RPPA-based proteomics recognizes distinct epigenetic signatures in chronic lymphocytic leukemia with clinical consequences

van Dijk, Anneke D; Griffen, Ti'ara L; Qiu, Yihua H; Hoff, Fieke W; Toro, Endurance; Ruiz, Kevin; Ruvolo, Peter P; Lillard, James W; de Bont, Eveline S J M; Burger, Jan A

Published in:
Leukemia

DOI:
10.1038/s41375-021-01438-4

Publication date:
2022

ARTICLE

CHRONIC LYMPHOCYTIC LEUKEMIA

RPPA-based proteomics recognizes distinct epigenetic signatures in chronic lymphocytic leukemia with clinical consequences

Anneke D. van Dijk[1,5]✉, Ti'ara L. Griffen[2,5], Yihua H. Qiu[3], Fieke W. Hoff[1], Endurance Toro[3], Kevin Ruiz[3], Peter P. Ruvolo[4], James W. LillardJr[2], Eveline S. J. M. de Bont[1], Jan A. Burger[3], William Wierda[3] and Steven M. Kornblau[3]

The chronic lymphocytic leukemia (CLL) armamentarium has evolved significantly, with novel therapies that inhibit Bruton Tyrosine Kinase, PI3K delta and/or the BCL2 protein improving outcomes. Still, the clinical course of CLL patients is highly variable and most previously recognized prognostic features lack the capacity to predict response to modern treatments indicating the need for new prognostic markers. In this study, we identified four epigenetically distinct proteomic signatures of a large cohort of CLL and related diseases derived samples ($n = 871$) using reverse phase protein array technology. These signatures are associated with clinical features including age, cytogenetic abnormalities [trisomy 12, del(13q) and del(17p)], immunoglobulin heavy-chain locus (IGHV) mutational load, ZAP-70 status, Binet and Rai staging as well as with the outcome measures of time to treatment and overall survival. Protein signature membership was identified as predictive marker for overall survival regardless of other clinical features. Among the analyzed epigenetic proteins, EZH2, HDAC6, and loss of H3K27me3 levels were the most independently associated with poor survival. These findings demonstrate that proteomic based epigenetic biomarkers can be used to better classify CLL patients and provide therapeutic guidance.

Leukemia; https://doi.org/10.1038/s41375-021-01438-4

INTRODUCTION

Chronic lymphocytic leukemia (CLL), the most common hematological malignancy, is characterized by the clonal expansion of B-lymphocytes in the bone marrow, peripheral blood, spleen, and other lymphoid organs [1]. The time until first treatment (TTFT) is highly variable, ranging from immediately after diagnosis to decades later [2]. Traditional prognostic variables and models have utilized immunoglobulin heavy chain variable region (IGHV) mutational, cytogenetic aberrations, and protein tyrosine kinase ZAP-70 expression [3] More favorable prognosis is associated with IGHV mutation, del[13q], trisomy 12, and/or low ZAP-70 expression while unmutated IGHV, del[11q], del[17p]/TP53 mutation, and high ZAP-70 are unfavorable. The international prognosis score for early-stage CLL (IPS-E) depends on lymphocytosis, nodal involvement, and IGHV status [4]. The development of therapies that interfere with B-cell receptor signaling via Bruton Tyrosine Kinase (BTK) and PI3K delta inhibitors (ibrutinib and idelalisib) along with the addition of BCL2 inhibitors (venetoclax) has significantly added to the CLL armamentarium, altered therapy patterns, and improved outcome in CLL. But, as with all modalities, there are non-responders, and resistance can emerge. Most previously recognized prognostic features have limited predictive capacity for modern therapies, so identifying new predictive markers for these new therapies are needed.

