Proteomic markers with prognostic impact on outcome of chronic lymphocytic leukemia patients under chemo-immunotherapy: results from the HOVON 109 study

Fatemeh Saberi Hosnijeh, Lina van der Straten, Arnon P. Kater, Marinus H.J. van Oers, Ward F.M. Posthuma, Martine E.D. Chamuleau, Mar Bellido, Jeanette K. Doorduin, Michel van Gelder, Mels Hoogendoorn, Fransien de Boer, G. Doreen te Raa, J. Martijn Kerst, Erik W.A. Marijt, Reinier A.P. Raymakers, Harry R. Koene, Martijn R. Schaafsma, Johan A. Dobber, Sanne H. Tonino, Sabina S. Kersting, Anton W. Langerak, and Mark-David Levin

Department of Immunology, Laboratory Medical Immunology, Erasmus MC, University Medical Center, Rotterdam, The Netherlands; Institute for Risk Assessment Sciences, Division of Environmental Epidemiology, Utrecht University, Utrecht, The Netherlands; Department of Internal Medicine, Albert Schweitzer Hospital, Dordrecht, The Netherlands; Department of Hematology and Lymphoma and Myeloma Center Amsterdam, Academic Medical Center, Amsterdam, The Netherlands; Department of Internal Medicine, Reinder de Graaf Hospital, Delft, The Netherlands; Department of Hematology, Leiden University Medical Center, Leiden, The Netherlands; Department of Hematology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands; Department of Hematology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands; Department of Hematology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands; Department of Hematology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands; Department of Hematology, VU University Medical Center, Amsterdam, The Netherlands; Department of Hematology, Academic Medical Center, Amsterdam, The Netherlands; Department of Hematology, University Hospital, Maastricht, The Netherlands; Department of Hematology, Medical Center Leeuwarden, The Netherlands; Department of Internal Medicine, Ibaiz Hospital, Rotterdam, The Netherlands; Department of Internal Medicine, Gelderland Valley Hospital, Ede, The Netherlands; Department of Medical Oncology, Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands; Department of Hematology, University Medical Center, Utrecht, The Netherlands; Department of Hematology, University Medical Center, Amsterdam, The Netherlands; Department of Hematology, Haga Hospital, The Hague, The Netherlands

(Received 10 June 2020; revised 23 July 2020; accepted 4 August 2020)

Despite recent identification of several prognostic markers, there is still a need for new prognostic parameters able to predict clinical outcome in chronic lymphocytic leukemia (CLL) patients. Here, we aimed to validate the prognostic ability of known (proteomic) markers measured pretreatment and to search for new proteomic markers that might be related to treatment response in CLL. To this end, baseline serum samples of 51 CLL patients treated with chemo-immunotherapy were analyzed for 360 proteomic markers, using Olink technology. Median event-free survival (EFS) was 23 months (range: 1.25–60.9). Patients with high levels of sCD23 (>11.27, \( p = 0.026 \)), sCD27 (>11.03, \( p = 0.04 \)), SPINT1 (>1.6, \( p = 0.001 \)), and LY9 (>8.22, \( p = 0.0003 \)) had a shorter EFS than those with marker levels below the median. The effect of sCD23 on EFS differed between immunoglobulin heavy chain variable gene-mutated and unmutated patients, with the shortest EFS for unmutated CLL patients with sCD23 levels above the median. Taken together, our results validate the prognostic impact of sCD23 and sCD27.

AWL and MDL share senior authorship. FSH, AWL, and MDL designed the research. FSH and AWL contributed to the data analysis, FSH performed the statistical analysis, wrote the article, and had primary responsibility for the final content of the article. AWL and MDL contributed to the data analysis and writing of the article. All authors contributed to data collection and acquisition, interpretation of the present findings, and approval of the final version of the article for publication.

Offprint requests to: Fatemeh Saberi Hosnijeh, Department of Immunology, Laboratory Medical Immunology, Erasmus MC, University Medical Center, Postbus 2040, CA, Rotterdam 3000, The Netherlands; E-mail: f.saberihosnijeh@erasasmusmc.nl
The natural history of chronic lymphocytic leukemia (CLL) is highly heterogeneous. Some patients do not require treatment for decades, while others require direct treatment after diagnosis and experience diminished life expectancy because of CLL [1]. Since the introduction of multiple novel treatments, there is a growing need for informative prognostic markers with clinical significance that can influence the choice of standard of care [1].

To date, several prognostic markers have been identified, including immunoglobulin heavy chain variable (IGHV) gene mutation status, ZAP70, chromosomal alterations, CD38, CD40L, and biochemical parameters (e.g., lactate dehydrogenase and β2-microglobulin [β2M]) [1]. Yet, the identification of new prognostic parameters that are able to predict clinical outcome after treatment is important for patient management and may be useful in guiding therapeutic decisions.

Proteomic data are increasingly being used for biomarker discovery and for gaining mechanistic insight into lymphoid malignancies. Among the general population, an immune system environment that is characterized by elevated levels of B-cell stimulatory cytokines has been suggested to contribute to the development of B-cell lymphoma, including CLL [2–5]. In a study including 105 newly diagnosed and untreated CLL patients, high levels of soluble CD23 (sCD23) at the time of initial diagnosis were a strong predictor of progressive disease within the first year of disease presentation [6]. Other studies obtained similar results implicating sCD23 as a suitable marker with prognostic potential in CLL at diagnosis [7–9].

The aims of this pilot study were (1) to validate the prognostic ability of known proteomic markers (in particular B-cell activation markers such as sCD23 and sCD27) measured pretreatment for treatment response in CLL patients, and (2) to search for new proteomic markers and their biological pathways that might be related to treatment response.

