Original Article

Integration of Proteomics and Transcriptomics Data Sets Identifies Prognostic Markers in Chronic Lymphocytic Leukemia

Suliman A. Alsagaby [1]

[1]Department of Medical Laboratories, College of Applied Medical Sciences, Majmaah University

Received on 2.10.2018; accepted on 2.12.2018

Corresponding author: Suliman A. Alsagaby, Department of Medical Laboratories, College of Applied Medical Sciences, Majmaah University, PO Box 1712, Majmaah 11932, Saudi Arabia.
Email: s.alsaqaby@mu.edu.sa

Abstract

Background and aim: Chronic lymphocytic leukemia (CLL) is a malignant disease of B-cells that is characterized by variable prognosis. This study aims to search for novel prognostic markers in CLL. **Methods:** Six publicly available data sets of omics (proteomics and/or transcriptomics) that were generated from CLL cells, normal B-cells and normal peripheral blood mononuclear cell (PBMCs) were integrated to perform a crude search for CLL-associated proteins. Next, two publicly available data sets of CLL transcriptomics based on two independent cohorts (130 and 107 CLL patients) with available clinical information were utilized to assess the relevance of the suggested CLL-associated proteins at the level of transcription (mRNA) to CLL prognosis. Finally, Pearson score (PS) was employed to conduct correlation analyses of gene expression. **Results:** Sixty candidates were suggested as CLL-associated proteins, from which 19 proteins (32%) were previously implicated in cancer, and over-represented CLL-related biological processes. Furthermore, amongst the 60 proteins, nine (15%) had transcripts that significantly predicted early therapy ($p = 0.0001$ to 0.05) and/or short survival in CLL patients ($p = 0.0002$ to 0.05). Four of these markers showed a prognostic significance in the two independent data sets of CLL transcriptomics, increasing their validity as predictors of CLL prognosis. Finally, genes of pathways known to contribute to the aggressiveness of CLL were significantly correlated with FCRLA ($PS \geq 0.50$; $p < 0.00001$), which is a currently reported marker of poor prognosis. **Conclusion:** This study shows the usefulness of integrating publicly available omics data sets to identify potential biomarkers.

Key words: Leukemia, CLL; Prognostic Markers, Overall Survival, Time of Therapy; Proteomics; Transcriptomics.

الملخص

الخلفية: سرطان الدم اللمفاوي المزمن (س.د.ل.م) مرض خبيث يصيب الخلايا اللمفاوية البائية ويتصف بنتائج سريرية مختلفة بشكل كبير بين المرضى المصابين. الهدف: تهدف هذه الدراسة للبحث عن علامات إنذار حيوية جديدة في المرض (س.د.ل.م). **طريقة البحث:** تم دمج ستة مجموعات بيانية مستقلة للبروتيومكس والترانسكربتومكس للخلايا السرطانية والخلايا الطبيعية للبحث عن بروتينات سرطانية. ثم بعد ذلك تم استخدام بيانات ترانسكربتومكس للخلايا السرطانية (س.د.ل.م) من مجموعتين من المرضى (١٣٠ مريض و١٠٧ مريض) وذلك لتقصي دور البروتينات السرطانية في إنذار السرطان (س.د.ل.م). وأخيرا تم توظيف مقياس بيرسون لدراسة الترافق في التعبير الجيني. **النتائج:** كشفت الدراسة النقاب عن ستين بروتين سرطاني منهم ١٩ سبق أن تم ربطهم بأمراض خبيثة مختلفة بدراسات سابقة. ايضا أظهرت تسعة من هذه بروتين السرطانية القدرة على التنبؤ بمسار المرض وبالتالي يمكن اقتراحها كعلامات حيوية يمكن أن تخدم في إنذار السرطان (س.د.ل.م). كما دلت تحاليل الترافق الجيني ترافق جينات في مسارات إشارية مهمة في شراسة السرطان (س.د.ل.م) مع أحد علامات الإنذار التي كشفتها الدراسة الحالية. **الخلاصة:** توضح هذه الدراسة فائدة دمج بيانات الأومكس السرطانية والطبيعية لإكتشاف علامات انذار سرطانية جديدة

Introduction

Chronic lymphocytic leukemia (CLL) is a malignant disease that affects B-cells and is characterized by lymphocytosis of mature-looking B-cells (>5000 cells/µl) that accumulate in peripheral blood, lymph nodes, bone marrow and the spleen [1]. The cause of CLL is still unknown, yet some hereditary factors have been associated with the risk of developing CLL [2]. While important advances in CLL therapy have been accomplished, the current treatment does not eradicate CLL but reduces the disease burden and improves the overall survival time of afflicted patients [3].

CLL is a heterogeneous disease with variable clinical outcomes. While some patients have an indolent form of CLL, others exhibit an aggressive form of the disease with an early need for therapy and short overall survival time [1]. Different prognostic markers have been widely implemented to predict the prognosis of CLL [4]. For instance, deletions in 11q and 17p; unmutated immunoglobulin heavy variable genes (IGHV; UM-CLL); and over-expression of CD38 and zeta-chain-associated protein kinase 70 (ZAP-70) are markers of poor prognosis [5-8]. In contrast, deletions in 13q and mutated IGHV (M-CLL); and low-expression of CD38 and ZAP70 are predictors of good prognosis [5-8]. Although the above prognostic markers have been well established and widely applied, providing an accurate prognosis of CLL is still challenging [9]. Therefore, the search for novel prognostic markers that may improve the prognostication of the disease is needed.

Cancer-associated genes and/or proteins can be a rich source to facilitate the search for prognostic markers. In the context of CLL, various molecules that are reported to possess prognostic value for the disease exhibit preferential expression in CLL cells as compared with normal B-cells [4]. A typical example is ZAP-70, which is not expressed in normal B-cells but highly expressed in CLL cells from a subset of patients who are characterized with inferior prognosis [8]. In addition, apoptosis regulator Bcl-2, which is an anti-apoptotic protein, is significantly over-expressed in CLL cells compared with normal B-cells and predicts poor prognosis of the disease [10]. In contrast to normal B-cells, CLL cells express elevated levels of C-X-C chemokine receptor type 4 (CXCR4), which is a poor prognostic marker in the disease [11]. Overall, these findings indicate that the search for cancer-associated genes and/or proteins can be an early step to facilitate the identification of prognostic markers.

The goal of the present study was to integrate proteomics and transcriptomics data sets to identify novel prognostic markers in CLL. As mentioned earlier, the search for disease-associated proteins may facilitate the discovery of prognostic markers. Consequently, this study integrated six publicly available data sets of proteomics and transcriptomics to conduct a crude search for CLL-associated proteins. Next, two publicly available data sets of CLL transcriptomics were used to evaluate the prognostic value of the cognate transcripts of the proposed CLL-associated

proteins in CLL patients. Collectively, nine transcripts were currently reported to significantly predict the prognosis of CLL.

Methods

Proteomics data sets

Five proteomics data sets were used for the crude search for CLL-associated proteins. These data sets include three CLL proteomics data sets [12-14], one proteomics data set of normal B-cells [15] and one proteomics data set of normal PBMCs [16]. The proteomics data sets were selected because they were generated from primary cells (i.e. clinical CLL samples and normal cells from healthy donors) using the bottom-up proteomics approach and liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) workflow. Further details about the five proteomics data sets are summarized in Supplementary Table 1. The proteomics data sets were obtained from the supporting information of the corresponding studies [12-16]. The data sets PXD001512-16 [14], PXD001657 [15] and PXD001415–23 [16] are also available at the ProteomeXchange Consortium http://proteomecentral.proteomexchange.org via the PRIDE partner repository https://www.ebi.ac.uk/pride/archive/. To insure a proper comparison between the proteomics data sets, the protein identifiers (protein-specific identifier) in the different data sets were cross-referenced with the UniProt entry identifiers (release: 2016_08) using the retrieve/ID mapping tool http://www.uniprot.org/uploadlists/ [17].

