Lymphoma Myeloma Leukemia 2016;16(11):593–606. https://doi.org/10.1016/j.clml.2016.08.008 PMid: 27601002.
[8] Woyach JA, Johnson AJ, Byrd JC. The B-cell receptor signaling pathway as a therapeutic target in CLL. Blood 2012;120(6):1175–84. https://doi.org/10.1182/blood-2012-02-362624 PMid: 22715122.
[9] Stevenson FK, Krysov S, Davies AJ, et al. B-cell receptor signaling in chronic lymphocytic leukemia. Blood 2011;118(16):4313–20. https://doi.org/10.1182/blood-2011-06-338855 PMid: 21816833.
[10] Hamblin TJ, Davis Z, Gardiner A, et al. Unmutated Ig VH genes are associated with a more aggressive form of chronic lymphocytic leukemia. Blood 1999;94(6):1848–54. https://doi.org/10.1182/blood.V94.6.1848 PMid: 10477713.
[11] Damle RN, Wasil T, Fais F, et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. Blood 1999;94(6):1840–7. https://doi.org/10.1182/blood.V94.6.1840 PMid: 10477712.
[12] Chan AC, Van Oers N, Tran A, et al. Differential expression of ZAP-70 and Syk protein tyrosine kinases, and the role of this family of protein tyrosine kinases in TCR signaling. J Immunol 1994;152(10):4758–66. PMid: 8176201.
[13] Chen L, Widhopf G, Huynh L, et al. Expression of ZAP-70 is associated with increased B-cell receptor signaling in chronic lymphocytic leukemia. Blood 2002;100(13):4609–14. https://doi.org/10.1182/blood-2002-06-1683 PMid: 12393534.
[14] Rassenti LZ, Huynh L, Toy TL, et al. ZAP-70 compared with immunoglobulin heavy-chain gene mutation status as a predictor of disease progression in chronic lymphocytic leukemia. New England J Med 2004;351(9):893–901. https://doi.org/10.1056/NEJMoa040857 PMid: 15329427.
[15] Deaglio S, Aydin S, Grand MM, et al. CD38/CD31 interactions activate genetic pathways leading to proliferation and migration in chronic lymphocytic leukemia cells. Mol Med 2010;16(3–4):87–91. https://doi.org/10.2119/molmed.2009.00146 PMid: 19956559.
[16] Burger JA, Burger M, Kipps TJ. Chronic lymphocytic leukemia B cells express functional CXCR4 chemokine receptors that mediate spontaneous migration beneath bone marrow stromal cells. Blood 1999;94(11):3658–67. https://doi.org/10.1182/blood.V94.11.3658 PMid: 10572077.
[17] Majid A, Lin TT, Best G, et al. CD49d is an independent prognostic marker that is associated with CXCR4 expression in CLL. Leuk Res 2011;35(6):750–6. https://doi.org/10.1016/j.leukres.2010.10.022 PMid: 21093051.
[18] Ibrahim S, Keating M, Do K-A, et al. CD38 expression as an important prognostic factor in B-cell chronic lymphocytic leukemia. Blood 2001;98(1):181–6. https://doi.org/10.1182/blood.V98.1.181 PMid: 11418478.
[19] Bulian P, Shanafelt TD, Fegan C, et al. CD49d Is the strongest flow cytometry–based predictor of overall survival in chronic lymphocytic leukemia. J Clin Oncol 2014;32(9):897–904. https://doi.org/10.1200/JCO.2013.50.8515 PMid: 24516016.
[20] Pepper C, Lin TT, Pratt G, et al. Mcl-1 expression has in vitro and in vivo significance in chronic lymphocytic leukemia and is associated with other poor prognostic markers. Blood 2008;112(9):3807–17. https://doi.org/10.1182/blood-2008-05-157131 PMid: 18599795.
[21] Hanada M, Delia D, Aiello A, et al. bcl-2 gene hypomethylation and high-level expression in B-cell chronic lymphocytic leukemia. Blood 1993;82(6):1820–8. https://doi.org/10.1182/blood.V82.6.1820.1820 PMid: 8104532.