It is hypothesized that altered regulation of epigenetics, either DNA methylation or histone modification, contributes to the development, progression, and prognosis of CLL. CLL cells harbor a similar methylome as other cancers with global hypomethylation promoting genome instability and local promoter hypermethylation causing gene silencing [5–7]. However, global DNA methylation profiles in CLL have been relatively stable over time in both the indolent and proliferative compartments and pre- and post-therapy, suggesting that recurrent methylation changes occur early in leukemogenesis [8]. Indeed, a more recent discovery is that the putative cellular origin of CLL can be tracked via epigenetic biomarkers and this appears to be a strong predictor for TTFT [9, 10]. Queiros et al. defined three epigenetic subgroups based on 5 CpG regions that predicted prognosis more accurately than IGHV status [10]. Another study revealed that CLL progression depends on aberrant DNA methylation at CpG sites near the polycomb 2 repressive complex 2 (PRC2) target genes concurrent with an increase in PRC2 activity [11], and not on genetic clonal evolution.

[1]Department of Pediatric Oncology/Hematology, University Medical Center Groningen, Groningen, the Netherlands. [2]Department of Microbiology, Biochemistry, and Immunology, Morehouse School of Medicine, Atlanta, GA, USA. [3]Department of Leukemia, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. [4]Department of Molecular Hematology and Therapy, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. [5]These authors contributed equally: Anneke D. van Dijk, Ti'ara L. Griffen. ✉email: a.d.van.dijk@umcg.nl

Received: 20 May 2021 Revised: 11 September 2021 Accepted: 21 September 2021
Published online: 08 October 2021

Histone modification involves post-translational modifications of histone proteins that impact chromatin structure and gene transcription. In CLL, dysregulated histone modification is linked to polycomb repressor complex 2 (PCR2) with defective overexpression of subunit EZH2 creating a potential therapeutic target [12–14]. These findings suggest that both DNA methylation and histone modification alternations are involved in CLL development, progression, and prognostication. To evaluate whether epigenetic profiles have prognostic or therapeutic potential, we used proteomic measurement of histone and chromatin-modifying enzymes (HME) and histone methylation marks (HMM) to identify epigenetic signatures in CLL. Previously, we found that high-HME protein levels predict adverse outcomes in adults [15] and pediatric acute myeloid leukemia (AML, manuscript submitted), and that loss of H3K27me3 is an independent poor prognostic factor in AML [16]. In this study, we show that simultaneous quantitative analysis of 37 epigenetic proteins in 871 CLL and related diseases patient derived samples by reverse phase protein array (RPPA) revealed four proteomic signatures with distinct clinical and biological features and prognostic effects on the outcome.

MATERIAL AND METHODS

Patient Population

Peripheral blood (PB, $n = 746$) and bone marrow (BM, $n = 52$) samples were collected from 798 patients with untreated CLL and 73 PB samples from patients with CLL related diseases including prolymphocytic leukemia (PLL) B cell ($n = 4$) and T-cell ($n = 16$) and Richter's transformation ($n = 5$) or mature small B-cell lymphomas (MSBL)-like malignancies including hairy cell leukemia (HCL, $n = 12$), HCL variant (HCLV, $n = 4$), large granular lymphocytic leukemia (LGLT, $n = 4$), monoclonal B-lymphocytosis (MBL, $n = 4$), mantle cell lymphoma (MCL, $n = 12$), marginal zone lymphoma (MZL, $n = 11$) and splenic lymphoma with villous lymphocytes (SLVL, $n = 1$). All patients were seen at the University of Texas M.D. Anderson Cancer Center (MDACC) between 2005 and 2019, and were acquired during routine diagnostic assessments in accordance with regulations and protocols (Lab 01-473, Lab 03-0893, Lab 04-0678, Lab 08-0431, and Lab 07-0719) approved by the MDACC Investigational Review Board (IRB). Informed consent was obtained in accordance with the Declaration of Helsinki.

Sample processing

Fresh ($n = 127$) samples were placed on ice and processed to yield mononuclear cells by Ficoll separation within 2 h of collection. Frozen samples ($n = 744$) were initially processed similarly but after thawing, placed in fresh media and Ficoll separated, to isolate viable mononuclear cells. A total of median 18% of the samples was discarded by the Ficoll separation including neutrophils, basophils, eosinophils, and red blood cells. Of the remaining 82%, there was a median of 2% of monocytes and 80% of lymphocytes. Purified samples were then normalized to 10,000 cells/μL and whole cells lysates prepared as described previously [17].