**Methods**

Study subjects were selected from the HOVON 109 clinical study, which is a phase I/II trial designed for efficacy and safety of first-line therapy involving chlorambucil, rituximab, and lenalidomide in elderly patients and young frail patients with advanced CLL [10]. Total treatment duration was 12 months, and all patients were followed until 5 years after registration. Of 63 patients enrolled in the HOVON 109 study, 51 with an available sample at baseline were included in our current pilot study. Clinical data were collected from the HOVON database.

Four commercially available proteomic panels (Oncology, Inflammation, Immune response, and Development; Olink Bioscience, Uppsala, Sweden) including 360 low-abundance serum proteins were selected for this study, covering multiple proteomic markers previously associated with incidence or progression of CLL in published clinical and population-based studies [2–9]. Serum samples were analyzed using a multiplex proximity extension assay. Data were expressed in the arbitrary unit NPX (Normalized Protein eXpression) on a log2 scale and linearized using the formula $2^{\text{NPX}}$, where a high NPX value corresponds to a high protein concentration (Supplementary Table E1, online only, available at www.exphem.org).

Statistical procedures are provided in detail in the Supplementary Data (online only, available at www.exphem.org). As the analyses of B-cell activation markers that were found to be predictive in the general population (sCD23 and sCD27) rely on a prior hypothesis, these were not corrected for multiple testing. All other $p$ values were corrected for multiple testing.

**Results**

Clinical characteristics of the CLL patients (30 males, 21 females; median age: 71 years) are summarized in Supplementary Table E2 (online only, available at www.exphem.org). Until 5 years after registration, 26 events were recorded. We first evaluated measured levels of the proteomic markers in the context of other known prognostic factors. Mutated IGHV status was significantly associated with lower levels of sCD23 and higher levels of nuclear factor of activated T cells 3 (NFATC3), while a positive association between β2M levels and 58 markers was established (Supplementary Table E3, online only, available at www.exphem.org). No marker remained significantly associated with Rai and cytogenetic aberrations after multiple testing correction.

In log-rank testing of known prognostic factors, male patients had shorter EFS as compared with female patients, and IGHV-mutated CLL patients had longer EFS than unmutated CLL patients (Supplementary Figure E1, online only, available at www.exphem.org). No significant EFS effects were seen for β2M levels, Rai stage, or well-defined chromosome aberrations. The latter may be due to the small number of cases with these aberrations in our pilot study.

From the previously studied B-cell activation proteomic markers (Supplementary Table E4, online only, available at www.exphem.org), only the sCD23 level was significantly associated with EFS (HR = 1.56, 95% CI = 1.02–2.4, $p = 0.04$) in the univariate model, while none appeared significantly correlated to EFS in the models adjusted for gender and IGHV status. As
sCD23 level was related to IGHV status, the interaction term was further included in the adjusted model that resulted in a borderline significant association for sCD23 (hazard ratio [HR] = 0.23, 95% confidence interval [CI] = 0.04 – 1.14, \( p = 0.07 \)).

When using median values as threshold, patients with sCD23 and sCD27 levels above the median did have a shorter EFS than those with marker levels below the median (Figure 1A,B). Notably, when combining these markers with significant CLL prognostic factors in this cohort (i.e., IGHV mutation status and gender), sCD23 or sCD27 levels above the median were associated with the lowest EFS in unmutated IGHV patients (Figure 1C,D). In contrast, survival distributions for male and female patients with sCD23 or sCD27 levels above the median did not differ significantly (Figure 1E,F).

Several newly studied proteomic markers were significantly associated with EFS in Cox regression models, albeit that upon multiple testing correction, even the top four proteomic markers from the models exhibited only a trend toward significance (Supplementary Table E5, online only, available at www.exphem.org). Nevertheless, the two markers with \( p \) values closest to significance, that is, serine peptidase inhibitor SPINT1 and surface antigen LY9, were associated with a significantly longer EFS in patients with marker levels equal or lower than the median (Figure 2). Moreover, CLL patients with higher levels of IFNLR1 had a shorter EFS than those patients with marker levels equal to or lower than the median.

Findings were independently validated via Lasso regression (details are in the Supplementary Data).

The stepwise Cox regression model including gender and IGHV mutation status plus sCD23 and the top four proteomic markers from the univariate and/or multivariable (model M2, Table 1) revealed a significant independent effect for sCD23 (HR = 0.27, 95% CI = 0.12 – 0.58, \( p = 0.0008 \)), SPINT1 (HR = 3.6, 95% CI = 1.12 – 11.5, \( p = 0.03 \)), LY9 (HR = 7.03, 95% CI = 1.6 – 30.8, \( p = 0.009 \)), and CLEC7A (HR = 0.43, 95% CI = 0.27 – 0.7, \( p = 0.0007 \)) over gender and IGHV status. The AUC of the model was higher (0.61) compared with that of the model including only gender and IGHV status (0.38). Because of the limited sample size of our study resulting in the wide confidence intervals, these findings should be interpreted with caution and require validation in larger studies.