Transcriptomics data sets

For the analyses performed to crudely search for CLL-associated proteins, a transcriptomics data set of normal B-cells (Series GEO accession number: GSE10325; n= 6 samples) [18] was used. This data set was selected because it was generated from primary cells (normal B-cells of healthy donors) using Affymetrix Human Genome U133A Array. Additionally, two independent data sets of CLL transcriptomics [19, 20] were used for the analyses aimed to assess the impact of the suggested CLL-associated proteins (at the level of transcription; mRNA) on CLL prognosis. Three criteria led to the selection of these two data sets. First, they included prognostic information for the individual patients whose samples were studied; the data set GSE39671 [19] contained data regarding the time to first treatment (TTFT), which measures the time between diagnosis and initial therapy; and the data set GSE22762 [20] included information about the overall survival (OS), which indicates the time between diagnosis and death. Second, the two data sets were based on two independent patients cohort consisting of more than 100 patients each (GSE39671 = 130 patients, GSE22762 = 107 patients). Third, both data sets were generated using the same oligonucleotide microarray platform (Affymetrix Human Genome U133 Plus 2.0 Array).

All the transcriptomics data sets were obtained from Gene Expression Omnibus (GEO) http://www.ncbi.nlm.nih.gov/geo/

[21] The DataSet SOFT files of the transcriptomics data sets were downloaded from GEO. Then, the ID references (probe IDs) of the Affymetrix Human Genome U133A Array and Affymetrix Human Genome U133 Plus 2.0 Array were cross-referenced with the UniProt entry identifiers (release: 2016_08) using g:Profiler http://biit.cs.ut.ee/gprofiler/gconvert.cgi and the retrieve/ID mapping tool http://www.uniprot.org/uploadlists/ [17, 22].

Integration of the data sets and statistical analysis

The UniProt entry identifiers corresponding to the proteomics and transcriptomics data sets and Venn diagram tool http://bioinfogp.cnb.csic.es/tools/venny_old/venny.php [23] were used for the integration analysis that aimed to search for CLL-associated proteins. The correlation analysis using Pearson score was performed using Excel software. Kaplan-Meier curves of TTFT and OS were created using Prism Graphpad software;
p values and hazard ratios (HR) were calculated using the log rank test. The probability scores of the pathway enrichment analysis was calculated using Reactome pathway knowledgebase [24]. The heatmap visualization of the correlation analyses was conducted using the heatmapper web-based tool http://www1.heatmapper.ca/ [25].

Gene ontology annotations

To gain insight into the biological processes assigned to the proteins of interest, Quick Gene Ontology tool (Quick GO) https://www.ebi.ac.uk/QuickGO/annotations was employed [26]. The analysis was limited to gene ontology (GO) of Homo sapiens and only GO terms assigned to the aspect "Biological Process" was searched. Given that different proteins have different biological processes, the present study reported only biological processes that were over-represented by at least 4% of the proteins of interest.

Pathway enrichment analysis

Reactome, which is a curated pathway database https://reactome.org/, was used to perform pathway enrichment analysis of transcripts of interest [24]. The analysis was conducted using the "Analyze Data" tool and was limited to human-specific pathways. Reactome reports enriched pathways with p value, which indicates the probability of a pathway being identified by chance [24]. In the present study, only pathways that were significantly enriched ($p \leq 0.05$) were reported.

Results

Cancer-associated proteins can be a rich source for the search for cancer prognostic markers. Five proteomics data sets were used for the crude search for CLL-associated proteins; specifically, there were three CLL proteomics data sets (696 proteins [12], 728 proteins [13] and 3521 proteins [14]), one proteomics data set of normal B-cells (3029 proteins [15]) and one proteomics data set of normal PBMCs (6885 proteins [16]). The three proteomics data sets of CLL were combined to give a total CLL proteome consisting of 3615 distinct proteins. In addition, the pro-

teomics data sets of normal B-cells and normal PBMCs were combined to give a total proteome of the normal cells (control cells) containing 9086 distinct proteins. Next, the proteome of the CLL cells was compared with the proteome of the normal cells. From the CLL proteome, 3479 proteins (96%) were common to the proteome of the normal cells, while 136 proteins (4%) appeared only in the proteome of CLL cells (Fig 1A). To increase the stringency of the search for CLL-associated proteins, the proteins that were specific to CLL proteome (136 proteins) were further compared with a transcriptomics data set of normal B-cells (7976 distinct transcripts; GEO accession number: GSE10325 [18]) to eliminate proteins with a positive transcript expression in normal B-cells. Consequently, the current study reported 60 proteins (2% of CLL proteome) that were evident only in the proteome of CLL cells and had no detectable mRNA in the transcriptome of normal B-cells (Fig 1B). These proteins were suggested as CLL-associated proteins and were used for the search for CLL prognostic markers (Supplementary Table 2).

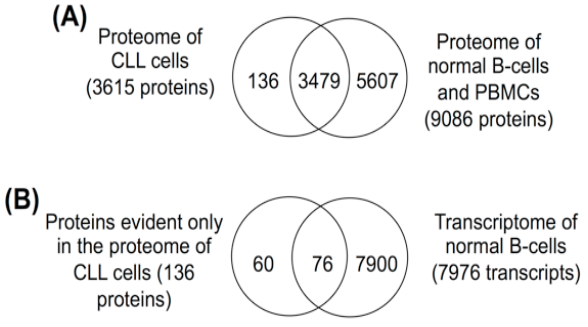

Fig.1 Integration of omics data sets. Proteomics and/or transcriptomics that were generated from CLL cells, normal B-cells and (PBMCs) were integrated for a crude identification of CLL-associated proteins

Existence of a relationship between the suggested CLL-associated proteins and cancer may indicate validity of the integration method applied in the present study. Thus, the literature was searched for the 60 proteins in the context of cancer. Interestingly, 19 proteins (32%) were previously shown to associate with cancer as opposed to normal tissues and/or implicated in the pathology and poor prognosis of different malignancies, including CLL (Supplementary Table 3). The Quick GO tool was also used to provide an insight into the biological processes of the proposed CLL-associated proteins (Table 1).

Table 1: Biological processes of the suggested CLL associated proteins

Biological process	Percentage of proteins
Signal transduction	12
Transcription, DNA-templated	12
Oxidation-reduction process	6
Protein ubiquitination	6
Phosphorylation	6
Kinase activity	6
Nucleosome assembly	6
Regulation of small GTPase mediated signal transduction	4
Positive regulation of MAPK cascade	4
Positive regulation of I-kappaB kinase/NF-kappaB (NF-κB) signaling	4
Positive regulation of B-cell proliferation	4
Inflammatory response	4
mRNA processing	4

Only biological processes reported for 4% or more of the proposed CLL-associated proteins are shown in the table. MAPK: mitogen-activated protein kinase.