RPPA

A custom RPPA was generated as previously described [17–19]. Patient lysates were printed onto slides along with normalization and expression controls in five serial dilutions. Protein lysis buffer (Biorad Lamelli buffer) was the negative control and a mixture of 11 leukemia cell lines were included as positive controls. Additionally, 5 normal $CD19^+$ samples were printed as normalization controls. Slides were probed with 384 strictly validated primary antibodies, then a secondary antibody conjugated to an infrared. Among the 384 antibodies utilized were targeting known histone and chromatin modification enzymes (HME) including ASH2L, BMI1, BRD4, CBX7, DNMT1, EZH2, HDAC1, HDAC2, HDAC3, HDAC6, hnRNPK, JMJD6, KAT2A, KDM1A, KMT2A, KMT2D, MEN1, NCL, NPM1, PAK1, SETD1A, SETD1B, SIRT1, SUZ12, WDR5, WTAP, and XPO1 along with antibodies recognizing total or post-translational methylated H3 or marks (HMM): H3K4me1, H3K4me2, H3K4me3, H3K9me2, H3K27me3, H3K36me3, and the histone variant H2AX phosphorylated on serine 139 (H2AX.pSer139) and 140 (H2AX.pSer140) and ubiquitinated histone H2B (HIST1H2B.Ub). Stained slides were analyzed using Microvigene® software (Version 3.4, Vigene Tech, Carlisle, MA). Protein expression levels were normalized relative to the mean expression on each array. To identify recurrent patterns of HME ($n = 27$) and HMM ($n = 10$) in CLL, an optimal number of protein clusters were determined by the Progeny clustering algorithm [20] coupled with k-means [21]. The measure of cluster stability was based on a co-occurrence probability matrix capturing true and false classifications. A cluster was defined as low or high HME/HMM if these were globally and relatively lower or higher expressed compared to their expression in normal CD19+, respectively. Samples did not cluster separately based on source (BM vs. PB).

Statistical analysis

Associations between protein clusters and clinical variables were calculated using Chi-Square and Fisher's exact test for categorical variables and the Kruskal–Wallis test for continuous variables. Survival curves were generated using the Kaplan–Meier estimator [22].

RESULTS

Characterization of four epigenetic-directed proteomic signatures in CLL

We analyzed the expression of 37 epigenetic regulating proteins in 871 samples derived from patients with CLL and related diseases. After performing the progeny cluster algorithm [20], patients were classified into four epigenetic proteomic signatures (ProSig) (Fig. 1A, top annotation). Proteins were clustered into eight correlated constellations (colored regions *y*-axis Fig. 1A). All ProSig shared lower expression of HIST1H2B ubiquitination, hnRNPK, H3K4me3, HDAC1 (violet constellation) H2AX.pSer140 (light blue constellation), and JMJD6 compared to normal $CD19^+$ B cells. ProSig1 (pink) is distinguished by lower expression of most HME (red, yellow, and light green constellations) and by lower expression of the most HMM (dark green, light and dark blue, and violet constellations), but relative overexpression of PAK1 (dark red), EZH2, DNMT1, HDAC6 (orange constellation). The other three signatures had a relatively uniform expression of most HME (yellow and light green constellations). ProSig2 (yellow) is distinguished by lower levels of most HMM (dark green and blue constellation) and higher levels of PAK1, ASH2L, KAT2A, BRD4, EZH2, DNMT1, and HDAC6 (dark red, red, and orange constellations). Patients in ProSig3 (green) displayed high expression of CBX7 and the HMM (light green and dark blue constellations) compared to the others. ProSig4 (red) had the lowest expression of several HME: ASH2L, EZH2, DNMT1, HDAC6, KMT2D, WDR5, BMI1 (orange, light and dark green constellations) and relatively high-HMM, except for H3K9me2, H2AX.pSer139 and H2AX.pSer140. Correlations between the 37 proteins are shown in Fig. 1B by hierarchical clustering.