Figure 1. Kaplan–Meier curves for EFS related to the B-cell activation proteomic markers.
Discussion

In this pilot study, we were able to validate sCD23 levels above the median as significantly associated with a shorter EFS, which is consistent with previous studies [7]. Soluble CD23 is released from activated B cells and can itself induce further B-cell stimulation as well

Table 1. Adjusted HRs for event-free survival for combination of the proteomic markers and known prognostic factors

|                | M1 (stepwise) |       | M2 (stepwise) |       |
|----------------|--------------|-------|--------------|-------|
|                | HR (95% CI)  | p     | HR (95% CI)  | p     |
| Female         | 0.40 (0.16–0.98) | 0.04  | 0.21 (0.06–0.66) | 0.008 |
| Unmutated IGHV | 4.25 (1.68–10.8) | 0.002 | 3.9 (1.28–11.9) | 0.016 |
| sCD23          |              |       | 0.27 (0.12–0.58) | 0.0009 |
| sCD23* IGHV    | NS           |       | NS           |       |
| SPINT1         |              |       | 3.6 (1.12–11.5) | 0.031 |
| LY9            |              |       | 7.03 (1.6–30.8) | 0.009 |
| IFNLR1         | NS           |       | NS           |       |
| ITM2A          | NS           |       | NS           |       |
| CLEC7A         |              |       | 0.43 (0.27–0.7) | 0.0007 |
| TCL1A          |              |       | 0.60 (0.48)  |       |
| $R^2$ (corrected) | 0.28 (0.24) |       | 0.40 (0.38)  |       |
| AUC (corrected) |              |       | 0.68 (0.61)  |       |

$AUC$=Area under the curve; CI=confidence interval; NS=nonsignificant.

$^aR^2$ and AUC corrected for overfitting by 100 times bootstrapping.
as function as a potent mitogenic growth factor. Notably, sCD23 seems to be associated with IGHV status, since mutated CLL patients had significantly lower levels of sCD23, as compared with unmutated patients. When both markers are used, patients with unmutated IGHV genes with sCD23 levels above the median can be regarded as very poor prognostic group. Similarly, patients with unmutated IGHV genes and high levels of sCD27 were found to have an even poorer prognosis. High sCD27 levels were described to be associated with higher Rai stage, β2M, and LDH among CLL patients [11,12]. Here we found that high levels of sCD27 were associated with β2M levels and inferior prognosis.

Our study further indicated that higher levels of SPINT1 are associated with shorter EFS time and higher β2M levels. SPINT1, an enzyme that is encoded by the Kunitz-type protease inhibitor 1 gene, modulates matriptase proteolytic activity. Matriptase was identified in Burkitt lymphoma cells [13] and later in CLL [14]. In fact, it is highly upregulated in CLL and promotes cancer cell invasion either directly by degrading matrix proteins or indirectly by activating growth factors or through yet unknown mechanisms [14].

In our cohort, CLL patients with high LY9 levels had an inferior EFS as compared with patients with low LY9 levels. LY9 is known to interact with SLAM-associated protein that has been implicated in autoimmunity. It has been reported that LY9 is a naturally processed antigen in CLL and can serve as tumor-associated antigen in this disease [15]. It was reported that LY9-specific cytotoxic T cells from CLL patients efficiently recognized native and CD40L-activated autologous malignant CLL cells via MHC-I molecules. These findings provide strong evidence that LY9 can be employed for the design of T cell-based immunotherapeutic strategies of LY9-expressing malignancies including CLL [15] and, thus, underline the impact of the results from our current study.

A major strength of this pilot study is the large set of novel proteomic markers, which we measured and which were previously not extensively described in CLL patients. Despite the relatively small number of available cases, which had an impact on statistical power, our pilot study identified SPINT1 and LY9 as promising independent prognostic proteomic markers next to sCD23 and sCD27 in patients treated for CLL. Further studies with larger sample sizes are required to validate these results. Also, as proteomic markers were solely measured before treatment, changes in marker levels during treatment should be evaluated in new CLL patient cohorts.

Acknowledgments
This work was supported by an EU TRANSCAN/Dutch Cancer Society grant (179; NOVEL consortium; to AWL) and an unrestricted research grant from Gilead Sciences, Netherlands BV (to AWL and FSH). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the article.

The authors acknowledge Martijn Kolijn, Erasmus MC, for critical discussions; Julie M. N. Dubois and Ingrid Derks, Academic Medical Center, Amsterdam, for sample preparation; Tamara Pesic, Erasmus MC, for IGHV sequencing; and Kirsten K. J. Gussinklo, Erasmus MC, for FISH analysis.

Conflict of interest disclosure
None of the authors declared a conflict of interest.

References
1. Kay NE, O’Brien SM, Pettitt AR, Stilgenbauer S. The role of prognostic factors in assessing ‘high-risk’ subgroups of patients with chronic lymphocytic leukemia. Leukemia. 2007;21:1885–1891.
2. Kaaks R, Sookthai D, Luczyska A, et al. Lag times between lymphoproliferative disorder and clinical diagnosis of chronic lymphocytic leukemia: a prospective analysis using plasma soluble CD23. Cancer Epidemiol Biomarkers Prev. 2015;24:538–545.
3. Hosnjieh FS, Portengen L, Spåth F, et al. Soluble B–cell activation marker of sCD27 and sCD30 and future risk of B–cell lymphomas: a nested case–control study and meta–analyses. Int J Cancer. 2016;138:2357–2367.
4. Epstein MM, Rosner B, Breen EC, et al. Pre-diagnosis plasma immune markers and risk of non-Hodgkin lymphoma in two prospective cohort studies. Haematologica. 2018;103:1679–1687.
5. Makgoeng SB, Bolanos RS, Jeon CY, et al. Markers of immune activation and inflammation, and non-Hodgkin lymphoma: a meta-analysis of prospective studies. JNCI Cancer Spectr. 2019;2:pyk082.
6. Knauf WU, Langenmayer I, Ehlers B, et al. Serum levels of soluble CD23, but not soluble CD25, predict disease progression in early stage B-cell chronic lymphocytic leukemia. Leuk Lymphoma. 1997;27:523–532.
7. Sarfati M, Chevret S, Chastang C, et al. Prognostic importance of serum soluble CD23 level in chronic lymphocytic leukemia. Blood. 1996;88:4259–4264.
8. Schwarzermeier JD, Shehata M, Hilgarth M, et al. The role of solubleCD23 in distinguishing stable and progressive forms of B-chronic lymphocytic leukemia. Leuk Lymphoma. 2002;43:549–554.
9. Saka B, Aktan M, Sami U, Oner D, Sanem O, Dincol G. Prognostic importance of soluble CD23 in B-cell chronic lymphocytic leukemia. Clin Lab Haematol. 2006;28:30–35.
10. Kater AP, van Oers MH, van Norden Y, et al. Feasibility and efficacy of addition of individualized-dose lenalidomide to chlorambucil and rituximab as first-line treatment in elderly and FCR-unfit patients with advanced chronic lymphocytic leukemia. Haematologica. 2019;104:147–154.
11. Molica S, Vitelli G, Levato D, et al. CD27 in B-cell chronic lymphocytic leukemia: cellular expression, serum release and correlation with other soluble molecules belonging to nerve growth factor receptors (NGF) superfamily. Haematologica. 1998;83:398–402.
12. Kara IO, Sahin B, Gunesacar R. Expression of soluble CD27 and interleukins-8 and -10 in B-cell chronic lymphocytic leukemia: correlation with disease stage and prognosis. Adv Ther. 2007;24:29–40.