The next goal was to use the proposed CLL-associated proteins for the search for CLL prognostic markers. To do that, two independent data sets of CLL transcriptomics were used (GEO accession numbers: GSE39671 [19] and GSE22762 [20]). The transcriptomics data set GSE39671 contained data regarding TTFT (number of patients = 130), while the transcriptomics data set GSE22762 included information concerning OS (number of patients = 107). Therefore, the two transcriptomics data sets were used independently to assess whether the transcripts corresponding to the suggested CLL-associated proteins predict TTFT and OS in CLL patients. The median level of transcripts corresponding to the proposed CLL-associated proteins was used to divide the patients into two groups: a low-expression group (patients with transcript expression smaller than the median) and a high-expression group (patients with transcript expression greater than the median). This was conducted for each one of the transcripts of interest in the two transcriptomics data sets. Next, Kaplan–Meier curves were constructed to compare TTFT and OS in the low expression and high expression groups of patients. Interestingly, the analyses showed that the cognate transcripts of 6 proteins and 7 proteins were significantly predictive of early therapy and short OS, respectively (Figs 2 and 3). Out of the six proteins whose transcripts indicated early therapy in the data set GSE39671 [19], four had transcripts that also predicted short OS in the data set GSE22762 [20], adding more validity to the reported findings.

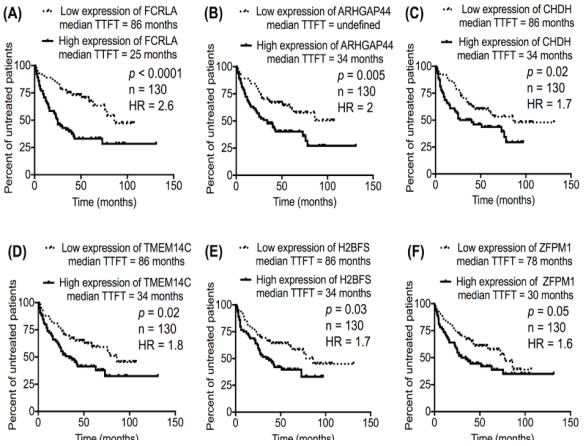

Fig. 2 Six of the CLL-associated proteins had transcripts significantly predictive of early therapy in CLL patients. Low expression: patients with a transcript expression smaller than the median; high expression: patients with a transcript expression higher than the median; TTFT: time to first treatment; HR: hazard ratio of high expression versus low expression.

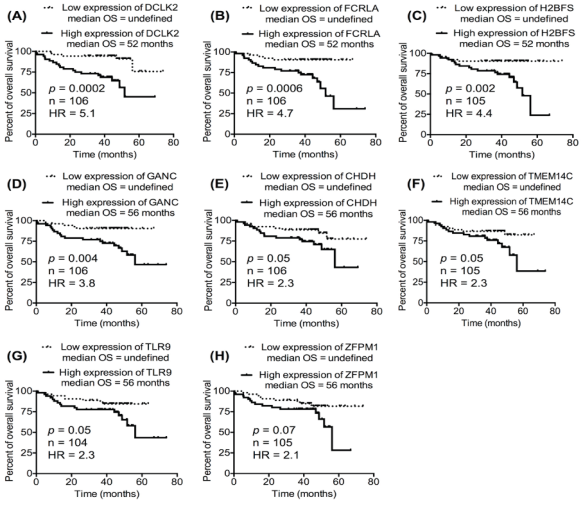

Fig. 3 Seven of the CLL-associated proteins had transcripts significantly indicative of short OS in CLL patients. Low expression: patients with a transcript expression below the median; high expression: patients with a transcript expression above the median; OS: overall survival time;
HR: hazard ratio of high expression versus low expression.

Next, the present study attempted to explain the association of FCRLA with high-risk CLL. The Pearson score (PS; Pearson correlation coefficient) was applied using Excel software to the data set of CLL transcriptomics (GSE39671; 130 patients [19]) in order to identify genes that correlate with the expression of FCRLA. Next, genes that exhibited correlation with FCRLA (PS ≥ 0.50; $p < 0.00001$) were subjected to a pathway enrichment analysis using Reactome. Interestingly, the analysis showed a significant enrichment ($p < 0.05$) for CLL-related pathways (Table 2). Heatmap-based visualization of the significant correlation between genes assigned to CLL-related pathways and FCRLA in the 130 CLL patients is shown in Fig 4.

Table 2: Significant enrichment of CLL-related pathways by the genes that correlated with the expression of FCRLA

Name of enriched pathway	*p* value
S Phase (cell cycle)	0.0008
Extension of Telomeres	0.002
Downstream signaling events of BCR.	0.007
Telomere lagging strand synthesis	0.012
Activation of NF-κB in B cells	0.015
Cellular response to hypoxia	0.038

BCR: B-cell receptor; NF-κB: nuclear factor-κB

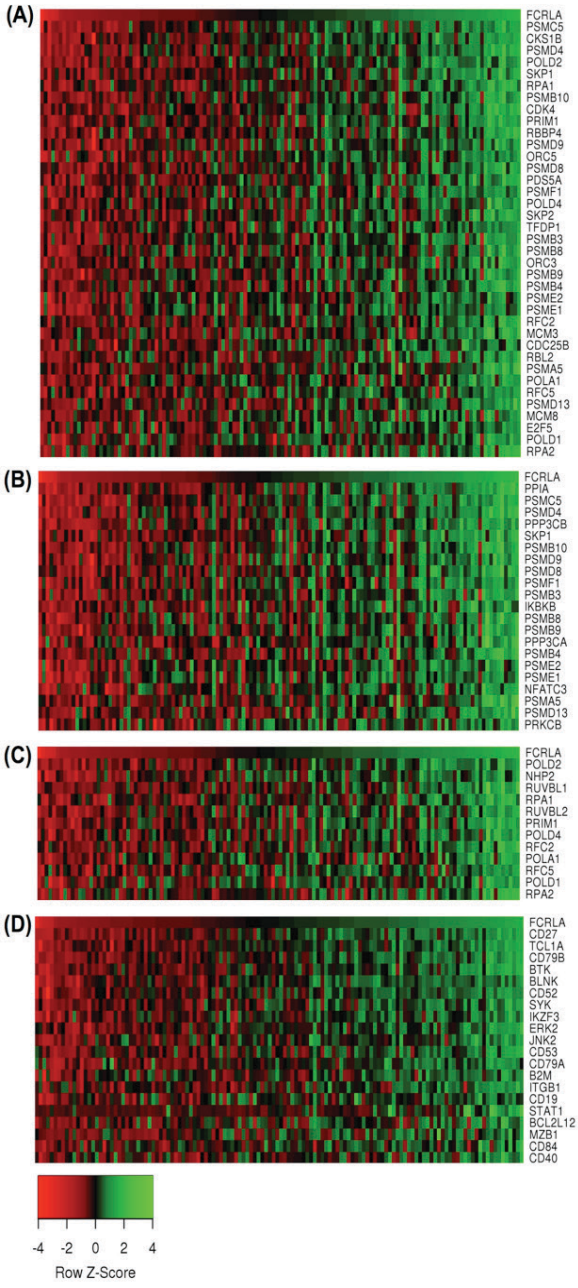

Fig. 4 Heatmap presentation of the correlation between genes assigned to CLL-related pathways and FCRLA. Excel software was used to sort the 130 CLL patients (from the data set GSE39671) horizontally based on the ascending expression FCRLA (from lowest expression to highest expression). Then, the genes that significantly enriched S phase (A), downstream signaling events of B-cell receptor (BCR) (B) and extension of telomeres (C); and other CLL-related genes (D) in the 130 CLL patients were sorted vertically according to their Pearson scores, with FCRLA being at the top of each list. Next, heatmapper web-based tool was used to construct a heatmap graphic based on the expression of the genes of interest and FCRLA in the 130 patients.

Discussion

The integration method of the omics data sets that was applied in the present study aimed for a rough identification of CLL-associated proteins from the massive CLL proteomics data [12-14]. This method targets proteins that were detected (present) in the CLL proteome but were not detected (absent) in the proteome of the normal cells (control cells). However, this method is not sensitive for the identification of differentially expressed proteins in CLL cells compared with the normal cells. Therefore, such proteins might have been missed because they were marked as "present proteins" in the proteomes of the malignant and the control cells. Although the proteome of the malignant cells was mainly derived from CLL cells, approximately 10% of the cellular population was PBMCs [12-14]. Therefore, the proteome of normal PBMCs [16] was included in the proteome of the control cells.