Most diagnoses were found in all four ProSig with the exception that 15/16 cases of HCL/HCLV and all cases of LGLT were exclusively seen in ProSig1. Unsurprisingly, the "pre-CLL" state of MBL, and the PB samples from Richter's transformed patients, likely comprised of circulating non-transformed B-CLL cells, were seen intermixed with the CLL cases. ProSig1 also had disproportionally more MCL and T-PLL cases. CLL was the dominant diagnosis in ProSig2 (45% of all cases, 94% CLL) and ProSig3 (42% of all cases, 96% CLL), referred to hereafter as "bulk CLL". These two clusters were characterized by high HME and low-HMM ("low") or high-HMM ("high") respectively. In contrast, only 9% of CLL cases were found in ProSig1 ("low-HME; low-HMM") and only 5% in ProSig4 ("low-HME; high-HMM") (Table 1).

Patient characteristics differ based on the proteomic epigenetic classification

ProSig1 was enriched (44.7%) for non-CLL MSBL variant diagnoses, including almost all HCL(V) and LGLT and a non-proportional number of cases of the MZL, and T-cell PLL cases (Table 1). CLL ($n = 798$) patient characteristics were compared between the proteomic signatures. Demographically, ProSig did not differ in age, gender or race. Almost half of ProSig1 patients had more advanced Binet stage C and Rai stage III-IV (high-risk) disease at

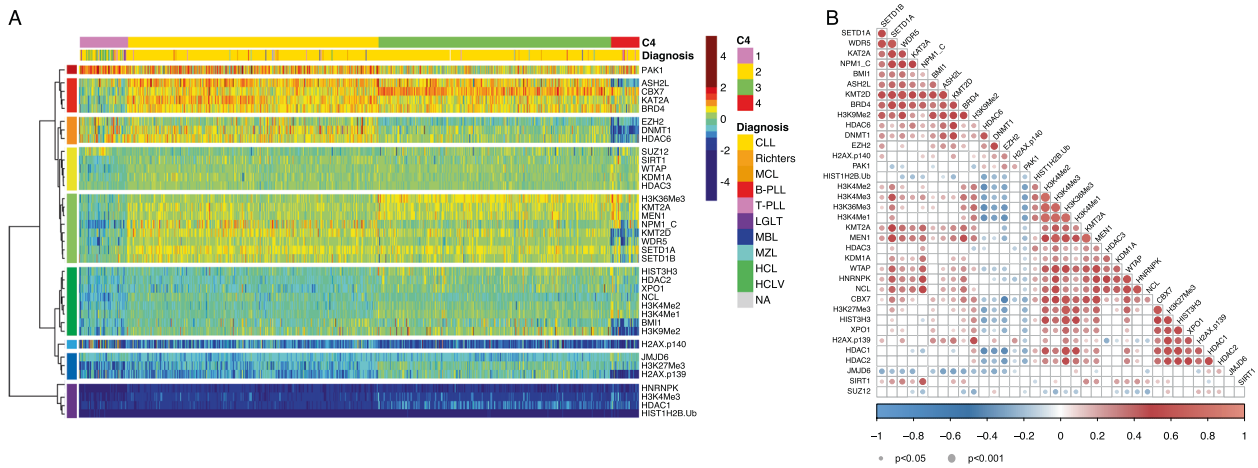

Fig. 1 HME and HMM are expressed in proteomic signatures in CLL. A RPPA-based heatmap showing expression of the 37 HME and HMM (vertical) from low (blue) to high (dark red) in 871 patients with CLL or MSBL-like malignancy (horizontal) relatively to $CD19^{+}$ cells. Patients were clustered in four clusters by the progeny clustering algorithm (pink, yellow and green, red). The same color scheme for Protein Signature membership is maintained throughout all subsequent figures. The proteins were clustered into nine constellations as shown by the colored boxes along the left *y*-axis (Dark red top, violet at bottom). **B** Correlation between the 37 HME and HMM by hierarchical clustering. Dot color represents Pearson's correlation coefficient and dot size was linked to *p*-value. Smaller *p*-value correlated with bigger dot and vice versa. Red coloration denotes positive correlation, blue negative correlation.

diagnosis; significantly higher than the overall study population (Binet 28% and Rai 30%) or the other clusters (both $p = 0.01$). Furthermore, ProSig1 had the lowest PB absolute lymphocyte count, BM lymphocytes percentage, and lower cell surface expression of CD19, CD23, and CD38, but higher CD22 (Table 1). This cluster contains more trisomy 12 cases (25%, $p < 0.001$) and chromosome nine abnormalities (8%, $p = 0.01$).