[22] McCarthy BA, Boyle E, Wang XP, et al. Surface expression of Bcl-2 in chronic lymphocytic leukemia and other B-cell leukemias and lymphomas without a breakpoint t(14; 18). Mol Med 2008;14(9–10):618–27. https://doi.org/10.2119/2008.00061.McCarthy PMid: 18633450.
[23] Alsagaby SA, Alhumaydhi FA. Proteomics insights into the pathology and prognosis of chronic lymphocytic leukemia. Saudi Med J 2019;40(4):317–27. https://doi.org/10.15537/smj.2019.4.23598 PMid: 30957124.

[24] Chen X, Dong XS, Gao HY, et al. Suppression of HSP27 increases the anti-tumor effects of quercetin in human leukemia U937 cells. Mol Med Report 2016;13 (1):689–96. https://doi.org/10.3892/mmr.2015.4600 PMid: 26648539.
[25] Voss T, Ahorn H, Haberl P, et al. Correlation of clinical data with proteomics profiles in 24 patients with B-cell chronic lymphocytic leukemia. Int J Cancer 2001;91(2):180–6. https://doi.org/10.1002/1097-0215(200002)9999:9999<::AID-IJC1037>3.0.CO;2-J.
[26] Scielzo C, Bertilaccio MT, Simonetti G, et al. HS1 has a central role in the trafficking and homing of leukemic B cells. Blood 2010;116(18):3537–46. https://doi.org/10.1182/blood-2009-12-258814 PMid: 20530793.
[27] Perrot A, Pionneau C, Nadaud S, et al. A unique proteomic profile upon surface IgM ligation in unmutated chronic lymphocytic leukemia. Blood 2011;118(4): e1–e15. https://doi.org/10.1182/blood-2011-02-335125 PMid: 21602524.
[28] Gaudio E, Spizzo R, Paduano F, et al. Tcl1 interacts with Atm and enhances NF-κB activation in hematologic malignancies. Blood 2012;119(1):180–7. https://doi.org/10.1182/blood-2011-08-374561 PMid: 22065599.
[29] Bichi R, Shinton SA, Martin ES, et al. Human chronic lymphocytic leukemia modeled in mouse by targeted TCL1 expression. Proc Natl Acad Sci 2002;99 (10):6955–60. https://doi.org/10.1073/pnas.102181599 PMid: 12011454.
[30] Eagle GL, Zhuang J, Jenkins RE, et al. Total proteome analysis identifies migration defects as a major pathogenetic factor in immunoglobulin heavy chain variable region (ighv)-unmutated chronic lymphocytic leukemia. Mol Cell Proteomics 2015;14(4):933–45. https://doi.org/10.1074/mcp.M114.044479 PMid: 25645933.
[31] Herold T, Jurinovic V, Metzeler KH, et al. An eight-gene expression signature for the prediction of survival and time to treatment in chronic lymphocytic leukemia. Leukemia 2011;25(10):1639–45. https://doi.org/10.1038/leu.2011.125 PMid: 21625232.
[32] Chuang H-Y, Rassenti L, Salcedo M, et al. Subnetwork-based analysis of chronic lymphocytic leukemia identifies pathways that associate with disease progression. Blood 2012;120(13):2639–49. https://doi.org/10.1182/blood-2012-03-416461 PMid: 22837534.
[33] Babicki S, Arndt D, Marcu A, et al. Heatmapper: web-enabled heat mapping for all. Nucleic Acids Res 2016;44(W1):W147–53. https://doi.org/10.1093/nar/gkw419 PMid: 27190236.
[34] Croft D, Mundo AF, Haw R, et al. The Reactome pathway knowledgebase. Nucleic Acids Res 2014;42(D1):D472–7. https://doi.org/10.1093/nar/gkt1102 PMid: 24243840.
[35] Binns D, Dimmer E, Huntley R, et al. QuickGO: a web-based tool for Gene Ontology searching. Bioinformatics 2009;25(22):3045–6. https://doi.org/10.1093/bioinformatics/btp536 PMid: 19744993.