Direct comparison between independent proteomics data sets could misidentify disease-irrelevant proteins as proteins of interest. Therefore, the present study applied three actions in an attempt to avoid false identification. First, only proteomics data sets generated using the bottom-up approach and LC-MS/MS workflow were included to minimize technical variations. Second, the proteome coverage of the normal cells was larger than that of the CLL cells (9086 proteins compared with 3615 proteins, respectively). Therefore, this reduced the false identification of random proteins as CLL-specific proteins due to limited proteome coverage in the normal cells. Finally, for a candidate to be suggested as a CLL-associated protein, it had to be evident only in the CLL proteome and undetectable in the transcriptome of normal B-cells. Consequently, only 2% (60 proteins) of the CLL proteome was suggested as CLL-associated proteins, of which 32% (19 proteins) were previously linked to different neoplasms, including CLL (summarized in Supplementary Table 3), and 15% (9 proteins) had transcript expression that was currently shown to significantly predict the prognosis of CLL. In addition, the proposed CLL-associated proteins over-represented biological processes that were implicated in the progression of CLL, such as positive regulation of MAPK cascade, signaling of NF-κB and B-cell proliferation [4].

Nine prognostic markers were currently reported, of which five were previously linked to CLL. For example, an increased expression of TLR9 significantly predicted short OS in the present study supporting a previous report that showed an elevated expression of TLR9 in CLL cells compared with normal B-cells and a higher expression of TLR9 in patients with progressive CLL compared with those having a stable form of the disease [27]. In addition, an up-regulated expression of DCLK2 was described in the present work as a marker of short OS. Consistently, an independent transcriptomics study found an increased expression of DCLK2 in CLL samples from high-risk patients (UM-CLL and ZAP-70 positive) compared with low risk group (M-

CLL and ZAP-70 negative) [28]. The present study also reported a significant prediction of early therapy and/or short OS by an elevated expression of FCRLA, ARHGAP44 and CHDH. Interestingly, another independent transcriptomics study reported high expression of these three markers as risk factors of developing CLL, suggesting an involvement of these markers in the initiation of the disease [29]. Furthermore, the protein products of DCLK2, FCRLA and CHDH were also shown in the CLL proteomics data set [14], which was used in the present study, to be overexpressed in patients with high-risk CLL (UM-CLL) compared with patients having low-risk CLL (M-CLL).

Among the reported prognostic markers, increased expression of FCRLA was the strongest predictor of early therapy and second most accurate indicator of short OS in CLL patients using the two independent CLL transcriptomics data sets. These findings seem to fit with the role of FCRLA in B-cells. FCRLA was implicated in the activation and proliferation of B-cells; elevated expression of FCRLA was shown to mark activated B-cells and the highest level of FCRLA expression was reported in B-cells that reside in the proliferation centers of lymphoid tissues [30, 31]. Furthermore, FCRLA interacts with intracellular Ig and is required for the proper assembly of IgG and IgM [32, 33]. In consistence with the role of FCRLA in B-cells, the correlation analysis reported in the current study demonstrated a significant enrichment of CLL-related pathways; such as S phase (cell cycle), downstream signaling events of B-cell receptor (BCR), extension of telomeres and activation of NF-ϰB, by the genes that significantly correlated with the expression of FCRLA in the 130 CLL patients [34-37]. In addition, genes that are known to derive the proliferation and/or survival of CLL cells, in particular those that participate in the signal transduction of BCR like CD79A, CD79B, BLNK, BTK and SYK; and other genes, such as TCL1A, CD40, IKZF3 and CD27, exhibited a significant correlation with the expression of FCRLA [38-44]. This may provide an insight into the significant prediction of the aggressive form of CLL by the elevated expression of FCRLA.

A number of points should be considered while viewing the results of the current report. First, the integration of the omics data sets aimed for a rough identification of candidates that could be proposed as CLL-associated proteins. However, to draw a definitive conclusion of the association of those proteins with CLL, either overexpression or specific expression of those proteins in CLL cells compared with normal B-cells has to be confirmed by a specific protein detection technique like western blotting. Second, the significant identification of the 9 prognostic markers was based on TTFT data of 130 patients (GSE39671) and OS information of 107 patients (GSE22762) whose transcriptomics data sets were published before the era of novel treatment of CLL. Therefore, the clinical usefulness of these markers in the context of the new modalities of CLL therapy merits investigation. Third, a comparison between

the present markers and those commonly applied in CLL, such as IGHV, ZAP-70, CD38 and the chromosomal aberrations, was not possible to conduct due to the unavailability of the common makers of the 237 patients. Consequently, studying the present markers in parallel with those frequently applied is necessary to assess whether the present markers add prognostic information to what can be known by the commonly used markers. Fourth, conventionally, real-time PCR is used to confirm findings generated using DNA microarray-based transcriptomics. Therefore, measuring the expression of the 9 markers using real-time PCR is worthwhile to validate the expression patterns of the 9 transcripts in CLL samples.

Conclusion

The integration method applied on the omics data sets suggested 60 candidates as CLL-associated proteins, of which 32% (19 proteins) were previously implicated in neoplasms, including CLL. Moreover, the increased transcript expression of 9 of the suggested CLL-associated proteins significantly predicted early therapy and/or short OS in CLL patients. Interestingly, 4 of the 9 transcripts retained prognostic significance in the two independent transcriptomics data sets (237 CLL patients). Therefore, they may have the potential to serve as prognostic markers in the disease. Importantly, further assessment of the reported prognostic markers (using real-time PCR) in parallel with those frequently applied in a cohort of patients who have been treated with the new modality of CLL therapy is worthwhile to validate the utility of the present markers in the prognosis of CLL.

References:

1. Dighiero G. CLL biology and prognosis. ASH Education Program Book 2005;2005:278-284.
2. Strefford JC. The genomic landscape of chronic lymphocytic leukaemia: biological and clinical implications. British journal of haematology 2015;169:14-31.
3. Hallek M. Chronic lymphocytic leukemia: 2017 update on diagnosis, risk stratification, and treatment. American journal of hematology 2017;92:946-965.
4. Alsagaby SA, Brennan P and Pepper C. Key Molecular Drivers of Chronic Lymphocytic Leukemia. Clinical Lymphoma Myeloma and Leukemia 2016;16:593-606.
5. Döhner H, Stilgenbauer S, Benner A, Leupolt E, Kröber A, Bullinger L, et al. Genomic aberrations and survival in chronic lymphocytic leukemia. New England Journal of Medicine 2000;343:1910-1916.
6. Hamblin TJ, Davis Z, Gardiner A, Oscier DG and Stevenson FK. Unmutated Ig VH genes are associated with a more aggressive form of chronic lymphocytic leukemia. Blood 1999;94:1848-1854.
7. Pepper C, Majid A, Lin TT, Hewamana S, Pratt G, Walewska R, et al. Defining the prognosis of early stage chronic lymphocytic leukaemia patients. British journal

of haematology 2012;156:499-507.