13. Chou FP, Chen YW, Zhao XF, et al. Imbalanced matriptase pericellular proteolysis contributes to the pathogenesis of malignant B-cell lymphomas. Am J Pathol. 2013;183:1306–1317.

14. Gao L, Liu M, Dong N, et al. Matriptase is highly upregulated in chronic lymphocytic leukemia and promotes cancer cell invasion. Leukemia. 2013;27:1191–1194.

15. Bund D, Mayr C, Kofler DM, Hallek M, Wendtner CM. Human Ly9 (CD229) as novel tumor-associated antigen (TAA) in chronic lymphocytic leukemia (B-CLL) recognized by autologous CD8+ T cells. Exp Hematol. 2006;34:860–869.
SUPPLEMENTARY INFORMATION

Supplementary Methods
Written informed consent was obtained before enrollment in the trial. The study was approved by an accredited Ethical Committee and Institutional Review Board and was performed according to the Declaration of Helsinki, the International Conference on Harmonization Good Clinical Practice Guidelines and the European Union Clinical Trial Directive (2001/20/EG). The study was registered with EuraCT number 2010-022294-34 [1].

Protein measurements. Serum samples were analyzed using a multiplex proximity extension assay. In brief, 1 µL sample was incubated in the presence of proximity antibody pairs tagged with DNA-reporter molecules. Once the pair of antibodies is bound to their corresponding antigens, the respective DNA tails form an amplicon by proximity extension, which was quantified by high-throughput real-time PCR (BioMark™ HD System, Fluidigm Corporation). Protein abundance directly correlates with the generated fluorescent signal, which is expressed in quantitation cycles produced by the BioMark’s Real-Time PCR Software following the Proseek Multiplex protocol. To minimize variation within and between runs, data (Ct values) were normalized using both an internal control (extension control) and an interplate control, and then transformed using a pre-determined correction factor.

The limit of detection was determined for each biomarker based on the mean value of 4 controls analyzed in each run. Thirty-five markers were excluded from statistical analyses because of a high non-detection rate in the cohort (i.e., >30% of the cases) (Supplementary Table E1, footnote). For 35 samples, the non-detection rate ranged from 2% to 30% (median 7%) and those were set to the value of the lower limit of quantification divided by the square root of two. Around 83% of markers were detected in all samples. The list of biomarkers (n=322) included in the statistical analyses and their median value is shown in Supplementary Table E1.

Statistical analyses. Differences in marker distributions across different levels of prognostic factors were evaluated by Wilcoxon tests. Kaplan–Meier plot and Log-rank test were used to examine the survival distribution for protein markers (< median or >median) and other known prognostic factors. The median of the protein markers are shown in Supplementary Table E1.

Cox proportional hazard models were used for testing the effects of the markers for event-free survival (EFS; time from registration to induction failure, progression, or death from any cause, whichever comes first). Induction failure was defined as not having achieved at least a PR during/after a maximum of 12 cycles. Unadjusted analyses were carried out for each proteomic marker. Due to the limited sample size of the study, the Cox model of proteomic markers was further adjusted only for significant known prognostic factors (gender and IGHV status).

Of interest was the added prognostic value of proteomic markers that might help to assess prognosis in clinical practice. Two models, one containing only significant known prognostic markers and the other containing the first model plus significant proteomic markers were compared by estimating the area under a ROC (receiver operating characteristic) curve (AUC) for EFS. R square and AUC corrected for overfitting by 100 bootstrapping were reported.

Validation analysis. As standard regression models perform poorly in a situation with a data set containing a number of variables superior to the number of samples, we additionally applied the least absolute shrinkage and selection operator (Lasso) technique for variable selection [2]. It is a powerful method that perform two main tasks: regularization and feature selection. In order to do so the method applied a shrinking (regularization) process in which the coefficients of the regression variables were penalized, thus shrinking some of them to zero. During the feature selection process the variables that still have a non-zero coefficient after the shrinking process were selected to be part of the model. Optimal tuning parameter λ, which controls the strength of the penalty, was obtained by 5 folds cross-validation.

Supplementary Results

Validation analysis by means of Lasso regression. In total 8 proteomic markers were selected in association with EFS in Lasso analysis (Supplementary Table E6). Interestingly, five proteomic markers suggested in our Cox regression models (LY9, SPINT1, ITM2A, IFNLR1, and CLEC7A) were among the selected variables, thus supporting the validity of these markers as possible prognostic markers.