[36] Darwiche W, Gubler B, Marolleau J-P, et al. Chronic lymphocytic leukemia B-cell normal cellular counterpart: clues from a functional perspective. Front Immunol 2018;9:683. https://doi.org/10.3389/fimmu.2018.00683 PMid: 29670635.
[37] Dono M, Cerruti G, Zupo S. The CD5+ B-cell. Int J Biochem Cell Biol 2004;36 (11):2105–11. https://doi.org/10.1016/j.biocel.2004.05.017 PMid: 15313456.
[38] Alsagaby SA. Integration of proteomics and transcriptomics data sets identifies prognostic markers in chronic lymphocytic leukemia. Majmaah J Health Sci 2019;7(1):1–22. https://doi.org/10.5455/mjhs.2019.01.002.
[39] Alsagaby SA. Transcriptomics-based validation of the relatedness of heterogeneous nuclear ribonucleoproteins to chronic lymphocytic leukemia as potential biomarkers of the disease aggressiveness. Saudi Med J 2019;40 (4):328–38. https://doi.org/10.15537/smj.2019.4.23380 PMid: 30957125.
[40] Budczies J, Klauschen F, Sinn BV, et al. Cutoff Finder: a comprehensive and straightforward Web application enabling rapid biomarker cutoff optimization. PLoS One 2012;7(12). https://doi.org/10.1371/journal.pone.0051862 PMid: 23251644.
[41] Ponomarenko EA, Poverennaya EV, Ilgisonis EV, et al. The size of the human proteome: the width and depth. Int J Anal Chem 2016;2016:7436849. https://doi.org/10.1155/2016/7436849 PMid: 27298622.
[42] Yates JR, Ruse CI, Nakorchevsky A. Proteomics by mass spectrometry: approaches, advances, and applications. Annu Rev Biomed Eng 2009;11 (1):49–79. https://doi.org/10.1146/annurev-bioeng-061008-124934 PMid: 19400705.
[43] Drissi R, Dubois ML, Boisvert FM. Proteomics methods for subcellular proteome analysis. FEBS J 2013;280(22):5626–34. https://doi.org/10.1111/febs.12502 PMid: 24034475.
[44] Stasyk T, Huber LA. Zooming in: fractionation strategies in proteomics. Proteomics 2004;4(12):3704-16. https://doi.org/10.1002/pmic.200401048 PMid: 15540207.
[45] Alsagaby SA. Understanding the fundamentals of proteomics. Curr Top Pept Protein Res 2019;20(3):25–33. https://doi.org/10.31300/CTPPR.20.2019.25-33.
[46] Wise JF, Berkova Z, Mathur R, et al. Nucleolin inhibits Fas ligand binding and suppresses Fas-mediated apoptosis in vivo via a surface nucleolin-Fas complex. Blood 2013;121(23):4729–39. https://doi.org/10.1182/blood-2012-12-471094 PMid: 23599269.
[47] Otake Y, Soundararajan S, Sengupta TK, et al. Overexpression of nucleolin in chronic lymphocytic leukemia cells induces stabilization of bcl2 mRNA. Blood 2007;109(7):3069–75. https://doi.org/10.1182/blood-2006-08-043257.
[48] Chen C-M, Chiang S-Y, Yeh N-H. Increased stability of nucleolin in proliferating cells by inhibition of its self-cleaving activity. J Biol Chem 1991;266 (12):7754–8. https://doi.org/10.1016/S0021-9258(20)89514-3.
[49] Messmer BT, Messmer D, Allen SL, et al. In vivo measurements document the dynamic cellular kinetics of chronic lymphocytic leukemia B cells. J Clin Invest 2005;115(3):755–64. https://doi.org/10.1172/JCI23409 PMid: 15711642.
[50] Johnston HE, Carter MJ, Larrayoz M, et al. Proteomics profiling of CLL versus healthy B-cells identifies putative therapeutic targets and a subtype-independent signature of spliceosome dysregulation. Mol Cell Proteomics 2018;17(4):776–91. https://doi.org/10.1074/mcp.RA117.000539 PMid: 29367434.