8. Rassenti LZ, Huynh L, Toy TL, Chen L, Keating MJ, Gribben JG, et al. ZAP-70 compared with immunoglobulin heavy-chain gene mutation status as a predictor of disease progression in chronic lymphocytic leukemia. New England Journal of Medicine 2004;351:893-901.
9. Chen C and Puvvada S. Prognostic Factors for Chronic Lymphocytic Leukemia. Current hematologic malignancy reports 2016;11:37-42.
10. Kitada S, Andersen J, Akar S, Zapata JM, Takayama S, Krajewski S, et al. Expression of apoptosis-regulating proteins in chronic lymphocytic leukemia: correlations with in vitro and in vivo chemoresponses. Blood 1998;91:3379-3389.
11. Ghobrial IM, Bone ND, Stenson MJ, Novak A, Hedin KE, Kay NE, et al. Expression of the Chemokine Receptors CXCR4 and CCR7 and Disease Progression in B-Cell Chronic Lymphocytic Leukemia/ Small Lymphocytic Lymphoma. Mayo Clinic Proceedings 2004;79:318-325.
12. Barnidge DR, Tschumper RC, Jelinek DF, Muddiman DC and Kay NE. Protein expression profiling of CLL B cells using replicate off-line strong cation exchange chromatography and LC–MS/MS. Journal of Chromatography B 2005;819:33-39.
13. Alsagaby SA, Khanna S, Hart KW, Pratt G, Fegan C, Pepper C, et al. Proteomics-based strategies to identify proteins relevant to chronic lymphocytic leukemia. Journal of proteome research 2014;13:5051-5062.
14. Eagle GL, Zhuang J, Jenkins RE, Till KJ, Jithesh PV, Lin K, et al. Total proteome analysis identifies migration defects as a major pathogenetic factor in immunoglobulin heavy chain variable region (ighv)-unmutated chronic lymphocytic leukemia. Molecular & Cellular Proteomics 2015;14:933-945.
15. Hoek KL, Samir P, Howard LM, Niu X, Prasad N, Galassie A, et al. A cell-based systems biology assessment of human blood to monitor immune responses after influenza vaccination. PloS one 2015;10:e0118528.
16. Bileck A, Kreutz D, Muqaku B, Slany A and Gerner C. Comprehensive assessment of proteins regulated by dexamethasone reveals novel effects in primary human peripheral blood mononuclear cells. Journal of proteome research 2014;13:5989-6000.
17. Pundir S, Martin MJ and O'Donovan C. UniProt tools. Current protocols in bioinformatics 2016;1.29. 1-1.29. 15.
18. Hutcheson J, Scatizzi JC, Siddiqui AM, Haines GK, Wu T, Li Q-Z, et al. Combined deficiency of proapoptotic regulators Bim and Fas results in the early onset of systemic autoimmunity. Immunity 2008;28:206-217.
19. Chuang H-Y, Rassenti L, Salcedo M, Licon K, Kohlmann A, Haferlach T, et al. Subnetwork-based analysis of chronic lymphocytic leukemia identifies path-

ways that associate with disease progression. Blood 2012;120:2639-2649.

20. Herold T, Jurinovic V, Metzeler KH, Boulesteix A-L, Bergmann M, Seiler T, et al. An eight-gene expression signature for the prediction of survival and time to treatment in chronic lymphocytic leukemia. Leukemia 2011;25:1639-1645.
21. Clough E and Barrett T. The gene expression omnibus database. Statistical Genomics: Methods and Protocols 2016;1418:93-110.
22. Reimand J, Arak T, Adler P, Kolberg L, Reisberg S, Peterson H, et al. g: Profiler—a web server for functional interpretation of gene lists (2016 update). Nucleic acids research 2016;44:W83–W89.
23. Oliveros JC. VENNY. An interactive tool for comparing lists with Venn Diagrams. h ttp. bioinfogp. cnb. csic. es/tools/venny/index. html 2007.
24. Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G, et al. The Reactome pathway knowledgebase. Nucleic acids research 2014;42:D472-D477.
25. Babicki S, Arndt D, Marcu A, Liang Y, Grant JR, Maciejewski A, et al. Heatmapper: web-enabled heat mapping for all. Nucleic acids research 2016;44:W147-W153.
26. Binns D, Dimmer E, Huntley R, Barrell D, O'Donovan C and Apweiler R. QuickGO: a web-based tool for Gene Ontology searching. Bioinformatics 2009;25:3045-3046.
27. Barcellini W, Imperiali FG, Zaninoni A, Reda G, Consonni D, Fattizzo B, et al. Toll-like receptor 4 and 9 expression in B-chronic lymphocytic leukemia: relationship with infections, autoimmunity and disease progression. Leukemia & lymphoma 2014;55:1768-1773.
28. Trojani A, Di Camillo B, Tedeschi A, Milena L, Montesano S, Ricci F, et al. Microarray Gene Expression Signature Indicates ARSD As a New Marker Associated with Igvh Mutational Status, ZAP-70 and Disease Progression in CLL. Blood 2011;118:3447.
29. Chadeau-Hyam M, Vermeulen RCH, Hebels D, Castagné R, Campanella G, Portengen L, et al. Prediagnostic transcriptomic markers of Chronic lymphocytic leukemia reveal perturbations 10 years before diagnosis. Annals of Oncology 2014;25:1065-1072.
30. Reshetnikova ES, Mechetina LV, Volkova OY, Guselnikov SV, Chikaev NA, Kövesdi D, et al. Differential expression of FCRLA in naive and activated mouse B cells. Cellular immunology 2012;272:182-192.
31. Capone M, Bryant JM, Sutkowski N and Haque A. Fc Receptor-Like Proteins in Pathophysiology of B-cell Disorder. Journal of clinical & cellular immunology 2016;7:427-436.
32. Wilson TJ, Gilfillan S and Colonna M. Fc receptor-like A associates with intracellular IgG and IgM but is dispensable for antigen-specific immune responses. The Journal of Immunology 2010;185:2960-

2967.

33. Santiago T, Kulemzin SV, Reshetnikova ES, Chikaev NA, Volkova OY, Mechetina LV, et al. FCRLA is a resident endoplasmic reticulum protein that associates with intracellular Igs, IgM, IgG and IgA. International immunology 2010;23:43-53.
34. Woyach JA, Johnson AJ and Byrd JC. The B-cell receptor signaling pathway as a therapeutic target in CLL. Blood 2012;120:1175-1184.
35. Pepper C, Hewamana S, Brennan P and Fegan C. NF-κB as a prognostic marker and therapeutic target in chronic lymphocytic leukemia. Future Oncology 2009;5:1027-1037.
36. Damle RN, Batliwalla FM, Ghiotto F, Valetto A, Albesiano E, Sison C, et al. Telomere length and telomerase activity delineate distinctive replicative features of the B-CLL subgroups defined by immunoglobulin V gene mutations. Blood 2004;103:375-382.
37. Messmer BT, Messmer D, Allen SL, Kolitz JE, Kudalkar P, Cesar D, et al. In vivo measurements document the dynamic cellular kinetics of chronic lymphocytic leukemia B cells. The Journal of clinical investigation 2005;115:755-764.
38. Grdisa M. Influence of CD40 ligation on survival and apoptosis of B-CLL cells in vitro. Leukemia research 2003;27:951-956.
39. Granziero L, Ghia P, Circosta P, Gottardi D, Strola G, Geuna M, et al. Survivin is expressed on CD40 stimulation and interfaces proliferation and apoptosis in B-cell chronic lymphocytic leukemia. Blood 2001;97:2777-2783.
40. Stevenson FK, Krysov S, Davies AJ, Steele AJ and Packham G. B-cell receptor signaling in chronic lymphocytic leukemia. Blood 2011;118:4313-4320.
41. Bichi R, Shinton SA, Martin ES, Koval A, Calin GA, Cesari R, et al. Human chronic lymphocytic leukemia modeled in mouse by targeted TCL1 expression. Proceedings of the National Academy of Sciences 2002;99:6955-6960.
42. Billot K, Soeur J, Chereau F, Arrouss I, Merle-Béral H, Huang M-E, et al. Deregulation of Aiolos expression in chronic lymphocytic leukemia is associated with epigenetic modifications. Blood 2011;117:1917-1927.
43. Kara IO, Sahin B and Gunesacar R. Expression of soluble CD27 and interleukins-8 and-10 in B-cell chronic lymphocytic leukemia: correlation with disease stage and prognosis. Advances in therapy 2007;24:29-40.
44. Poeta GD, Principe MID, Zucchetto A, Buccisano F, Simotti C, Maurillo L, et al. High CD79b Expression Predicts a Poor Outcome in B-Cell Chronic Lymphocytic Leukemia (B-CLL). Blood 2008;112:1054.