Fourteen proteomic markers [sCD23, TCL1A, ITM2A, CLEC7A, LY9, VEGFR-3, IFNLR1, SPINT1, TNF receptor-associated factor 2 (TRAF2), Semaphorin-7A (SEMA7A), signal-regulatory protein beta-1 (SIRPB1), C-type lectin domain family 4 member C (CLEC4C), tripeptidyl peptidase 1 (TPP1), monocye chemoattractant protein 1 (MCP1)] found to be associated with EFS in the adjusted Cox regression models at p<0.05 and/or in Lasso analysis were subjected to computational functional analyses. Molecular functions, biological process, and cellular components related to the proteins are shown in Supplementary Figure E2. Due to the limited sample size and protein numbers, further pathway analyses and statistical enrichment detection were not possible.

Reference

1 Kater AP, van Oers MH, van Norden Y, et al. Feasibility and efficacy of addition of individualized-dose lenalidomide to chlorambucil and rituximab as first-line treatment in elderly and FCR-unfit patients with advanced chronic lymphocytic leukemia. *Haematologica* 2019; 104(1): 147-154.

2 Simon N, Friedman J, Hastie T, Tibshirani R. Regularization paths for Cox’s proportional hazards model via coordinate descent. *Journal of statistical software* 2011; 39(5): 1.
Supplementary Table E1. Proteomic markers with their median values included in the statistical analyses

| Marker Label | Marker M* | Marker Label | Marker M* | Marker Label | Marker M* |
|--------------|-----------|--------------|-----------|--------------|-----------|
| DevOID01400  | CTSF      | 2.32         | DevOID01447 | CD109        | 2.56      |
| DevOID01401  | MATN2     | 2.68         | DevOID01448 | VSVG4        | 5.13      |
| DevOID01402  | HTRA2     | 0.68         | DevOID01449 | CRIM1        | 1.71      |
| DevOID01403  | MFGE8     | 3.72         | DevOID01450 | CGA          | 4.77      |
| DevOID01405  | ACAN      | 2.97         | DevOID01451 | CLEC14A      | 3.82      |
| DevOID01406  | ITGB1     | 6.14         | DevOID01452 | CA6          | 0.37      |
| DevOID01407  | DKK3      | 4.23         | DevOID01453 | MIF          | 7.56      |
| DevOID01408  | SPINT2    | 2.46         | DevOID01454 | CRELD2       | 2.50      |
| DevOID01409  | APP       | 4.16         | DevOID01455 | ITGA5        | 1.79      |
| DevOID01410  | HAVCR2    | 3.21         | DevOID01456 | PEAR1        | 5.05      |
| DevOID01411  | DSC2      | 4.59         | DevOID01457 | MESDC2       | 2.14      |
| DevOID01412  | TMSB10    | 6.03         | DevOID01458 | CD99L2       | 1.59      |
| DevOID01413  | ANGPTL4   | 2.43         | DevOID01459 | SCARF1       | 6.74      |
| DevOID01414  | INHBc     | 1.56         | DevOID01460 | LGMN         | 3.02      |
| DevOID01415  | SPINK1    | 3.37         | DevOID01461 | SEMA7A       | 6.72      |
| DevOID01416  | CLEC11A   | 3.69         | DevOID01462 | COLEC12      | 3.57      |
| DevOID01417  | PDGFRB    | 3.95         | DevOID01463 | GUSB         | 3.42      |
| DevOID01418  | COCH      | 5.47         | DevOID01464 | FUT3/FUT5    | 2.44      |
| DevOID01419  | FCR5L5    | 5.72         | DevOID01465 | STIP1        | 0.64      |
| DevOID01420  | PELA      | 2.84         | DevOID01466 | B4GAT1       | 4.18      |
| DevOID01421  | B4GALT1   | 3.07         | DevOID01467 | CD69         | 5.76      |
| DevOID01422  | CD300LG   | 3.38         | DevOID01468 | CRHBP        | 2.07      |
| DevOID01423  | PARK7     | 4.27         | DevOID01469 | MSMB         | 2.64      |
| DevOID01424  | BLVRB     | 0.11         | DevOID01470 | SPINT1       | 1.59      |
| DevOID01426  | CD74      | 2.69         | DevOID01471 | FSTL3        | 2.23      |
| DevOID01427  | CCL21     | 0.58         | DevOID01472 | IL13RA1      | 1.65      |
| DevOID01428  | PTPRF     | 2.72         | DevOID01473 | CD58         | 2.26      |
| DevOID01429  | BCAM      | 3.36         | DevOID01474 | SCGB3A1      | 2.34      |
| DevOID01430  | SIRPB1    | 3.17         | DevOID01475 | NOV          | 3.26      |
| DevOID01431  | IGF2R     | 6.25         | DevOID01476 | CNTN4        | 3.52      |
| DevOID01432  | P4HB      | 0.28         | DevOID01477 | CA2          | 3.85      |
| DevOID01433  | FUCa1     | 5.83         | DevOID01478 | XG           | 4.00      |
| DevOID01434  | ESAM      | 3.21         | DevOID01479 | ARSA         | 2.93      |
| DevOID01436  | BST1      | 1.28         | DevOID01480 | PPIB         | 3.93      |
| DevOID01437  | CST6      | 4.85         | DevOID01481 | SPINK5       | 0.04      |
| DevOID01438  | MYOC      | 3.19         | DevOID01482 | OMD          | 2.08      |
| DevOID01439  | SNAP29P   | 4.92         | DevOID01483 | PEBP1        | 5.74      |
| DevOID01440  | WFKK2N2   | 3.37         | DevOID01484 | PAMR1        | 3.78      |
| DevOID01441  | CDON      | 2.20         | DevOID01485 | ROBO1        | 2.08      |
| DevOID01442  | CD177     | 4.79         | DevOID01486 | CD23         | 11.3      |
| DevOID01443  | NID2      | 1.27         | DevOID01487 | LAMA4        | 2.45      |
| DevOID01444  | DAG1      | 2.75         | DevOID01488 | LAR1         | 4.55      |
| DevOID01445  | CD97      | 5.71         | DevOID01489 | RELT         | 5.37      |
| ImmOID00985  | CKA4P     | 6.32         | ImmOID00483 | IL-17C       | 0.99      |
| ImmOID00987  | HEX5M1    | 4.31         | ImmOID00484 | MCP-1        | 11.0      |
| ImmOID00988  | CLEC4D     | 2.29        | ImmOID00485 | CXCL11       | 8.67      |
| ImmOID00989  | PKRCC       | 0.27        | ImmOID00487 | AXIN1        | 3.49      |
| ImmOID00990  | MGMT      | 4.26         | ImmOID00490 | CXCL9        | 9.31      |
| ImmOID00991  | TREM1     | 2.03         | ImmOID00491 | CST5         | 5.89      |
| ImmOID00992  | CXADR     | 1.55         | ImmOID00494 | OSM          | 5.38      |
| ImmOID00994  | SRP2K2    | 1.18         | ImmOID00496 | CXCL1        | 8.40      |
| ImmOID00995  | KLRLD1    | 5.69         | ImmOID00498 | CCL4         | 8.93      |
| ImmOID00996  | BACH1     | 1.84         | ImmOID00499 | CD6          | 6.90      |
| ImmOID00997  | PIK3AP1   | 5.43         | ImmOID00501 | IL18         | 8.73      |
| ImmOID00999  | STC1      | 5.25         | ImmOID00502 | SLAMF1       | 2.04      |
| ImmOID01000  | FAM3B     | 4.04         | ImmOID00504 | MCP-4        | 3.92      |
| ImmOID01002  | SH2D1A    | 1.39         | ImmOID00505 | CCL11        | 7.63      |
| ImmOID01003  | ICA1      | 1.44         | ImmOID00506 | TNFSF14      | 7.08      |
| ImmOID01004  | DFAA      | 5.64         | ImmOID00507 | FGF-23       | 1.13      |
| ImmOID01005  | DCBLD2    | 2.35         | ImmOID00508 | IL-10RA      | 0.87      |
| ImmOID01006  | FCRL6     | 2.07         | ImmOID00509 | FGF-5        | 0.64      |