The ProSig4 patients (low-HME, high-HMM) had the highest number of circulating lymphocytes at presentation and 71% were classified as Binet stage A. The bulk CLL signatures ProSig2 and ProSig3 differed with respect to ZAP-70 positivity at 64% vs. 40%, respectively. In contrast, ProSig3 patients were more likely to have mutated IGHV (59%) compared to ProSig2 (42%).

CLL patients with MBSL-like signature have poor outcomes

Among the MSBL, HCL, and HCLV, LGLT are prognostically better and PLL and MCL prognostically worse than CLL (Supplementary Fig. 1A and as described earlier [1, 23]). Therefore, to ascertain the effect of HME/HMM on outcome in CLL, outcome analysis was performed with CLL patients only. Although there was no difference in the percentage of CLL patients that eventually received first or second treatment, patients in ProSig1 had a significantly shorter mean TTFT of 3 years compared to 4.2, 6, and 5 years in ProSig2, 3, and 4, respectively (Table 1) as well as the shortest TTST (2.5 years vs 3.8 (ProSig2), 5 (ProSig3) and 3.5 (ProSig4) years) and mean overall survival (OS, 4.7 years vs 6.4 (ProSig2), 8.5 (ProSig3), and 6.4 (ProSig4) years). Median survival was not reached for patients in ProSig4, but was 12.5 years, 20.3 years, and 35.1 years for ProSig 1, 2, and 3, respectively (Fig. 2A, $p < 0.0001$). Also, TTFT was significantly shorter for patients within ProSig1 (median 5.5 years) compared to ProSig2 (median 6.7 years), 3 (median 9.7 years) or 4 (median 9.4 years) (Fig. 2B, $p = 0.007$). Outcome comparison between the two bulk CLL clusters revealed adverse outcome for those in ProSig2 which has lower HMM. Within ProSig1, as expected the PLL and MCL diagnoses had the worst OS compared to CLL, HCL(V), or other MSBL diagnoses ($p < 0.001$, Supplementary Fig. 1B).

Protein cluster membership associates with survival and second treatment probability per treatment regime

Most CLL patients are initially managed using a watch and wait (WAW) approach, receiving therapy upon progression. Standard CLL therapy has been rapidly evolving since the introduction of Bruton's tyrosine kinase inhibitors (BTKi) and more recently the availability of Bcl-2 blockade with venetoclax. Previously, the standard was chemoimmunotherapy with fludarabine, cyclophosphamide, and rituximab (FCR) [24]. To examine whether the epigenetic proteomic signatures were prognostic for both old and new treatment regimes, we analyzed outcomes based on treatment: chemotherapy-, BTKi- and antibody/targeted-based. A total of 322 of the 798 CLL patients received treatment (40%), 175 received FCR or other chemotherapeutics ("Chemo"), 110 had BTKi-based (Ibrutinib ± venetoclax or ± Rituximab/Nivolumab, "BTKi") and 46 patients were treated with an antibody/targeted-based treatment plan ("Antibody"). Eighteen patients were classified in both the "Chemo" and "BTKi" groups as they received ibrutinib, fludarabine, cyclophosphamide, and Obinutuzumab (iFCG) and nine patients were excluded as other therapeutics were given (e.g. anti-cancer vaccines or CAR-T). We analyzed OS after the first treatment (OS-post-Tx) and TTST per treatment regime stratified by ProSig membership. Although the case numbers are small, we observed that patients in ProSig1 had the shortest OS-post-Tx after BTKi treatment (median 2 years vs. not reached, $p < 0.0001$, Fig. 3B). Notably, the ProSig1 OS-post-Tx was longer with chemotherapy-based treatments (median 8.9 years) than after BTKi, but was still shorter than other clusters (median 11.4 years in ProSig2, 17.4 years in ProSig3 and not reached in ProSig4, Fig. 3A). Patients in ProSig4 with low-HME and high HME all needed second therapy relatively shortly after antibody/targeted therapy (median TTST of 3.3 years, Fig. 3F). Overall, bulk CLL ProSig2 with lower HMM had worse OS and TTFT (yellow curve) compared to ProSig3 with higher expressed HMM (green curve) (Fig. 2). ProSig2 had shorter OS-post-Tx compared to ProSig3 after chemo, but did equally well after BTKi (Fig. 3A, B), ProSig2 OS-post-Tx and TTST were slightly inferior to ProSig3 for antibody-based therapy as well, but numbers are very small. Thus, lower HMM is linked to poor prognosis after chemotherapy-based treatment but not after BTK inhibitors. Previously, studies have shown that ibrutinib-treated patients with del[11q] had a significantly longer progression-free survival (PFS) than ibrutinib-treated patients without del[11q] and that chlorambucil-treated patients with del[11q] had a significantly shorter PFS than chlorambucil-treated patients without del[11q] [25, 26]. Therefore, we compared OS-post-Tx in chemotherapy and BTKi treated