Supplementary Tables (1, 2 and 3)

Supplementary Table 1: Technical information about the proteomics data sets used in the present study

Source of sample	Number of samples	Type of protein extract	Mass label	Protein database	Mass spectrometry	Ref.
PBMCs isolated from CLL patient	1 sample from an untreated patient	Crude-membrane extract	N/A	Human SwissProt database	Q-TOF API US (Waters)	[1]
PBMCs isolated from CLL patients	12 samples from 12 patients with different prognoses, some of whom were treated.	Cytosolic and nuclear extracts	iTRAQ	Human SwissProt database	4800 MALDI-TOF/TOF (Applied Biosystems)	[2]
PBMCs isolated from CLL patients	18 samples from 18 patients with different prognoses, some of whom were treated.	Whole cell lysate	iTRAQ	Human SwissProt database	TripleTOF 5600 system (AB SCIEX)	[3]
Normal B-cells isolated from PBMCs of healthy donor	2 sample from 2 healthy donors	Whole cell lysate	iTRAQ	Human Ensembl protein database	LTQ-OrbitrapXL (Thermo Fisher Scientific)	[4]
Normal PBMCs isolated from healthy donors	3 samples from 3 healthy donors	Supernatant, cytosolic and nuclear extracts	N/A	Human SwissProt database	QExactive Orbitrap (Thermo Fisher Scientific)	[5]

Protein database indicates the database that was used in each study for the identification of proteins based on the MS and MS/MS spectra. Q-TOF: Quadrupole-time of flight mass spectrometry; MALDI-TOF/TOF: matrix assisted laser desorption/ionization-time of flight tandem mass spectrometry; TripleTOF: triple time of flight mass spectrometry; LTQ-OrbitrapXL: linear trap quadrupole-orbitrap XL mass spectrometry; iTRAQ: isobaric tags for relative and absolute quantitation; N/A indicates that mass label was not used in the corresponding study; PBMCs: peripheral blood mononuclear cell. PBMCs isolated from CLL patients were ≥90% CD5+ and CD19+ (CLL cells).

Supplementary Table 2: The proposed CLL-associated proteins.

UniProt identifier	Protein name
Q8NHW5	60S acidic ribosomal protein P0-like
Q9UKP5	A disintegrin and metalloproteinase with thrombospondin motifs 6 (ADAM-TS 6) (ADAM-TS6) (ADAMTS-6) (EC 3.4.24)
P49753	Acyl-coenzyme A thioesterase 2, mitochondrial (Acyl-CoA thioesterase 2) (EC 3.1.2.2) (Acyl-coenzyme A thioester hydrolase 2a) (CTE-Ia) (Long-chain acyl-CoA thioesterase 2) (ZAP128)
Q9HCE9	Anoctamin-8 (Transmembrane protein 16H)
P03928	ATP synthase protein 8 (A6L) (F-ATPase subunit 8)
Q7Z6A9	B- and T-lymphocyte attenuator (B- and T-lymphocyte-associated protein) (CD antigen CD272)
Q9BX70	BTB/POZ domain-containing protein 2
Q9Y4F5	Centrosomal protein of 170 kDa protein B (Centrosomal protein 170B) (Cep170B)
Q8NE62	Choline dehydrogenase, mitochondrial (CDH) (CHD) (EC 1.1.99.1)
Q9Y5K3	Choline-phosphate cytidylyltransferase B (EC 2.7.7.15) (CCT-beta) (CTP:phosphocholine cytidylyltransferase B) (CCT B) (CT B) (Phosphorylcholine transferase B)
Q8IWD4	Coiled-coil domain-containing protein 117
Q7LFL8	CXXC-type zinc finger protein 5 (CF5) (Putative MAPK-activating protein PM08) (Putative NF-kappa-B-activating protein 102) (Retinoid-inducible nuclear factor) (RINF)
P00414	Cytochrome c oxidase subunit 3 (Cytochrome c oxidase polypeptide III)
Q8NCM8	Cytoplasmic dynein 2 heavy chain 1 (Cytoplasmic dynein 2 heavy chain) (Dynein cytoplasmic heavy chain 2) (Dynein heavy chain 11) (hDHC11) (Dynein heavy chain isotype 1B)
P49619	Diacylglycerol kinase gamma (DAG kinase gamma) (EC 2.7.1.107) (Diglyceride kinase gamma) (DGK-gamma)
P0C6T2	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 4
Q7L513	Fc receptor-like A
Q96LA6	Fc receptor-like protein 1 (FcR-like protein 1) (FcRL1) (Fc receptor homolog 1) (FcRH1) (IFGP family protein 1) (hIFGP1) (Immune receptor translocation-associated (protein 5) (CD antigen CD307a
Q96P31	Fc receptor-like protein 3 (FcR-like protein 3) (FcRL3) (Fc receptor homolog 3) (FcRH3) (IFGP family protein 3) (hIFGP3) (Immune receptor translocation-associated protein 3) (SH2 domain-containing phosphatase anchor protein 2) (CD antigen (CD307c

Q6DKI2	(Galectin-9C (Gal-9C) (Galectin-9-like protein B
Q9UK08	(Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-8 (Gamma-9
P23527	(Histone H2B type 1-O (Histone H2B.2) (Histone H2B.n) (H2B/n
P57053	(Histone H2B type F-S (Histone H2B.s) (H2B/s
Q6PFW1	Inositol hexakisphosphate and diphosphoinositol-pentakisphosphate kinase 1 (EC 2.7.4.21) (EC 2.7.4.24) (Diphosphoinositol pentakisphosphate kinase 1) (Histidine acid phosphatase domain-containing protein 2A) (IP6 kinase) (Inositol pyrophosphate syn- (thase 1) (InsP6 and PP-IP5 kinase 1) (VIP1 homolog) (hsVIP1
Q9H293	(Interleukin-25 (IL-25) (Interleukin-17E) (IL-17E
Q9P2G3	Kelch-like protein 14 (Protein interactor of Torsin-1A) (Printor) (Protein interactor of (torsinA
O75525	KH domain-containing, RNA-binding, signal transduction-associated protein 3 (RNA- binding protein T-Star) (Sam68-like mammalian protein 2) (SLM-2) (Sam68-like phos- (photyrosine protein
Q6IPR1	LYR motif-containing protein 5
O15481	(Melanoma-associated antigen B4 (MAGE-B4 antigen
Q9UPN3	Microtubule-actin cross-linking factor 1, isoforms 1/2/3/5 (620 kDa actin-binding pro- (tein) (ABP620) (Actin cross-linking family protein 7) (Macrophin-1) (Trabeculin-alpha
Q08AG7	Mitotic-spindle organizing protein 1 (Mitotic-spindle organizing protein associated with a (ring of gamma-tubulin 1
Q7Z5P9	(Mucin-19 (MUC-19
Q8TET4	(Neutral alpha-glucosidase C (EC 3.2.1.20
Q86XR2	(Niban-like protein 2 (B-cell novel protein 1) (Protein FAM129C
Q9H0G5	Nuclear speckle splicing regulatory protein 1 (Coiled-coil domain-containing protein 55) ((Nuclear speckle-related protein 70) (NSrp70
Q13515	Phakinin (49 kDa cytoskeletal protein) (Beaded filament structural protein 2) (Lens fiber cell beaded filament protein CP 47) (CP47) (Lens fiber cell beaded filament protein CP (49) (CP49) (Lens intermediate filament-like light) (LIFL-L
A4D1U4	Protein LCHN
Q9NQ39	Putative 40S ribosomal protein S10-like
Q5VTE0	Putative elongation factor 1-alpha-like 3 (EF-1-alpha-like 3) (Eukaryotic elongation fac- tor 1 A-like 3) (eEF1A-like 3) (Eukaryotic translation elongation factor 1 alpha-1 pseu- (dogene 5
Q6DN03	(Putative histone H2B type 2-C (Histone H2B.t) (H2B/t
A8MVU1	Putative neutrophil cytosol factor 1C (NCF-1C) (Putative SH3 and PX domain-contain- (ing protein 1C