(continued)
| Marker      | Label | M* | Marker      | Label | M* | Marker      | Label | M* |
|------------|-------|----|------------|-------|----|------------|-------|----|
| ImmOID01007 | NCR1  | 1.82 | ImmOID00510 | MMP-1 | 13.6 | OncOID00659 | CEACAM1 | 6.66 |
| ImmOID01010 | IFNLR1 | 3.52 | InfOID01011 | LIF-R | 2.60 | OncOID00660 | MSLN  | 4.01 |
| ImmOID01011 | DAPPI | 1.30 | InfOID00512 | FGF-21 | 4.58 | OncOID00661 | TNFSF13 | 7.59 |
| ImmOID01013 | SIT1  | 3.79 | InfOID00513 | CCL19 | 8.59 | OncOID00662 | EGF   | 11.1 |
| ImmOID01014 | MASP1 | 1.08 | InfOID00514 | IL-15RA | 0.56 | OncOID00663 | TNFRSF6B | 6.60 |
| ImmOID01015 | LAMP3 | 3.72 | InfOID00515 | IL-10R | 7.30 | OncOID00664 | SYND1 | 6.92 |
| ImmOID01016 | CLEC7A | 3.36 | InfOID00516 | IL-18R1 | 6.84 | OncOID00665 | TGFR-2 | 6.71 |
| ImmOID01017 | CLEC6A | 2.38 | InfOID00517 | PD-L1 | 4.50 | OncOID00666 | IL6   | 3.83 |
| ImmOID01018 | DDX58 | 4.17 | InfOID00519 | Beta-NGF | 1.53 | OncOID00667 | CD48  | 7.85 |
| ImmOID01019 | IL12RBI | 2.32 | InfOID00520 | CXCL5 | 9.74 | OncOID00668 | SCAMP3 | 6.67 |
| ImmOID01020 | TANK  | 2.48 | InfOID00521 | TRANCE| 3.97 | OncOID00669 | LY9   | 8.22 |
| ImmOID01021 | ITGA11 | 1.35 | InfOID00522 | HGF  | 8.63 | OncOID00670 | IFN-gamma-R1 | 5.34 |
| ImmOID01023 | LG3   | 2.53 | InfOID00523 | IL-12B | 4.65 | OncOID00671 | ITGAV | 3.60 |
| ImmOID01025 | CD83  | 3.56 | InfOID00527 | MMP-10| 5.44 | OncOID00672 | TRAIL | 8.07 |
| ImmOID01026 | ITGB6 | 1.63 | InfOID00528 | IL10  | 3.27 | Onc OID00673 | hK11  | 6.72 |
| ImmOID01027 | BTN3A2 | 1.98 | InfOID00530 | CCL23 | 9.95 | OncOID00674 | GPC1  | 5.05 |
| InfOID00471 | IL8   | 10.9 | InfOID00531 | CD5  | 8.18 | OncOID00675 | TFPI-2 | 8.10 |
| InfOID00472 | VEGFA | 9.92 | InfOID00532 | CCL3  | 7.86 | OncOID00676 | hK8   | 7.29 |
| InfOID00474 | MCP-3 | 3.08 | InfOID00533 | Fli3L | 9.08 | OncOID00677 | VEGFR-2 | 7.14 |
| InfOID00475 | GDNF  | 0.90 | InfOID00534 | CXCL6 | 8.46 | OncOID00678 | LYPD3 | 4.70 |
| InfOID00476 | CDCP1 | 3.33 | InfOID00535 | CXCL10| 8.78 | OncOID00679 | PODXL | 3.60 |
| InfOID00477 | CD24A | 5.69 | InfOID00536 | 4E-BP1| 7.53 | OncOID00680 | S100A4| 3.26 |
| InfOID00478 | IL7   | 2.53 | InfOID00538 | SIRT2 | 3.35 | OncOID00681 | IFNGR1| 3.52 |
| InfOID00479 | OPG   | 10.6 | InfOID00539 | CCL28 | 0.94 | OncOID00682 | ERBB2 | 7.79 |
| InfOID00480 | LAP   | 8.00 | InfOID001213 | DNER | 7.36 | OncOID00683 | ERBB3 | 8.08 |
| InfOID00481 | uPA   | 10.2 | InfOID00541 | EN-RAGE| 6.03 | OncOID00684 | SCF   | 9.28 |
| OncOID00685 | SPARC | 6.06 | OncOID00708 | 5'-NT | 10.8 | OncOID00730 | CD207 | 2.81 |
| OncOID00686 | GZMH  | 5.97 | OncOID00709 | CDKN1A| 2.46 | OncOID00731 | ICOSLG| 6.45 |
| OncOID00687 | TGF-alpha | 8.85 | OncOID00710 | DLL1 | 10.4 | OncOID00732 | WDFC2 | 8.68 |
| OncOID00688 | FURIN | 9.21 | OncOID00711 | MK  | 5.92 | OncOID00733 | CXCL13| 9.70 |
| OncOID00689 | CYR61  | 5.38 | OncOID00712 | ABL1 | 4.42 | OncOID00734 | MAD homolog 5 | 3.62 |
| OncOID00690 | hK14  | 6.94 | OncOID00713 | FGF-BP1| 5.19 | OncOID00735 | ADAM-TS 15| 1.56 |
| OncOID00691 | FADD  | 2.65 | OncOID00714 | TLR3 | 5.51 | OncOID00736 | CD70  | 6.96 |
| OncOID00692 | MetAP 2 | 5.38 | OncOID00715 | LYN  | 3.57 | OncOID00737 | RSPO3 | 3.45 |
| OncOID00693 | PVRL4 | 4.83 | OncOID00716 | RET  | 5.30 | OncOID00738 | FR-gamma | 6.53 |
| OncOID00694 | FASLG | 9.28 | OncOID00717 | VIM  | 6.17 | OncOID00739 | CEACAM5 | 1.24 |
| OncOID00695 | EPHA2 | 3.93 | OncOID00718 | TNFRSF19| 4.59 | OncOID00740 | VEGFR-3 | 7.22 |
| OncOID00696 | ITGB5 | 7.09 | OncOID00719 | CRNN | 4.47 | OncOID00741 | MUC-16 | 4.50 |
| OncOID00697 | Gal-1  | 7.07 | OncOID00720 | TCL1A | 11.2 | OncOID00742 | WIF-1 | 4.94 |
| OncOID00698 | SEZ6L | 3.48 | OncOID00721 | CD160 | 6.24 | OncOID00743 | GZMB  | 4.44 |
| OncOID00749 | GPMB  | 6.85 | OncOID00722 | TNFRSF4| 5.77 | OncOID00744 | FCRLB | 1.42 |
| OncOID00700 | CAIX  | 4.02 | OncOID00723 | MIC-A/B | 4.63 | OncOID00745 | ANXA1 | 4.90 |
| OncOID00701 | MIA   | 10.2 | OncOID00724 | WISP-1 | 6.27 | OncOID00746 | FR-alpha | 6.72 |
| OncOID00702 | CTSV  | 3.44 | OncOID00725 | CXCL7 | 3.91 | | | |
| OncOID00703 | CD27  | 11.0 | OncOID00726 | PPY  | 9.42 | | | |
| OncOID00704 | XPNEP2 | 8.56 | OncOID00727 | S100A11 | 4.34 | | | |
| OncOID00705 | ERBB4 | 5.26 | OncOID00728 | AREG | 3.40 | | | |
| OncOID00707 | ADAM 8 | 6.27 | OncOID00729 | ESM-1 | 8.85 | | | |