[51] Bracken CP, Wall SJ, Barré B, et al. Regulation of cyclin D1 RNA stability by SNIP1. Cancer Res 2008;68(18):7621–8. https://doi.org/10.1158/0008-5472.CAN-08-1217 PMid: 18794151.
[52] Abboudi Z, Patel K, Naresh KN. Cyclin D1 expression in typical chronic lymphocytic leukaemia. Eur J Haematol 2009;83(3):203–7. https://doi.org/10.1111/j.1600-0609.2009.01276.x PMid: 19467018.
[53] Damle RN, Batliwalla FM, Ghiotto F, et al. Telomere length and telomerase activity delineate distinctive replicative features of the B-CLL subgroups defined by immunoglobulin V gene mutations. Blood 2004;103(2):375–82. https://doi.org/10.1182/blood-2003-04-1345 PMid: 14504108.
[54] Mailand N, Lukas C, Kaiser BK, et al. Deregulated human Cdc14A phosphatase disrupts centrosome separation and chromosome segregation. Nat Cell Biol 2002;4(4):318–22. https://doi.org/10.1038/ncb777 PMid: 11901424.
[55] Kong X-x, Lv Y-r, Shao L-p, et al. HBx-induced MiR-1269b in NF-κB dependent manner upregulates cell division cycle 40 homolog (CDC40) to promote proliferation and migration in hepatoma cells. J Transl Med 2016;14(1):189. https://doi.org/10.1186/s12967-016-0949-y PMid: 27349221.
[56] Grill B, Wilson GM, Zhang K-X, et al. Activation/division of lymphocytes results in increased levels of cytoplasmic activation/proliferation-associated protein-1: prototype of a new family of proteins. J Immunol 2004;172(4):2389–400. https://doi.org/10.4049/jimmunol.172.4.2389 PMid: 14764709.
[57] Saffin J-M, Venoux M, Prigent C, et al. ASAP, a human microtubule-associated protein required for bipolar spindle assembly and cytokinesis. Proc Natl Acad Sci 2005;102(32):11302–7. https://doi.org/10.1073/pnas.0500964102 PMid: 16049101.
[58] Touat M, Ileana E, Postel-Vinay S, et al. Targeting FGFR signaling in cancer. Clin Cancer Res 2015;21(12):2684–94. https://doi.org/10.1158/1078-0432.CCR-14-2329 PMid: 26078430.
[59] Song Z, Lu P, Furman RR, et al. Activities of SYK and PLCγ2 predict apoptotic response of CLL cells to SRC tyrosine kinase inhibitor dasatinib. Clin Cancer Res 2010;16(2):587–99. https://doi.org/10.1158/1078-0432.CCR-09-1519 PMid: 20068106.
[60] Zhou J-D, Yao D-M, Li X-X, et al. KRAS overexpression independent of RAS mutations confers an adverse prognosis in cytogenetically normal acute myeloid leukemia. Oncotarget 2017;8(39):66087. https://doi.org/10.18632/oncotarget.19798 PMid: 29029494.
[61] Jamalpour M, Li X, Cavelier L, et al. Tumor SHB gene expression affects disease characteristics in human acute myeloid leukemia. Tumor Biol 2017;39(10). https://doi.org/10.1177/1010428317720643 PMid: 28982308.
[62] Dutta B, Ren Y, Hao P, et al. Profiling of the chromatin-associated proteome identifies hp1bp3 as a novel regulator of cell cycle progression. Mol Cell Proteomics 2014;13(9):2183–97. https://doi.org/10.1074/mcp.M113.034975 PMid: 24830416.
[63] Dutta B, Yan R, Lim SK, et al. Quantitative profiling of chromatome dynamics reveals a novel role for HP1BP3 in hypoxia-induced oncogenesis. Mol Cell Proteomics 2014;13(12):3236–49. https://doi.org/10.1074/mcp.M114.038232 PMid: 25100860.
[64] Pepper C, Hewamana S, Brennan P, et al. NF-κB as a prognostic marker and therapeutic target in chronic lymphocytic leukemia. Future Oncol 2009;5 (7):1027–37. https://doi.org/10.2217/fon.09.72 PMid: 19792971.