Q8NHM4	(Putative trypsin-6 (EC 3.4.21.4) (Serine protease 3 pseudogene 2) (Trypsinogen C
Q17R89	Rho GTPase-activating protein 44 (NPC-A-10) (Rho-type GTPase-activating protein (RICH2) (RhoGAP interacting with CIP4 homologs protein 2) (RICH-2
Q8N568	Serine/threonine-protein kinase DCLK2 (EC 2.7.11.1) (CaMK-like CREB regulatory kinase 2) (CL2) (CLICK-II) (CLICK2) (Doublecortin domain-containing protein 3B) ((Doublecortin-like and CAM kinase-like 2) (Doublecortin-like kinase 2
Q9P246	Stromal interaction molecule 2
Q6ZRP7	Sulfhydryl oxidase 2 (EC 1.8.3.2) (Neuroblastoma-derived sulfhydryl oxidase) (Qui-(escin Q6-like protein 1
Q86XK3	Swi5-dependent recombination DNA repair protein 1 homolog (Meiosis protein 5 homo-(log
Q8N103	T-cell activation Rho GTPase-activating protein (T-cell activation GTPase-activating (protein
Q8N4P2	(Tetratricopeptide repeat protein 30B (TPR repeat protein 30B
Q9NR96	(Toll-like receptor 9 (CD antigen CD289
Q9P0S9	Transmembrane protein 14C
Q6P9G4	Transmembrane protein 154
Q6ZMU5	(Tripartite motif-containing protein 72 (Mitsugumin-53) (Mg53
P09936	Ubiquitin carboxyl-terminal hydrolase isozyme L1 (UCH-L1) (Neuron cytoplasmic pro-(tein 9.5) (PGP 9.5) (PGP9.5) (Ubiquitin thioesterase L1
Q5SQH8	Uncharacterized protein C6orf136
Q9ULV0	Unconventional myosin-Vb
Q6VEQ5	WAS protein family homolog 2 (CXYorf1-like protein on chromosome 2) (Protein (FAM39B
Q6ZNA1	Zinc finger protein 836
Q96IR2	Zinc finger protein 845
Q8IX07	Zinc finger protein ZFPM1 (Friend of GATA protein 1) (FOG-1) (Friend of GATA 1) (Zinc (finger protein 89A) (Zinc finger protein multitype 1

The integration of the proteomics data sets of CLL cells, normal B-cells and normal PBMCs; as well as the transcriptomics data set of normal B-cells suggested 60 candidates as CLL-associated proteins. This table shows these proteins.

Supplementary Table 3: Reported role of the suggested CLL-associated proteins in cancer

UniProt identifier	Protein name (*gene name*)	Reported role in cancer
Q8N4P2	Tetratricopeptide repeat protein 30B (*TTC30B*)	• Promotes survival of acute myeloid leukemia (AML) [6].
Q9HCE9	Anoctamin-8 (*ANO8*)	• Enhances the growth and metastasis of gastrointestinal tumor and head/neck squamous cell carcinoma [7].
Q7LFL8	CXXC-type zinc finger protein 5 (*CXXC5*), also known as RINF	• Promotes survival of AML cells and associates with poor prognosis of the disease [8]. • In comparison with normal tissues, breast cancer and melanoma cells exhibit significantly increased expression of CF5 [9]. • Associates with poor prognosis of breast cancer [9].
P49619	Diacylglycerol kinase gamma (*DGKG*)	• Promotes survival, migration and invasion of breast cancer cells and colon cancer cells [10, 11].
Q6DKI2	Galectin-9C (*LGALS9C*)	• Associates with lymphoma, leukemia and colon cancer cells as compared with normal tissue [12, 13].
Q9H0G5	Nuclear speckle splicing regulatory protein 1 (*NSRP1*), also known as NSRP70	• Associates with acute lymphoblastic leukemia (ALL) and considered a good diagnostic marker of the disease [14]. • Associates with poor prognosis of ALL and AML [14].
Q7L513	Fc receptor-like A (*FCRLA*)	• Associates with the risk of developing CLL [15].
Q17R89	Rho GTPase activating protein 44 (*ARHGAP44*)	• Associates with the risk of developing CLL [15].
Q8NE62	Choline dehydrogenase (*CHDH*)	• Associates with the risk of developing CLL [15].
Q9P0S9	Transmembrane protein 14C (*TMEM14C*)	• Over-expressed in ovarian cancer compared with normal tissue and required for tumor growth and invasion of the malignant cells [16]. • Suppresses the pro-apoptotic protein BAX and protects glioblastoma tumor cells from apoptosis [17].

Q9NR96	Toll-like receptor 9 (*TLR9*)	• Highly expressed on CLL compared with normal B-cells and associates with high-risk CLL [18, 19].
P09936	Ubiquitin carboxyl-terminal hydrolase isozyme L1 (*UCHL1*)	• Promotes proliferation, migration and metastasis of colorectal cancer cells [20]. • Induces metastasis of prostate cancer cells [21]. • Promotes the development of B cell lymphoma [22].
Q7Z6A9	B- and T-lymphocyte attenuator (*BTLA*)	• Preferentially expressed in gastric carcinoma compared with the normal tissue and associates with high-risk of the disease [23]. • Strongly associates with CLL cells compared with normal B-cells [24].
Q8NCM8	Cytoplasmic dynein 2 heavy chain 1 (*DYNC2H1*)	• Associates with glioblastoma as compared with the normal tissue [25].
Q96LA6	Fc receptor-like 1 (*FCRL1*)	• Contributes to the progression of melanoma [26].
Q96P31	Fc receptor-like 3 (*FCRL3*)	• Associates with poor prognosis of T cell lymphoma [27].
O75525	RNA-binding protein T-Star (*KHDRBS3*), also known as SLM-2	• Contributes to the development of medulloblastoma [28].
Q86XK3	Swi5-dependent recombination DNA repair protein 1 homolog (*SFR1*)	• Promotes the progression of breast cancer [29].
Q8N568	Serine/threonine-protein kinase (*DCLK2*)	• Associates with poor prognosis of CLL [30].

The integration of the omics data sets of CLL cells, normal B-cells and normal PBMCs suggested 60 candidates as associated proteins. This table shows what is known in the literature about the relativeness of these proteins to malignant diseases.