*Median (M) normalized protein expression level on log2 scale; The excluded markers due to a high non-detection rate were NFASC, KPNA1, BDNF, IL-1 alpha, IL-2, IL-4, IL-5, IL-13, IL-20, IL-24, IL-33, TSLP, EIF5A, IL-22 RA1, ARTN, PLA2G4A, IL-20RA, IFN-gamma, SPRY2, ARNT, IL-2RB, NTF4, PAD2, TNF, LIF, NRTN, JUN, DKGX, FXYD5, NF2, IL-17A, PTNP6, ITGAV, NUDT5, CXCL12.*
### Supplementary Table E2. Characteristics of the study participants

| Characteristic               | N=51 |   |
|-----------------------------|------|---|
| Age, years*                 | 71 (60-83) |   |
| Gender                      |      |   |
| Female, n                   | 21   |   |
| Male, n                     | 30   |   |
| Rai stage, n                |      |   |
| 0                            | 2    |   |
| 1                            | 9    |   |
| 2                            | 9    |   |
| 3                            | 20   |   |
| 4                            | 11   |   |
| Chromosome aberration       |      |   |
| Del17p13                     | 7/49  |   |
| Del13q14                     | 24/51 |   |
| Del11q22                     | 7/51  |   |
| Tr12                         | 11/51 |   |
| IGHV mutation status        |      |   |
| Mutated                     | 25/47 |   |
| Unmutated                   | 22/47 |   |
| Beta 2-Microglobulin (β2M), μg/mL* | 3.8 (1.6 - 10.4) |   |
| Survival status             |      |   |
| Alive                       | 47   |   |
| Dead                        | 4    |   |
| Survival duration*          | 31.7 (15.8-62.1) |   |
| Progression free survival (PFS) status | 27 |   |
| No progression              | 27   |   |
| Progression                 | 24   |   |
| PFS duration*               | 24.4 (1.25-60.9) |   |
| Event free survival (EFS) status | 25 |   |
| No event                    | 25   |   |
| Event                       | 26   |   |
| EFS duration*               | 23 (1.25-60.9) |   |

*Median (range); Missing numbers: Del17p13 (n=2), Immunoglobulin heavy chain variable genes (IGHV) mutation status (n=4), β2M (n=10)

### Supplementary Table E4. Cox models for EFS related to individual B-cell activation proteomic markers found to be predictive in population-based studies

| Markers | Univariate | Multivariable* |
|---------|------------|---------------|
|         | HR (95% CI) | P            |
|         | HR (95% CI) | P            |
| CD23    | 1.56 (1.02-2.4) | 0.04        |
|         | 1.21 (0.80-1.86) | 0.37        |
| CD27    | 2.9 (0.71-11.9)  | 0.14        |
|         | 2.45 (0.53-11.4) | 0.25        |

*Each protein marker individually adjusted for gender and IGHV mutation status; There was a significant interaction effect between sCD23 and IGHV status (HR=2.75, 95% CI= 1.04-7.27, P=0.04) implying that the effect of sCD23 on EFS would be different between mutated and unmutated patients (HR of 0.23 for one unit increase in sCD23 in mutated male CLL patients vs. HR of 3.4 for unmutated male CLL patients).

### Supplementary Table E5. Top 4 proteomic markers from the univariate and multivariable Cox regression models for EFS

| Markers | Univariate |
|---------|------------|
|         | HR (95% CI) | P   |
| SPINT1  | 3.16 (1.64-6.05) | 0.0005 |
| LY9     | 5.1 (1.94-13.41) | 0.0009 |
| IFNLR1  | 2.85 (1.51-5.40) | 0.0013 |
| ITM2A   | 3.40 (1.61-7.12) | 0.0013 |

*Each protein marker individually adjusted for gender and IGHV mutation status; Bonferroni correction level: p=0.05/322=0.0002

### Supplementary Table E3. Proteomic markers significantly associated with well-known CLL prognostic factors (corrected for multiple testing)

| Sex | IGHV Mut. | β2M | β2M (cont.) | β2M (cont.) | β2M (cont.) | LDH | pblymph | Pblymph (cont.) |
|-----|-----------|-----|-------------|-------------|-------------|-----|---------|-----------------|
|     | ANGPTL4   | CD23| TNFRSF4     | TNFRSF6B    | FSTL3       | LILRB4| BCAM    | PSIP1           |
|     | CGLA      | NFATC3| ROBO1      | SPINT1      | CD74       | CD28 | ITGB1   | DFFA            |
|     | B4GALT1   | TNFB| RELT        | ANGPTL4     | LAIR1      | BCAM | TFP1-2  | ABL1            |
|     | RELT      | TNFRSF9 | EPHA2      | TPF1        | PD-L1      | CD5  | FSTL3   | S1T1            |
|     | EPHA2     | FCRLB| DLI1        | FCR-LB      | CD23       | MIC-A/B| EPHA2   | PPP1R9B         |
|     | DSC2      | CD40| DSC2        | SEMA7A      | CD5        | FGF-23| SRPK2   | LY75            |
|     | DSC2      | CD40| LAG3        | IL12RB1     | WISP-1     | CD244 | HTRA2   | TMSB10          |
|     | DSC2      | CD40| CD97        | CLEC4C      | CD27       | IL18  | SCF*    | IFRF            |
|     | ADAM8     | CXL9| HAVCR2      | MILR1       | CST5       | CD160 | TNFRSF4 | TXLNA           |
|     | IL-15RA   | ILT8| IL-15RA     | ILT8        | CD209      | CD209 | TNFRSF4 | CLL4            |
|     | TNFRSF9   | SH2D1A| TNFRSF9   | SH2D1A      | SLAMF1     | WFD2C | PILRA   | ICOSLG          |
|     | TNFRSF19  | GALNT3| IL-10RB   | CD83        | Pdgfrb     |       |        |                 |
|     | TNFRSF19  | GALNT3| IL-10RB   | CD83        | FGF-23     |       |        |                 |

*inverse association; Beta 2-Microglobulin (β2M); lactate dehydrogenase (LDH); peripheral B lymphocyte number (pblymph)
Supplementary Table E6. Proteomic markers selected by 5 folds cross-validated Lasso

| Proteomic marker | Coefficient at lambda.min |
|------------------|---------------------------|
| LY9              | 0.1757                    |
| ITM2A            | 0.1749                    |
| IFNLR1           | 0.1025                    |
| TPP1             | 0.0968                    |
| SPINT1           | 0.0913                    |
| CLEC4C           | 0.0325                    |
| MCP-1            | 0.0056                    |
| CLEC7A           | -0.2875                   |

Lambda minimum: 0.21
Alpha: 1

The lambda.min refers to value of $\lambda$ at the lowest cross-validation error.

Supplementary Figure E1. Kaplan-Meier curves for EFS related to well-known CLL prognostic factors.
Supplementary Figure E2. Molecular functions, biological process, and cellular components related to 14 proteins (with P<0.05 in adjusted Cox model and/or Lasso model) using the Protein Analysis Through Evolutionary Relationships (PANTHER) classification system.