References of the Supplementary Tables 1 and 3

1. Barnidge DR, Tschumper RC, Jelinek DF, Muddiman DC and Kay NE. Protein expression profiling of CLL B cells using replicate off-line strong cation exchange chromatography and LC–MS/MS. Journal of Chromatography B 2005;819:33-39.
2. Alsagaby SA, Khanna S, Hart KW, Pratt G, Fegan C, Pepper C, et al. Proteomics-based strategies to identify proteins relevant to chronic lymphocytic leukemia. Journal of proteome research 2014;13:5051-5062.
3. Eagle GL, Zhuang J, Jenkins RE, Till KJ, Jithesh PV, Lin K, et al. Total proteome analysis identifies migration defects as a major pathogenetic factor in immunoglobulin heavy chain variable region (ighv)-unmutated chronic lymphocytic leukemia. Molecular & Cellular Proteomics 2015;14:933-945.
4. Hoek KL, Samir P, Howard LM, Niu X, Prasad N, Galassie A, et al. A cell-based systems biology assessment of human blood to monitor immune responses after influenza vaccination. PloS one 2015;10:e0118528.
5. Bileck A, Kreutz D, Muqaku B, Slany A and Gerner C. Comprehensive assessment of proteins regulated by dexamethasone reveals novel effects in primary human peripheral blood mononuclear cells. Journal of proteome research 2014;13:5989-6000.
6. Lynch JT, Somerville TDD, Spencer GJ, Huang X and Somervaille TCP. TTC5 is required to prevent apoptosis of acute myeloid leukemia stem cells. Cell death & disease 2013;4:e573.
7. Wanitchakool P, Wolf L, Koehl GE, Sirianant L, Schreiber R, Kulkarni S, et al. Role of anoctamins in cancer and apoptosis. Phil. Trans. R. Soc. B 2014;369:20130096. http://dx.doi.org/10.1098/rstb.2013.0096.
8. Astori A, Fredly H, Aloysius TA, Bullinger L, Mansat-De Mas V, de la Grange P, et al. CXXC5 (retinoid-inducible nuclear factor, RINF) is a potential therapeutic target in high-risk human acute myeloid leukemia. Oncotarget 2013;4:1438-1448.
9. Knappskog S, Myklebust LM, Busch C, Aloysius T, Varhaug JE, Lønning PE, et al. RINF (CXXC5) is overexpressed in solid tumors and is an unfavorable prognostic factor in breast cancer. Annals of oncology 2011;2208-2215.
10. Torres-Ayuso P, Daza-Martín M, Martín-Pérez J, Ávila-Flores A and Mérida I. Diacylglycerol kinase α promotes 3D cancer cell growth and limits drug sensitivity through functional interaction with Src. Oncotarget 2014;5:9710-9726.
11. Cai K, Mulatz K, Ard R, Nguyen T and Gee SH. Increased diacylglycerol kinase ζ expression in human metastatic colon cancer cells augments Rho GTPase activity and contributes to enhanced invasion. BMC cancer 2014;14:208-218.
12. Heusschen R, Griffioen AW and Thijssen VL. Galectin-9 in tumor biology: a jack

of multiple trades. Biochimica et Biophysica Acta (BBA)-Reviews on Cancer 2013;1836:177-185.

13. Pena C, Mirandola L, Figueroa JA, Hosiriluck N, Suvorava N, Trotter K, et al. Galectins as therapeutic targets for hematological malignancies: a hopeful sweetness. Annals of translational medicine 2014;2:87-94.
14. Choi D-B, Park M-R, Kim H-R, Jun C-D, Kim H-J, Shim H, et al. Aberrant proteomic expression of NSRP70 and its clinical implications and connection to the transcriptional level in adult acute leukemia. Leukemia research 2014;38:1252-1259.
15. Chadeau-Hyam M, Vermeulen RCH, Hebels D, Castagné R, Campanella G, Portengen L, et al. Prediagnostic transcriptomic markers of Chronic lymphocytic leukemia reveal perturbations 10 years before diagnosis. Annals of Oncology 2014;25:1065-1072.
16. Zhang Q, Chen X, Zhang X, Zhan J and Chen J. Knockdown of TMEM14A expression by RNAi inhibits the proliferation and invasion of human ovarian cancer cells. Bioscience reports 2016;36:e00298.
17. Jin H, Kang ES, Kim HJ, Lee JH, Chang KC, Park J-Y, et al. TMEM14A inhibits N-(4-hydroxyphenyl) retinamide-induced apoptosis through the stabilization of mitochondrial membrane potential. Cancer letters 2011;309:190-198.
18. Grandjenette C, Kennel A, Faure GC, Béné MC and Feugier P. Expression of functional toll-like receptors by B-chronic lymphocytic leukemia cells. Haematologica 2007;92:1279-1281.
19. Barcellini W, Imperiali FG, Zaninoni A, Reda G, Consonni D, Fattizzo B, et al. Toll-like receptor 4 and 9 expression in B-chronic lymphocytic leukemia: relationship with infections, autoimmunity and disease progression. Leukemia & lymphoma 2014;55:1768-1773.
20. Zhong J, Zhao M, Ma Y, Luo Q, Liu J, Wang J, et al. UCHL1 acts as a colorectal cancer oncogene via activation of the β-catenin/TCF pathway through its deubiquitinating activity. International journal of molecular medicine 2012;30:430-436.
21. Jang MJ, Baek SH and Kim JH. UCH-L1 promotes cancer metastasis in prostate cancer cells through EMT induction. Cancer letters 2011;302:128-135.
22. Bheda A, Yue W, Gullapalli A, Shackelford J and Pagano JS. PU. 1-dependent regulation of UCH L1 expression in B-lymphoma cells. Leukemia & lymphoma 2011;52:1336-1347.
23. Feng X-Y, Wen X-Z, Tan X-J, Hou J-H, Ding Y, Wang KF, et al. Ectopic expression of B and T lymphocyte attenuator in gastric cancer: A potential independent prognostic factor in patients with gastric cancer. Molecular medicine reports 2015;11:658-664.
24. M'Hidi H, Thibult M-L, Chetaille B, Rey F, Bouadallah R, Nicollas R, et al. High ex-

pression of the inhibitory receptor BTLA in T-follicular helper cells and in B-cell small lymphocytic lymphoma/chronic lymphocytic leukemia. American journal of clinical pathology 2009;132:589-596.

25. Yokota T, Kouno J, Adachi K, Takahashi H, Teramoto A, Matsumoto K, et al. Identification of histological markers for malignant glioma by genome-wide expression analysis: dynein, α-PIX and sorcin. Acta neuropathologica 2006;111:29-38.
26. Koh SS, Wei J-PJ, Li X, Huang RR, Doan NB, Scolyer RA, et al. Differential gene expression profiling of primary cutaneous melanoma and sentinel lymph node metastases. Modern Pathology 2012;25:828-837.
27. Wysocka M, Kossenkov AV, Benoit BM, Troxel AB, Singer E, Schaffer A, et al. CD164 and FCRL3 Are Highly Expressed on CD4+ CD26– T Cells in Sézary Syndrome Patients. Journal of Investigative Dermatology 2014;134:229-236.
28. Lu Y, Ryan SL, Elliott DJ, Bignell GR, Futreal PA, Ellison DW, et al. Amplification and overexpression of Hsa-miR-30b, Hsa-miR-30d and KHDRBS3 at 8q24. 22-q24. 23 in medulloblastoma. PloS one 2009;4:e6159.
29. Feng Y, Singleton D, Guo C, Gardner A, Pakala S, Kumar R, et al. DNA homologous recombination factor SFR1 physically and functionally interacts with estrogen receptor alpha. PloS one 2013;8:e68075.
30. Trojani A, Di Camillo B, Tedeschi A, Milena L, Montesano S, Ricci F, et al. Microarray Gene Expression Signature Indicates ARSD As a New Marker Associated with Igvh Mutational Status, ZAP-70 and Disease Progression in CLL. Blood 2011;118:3447.