Table 1. Patient characteristics stratified by cluster membership.

Variable		All pts, $n = 871$	C1, $n = 75$	C2, $n = 390$	C3, $n = 362$	C4, $n = 44$	***P*-value**
		100%	8.6%	44.8%	41.6%	5.1%	
Diagnosis	CLL	798	41	368	351	38	**2.2E-16**
	Richters	5	2	1	2	0	
	B-PLL	4	0	2	1	1	
	T-PLL	16	5	6	1	4	
	HCL	12	11	1	0	0	
	HCLV	4	4	0	0	0	
	MCL	12	3	6	2	1	
	MZL	11	5	3	3	0	
	LGLT	4	4	0	0	0	
	MBL	4	0	3	1	0	
	SLVL	1	0	0	1	0	
Variable		CLL pts, $n =$ 798	C1, $n = 41$	C2, $n = 368$	C3, $n = 351$	C4, $n = 38$	***P*-value**
		100%	5.1%	46.1%	44.0%	4.8%	
Age (mean ± SD years)		64.7 ± 10.2	62.4 ± 10.6	64.4 ± 10.1	65.5 ± 10.2	62.8 ± 11.4	0.10
Gender	Female	310/798 (39%)	23/41 (56%)	138/ 368 (38%)	133/ 351 (38%)	16/38 (42%)	0.13
	Male	488/798 (61%)	18/51 (44%)	230/ 368 (63%)	218/ 351 (62%)	22/38 (58%)	
Race	White	713/775 (92%)	35/40 (88%)	330/ 357 (92%)	315/ 343 (92%)	33/35 (94%)	0.74
	Black	33/775 (4%)	3/40 (8%)	16/357 (4%)	14/343 (4%)	0/35 (0%)	
	Hispanic	22/775 (3%)	1/40 (3%)	3/357 (1%)	10/343 (3%)	2/35 (6%)	
	Asian	7/775 (1%)	1/40 (3%)	2 (1%)	4/343 (1%)	0/35 (0%)	
Binet stage	A	489/784 (62%)	17/37 (46%)	213/ 363 (59%)	232/ 346 (67%)	27/38 (71%)	**0.01**
	B	75/784 (10%)	2/37 (5%)	42/363 (12%)	30/346 (9%)	1/38 (3%)	
	C	220/784 (28%)	18/37 (49%)	108/ 363 (30%)	84/346 (24%)	10/38 (26%)	
Rai stage	0	268/784 (34%)	9/37 (24%)	106/ 363 (29%)	136/ 346 (39%)	17/38 (45%)	**0.01**
	I-II	281/784 (36%)	10/37 (27%)	137/ 363 (38%)	123/ 346 (36%)	11/38 (29%)	
	III-IV	235/784 (30%)	18/37 (49%)	120/ 363 (33%)	87/346 (25%)	10/38 (26%)	
Lymphocytes (x 10^3/uL)		75.8 ± 17.7	63 ± 19.9	75.4 ± 18	77.5 ± 17	78.2 ± 12.7	**9.1E-06**
Hemoglobin ()		13.5 ± 1.8	13.7 ± 1.8	13.3 ± 1.7	13.6 ± 1.8	13.4 ± 2	0.24
Platelets (x 10^9/L)		191 ± 72	227 ± 118	181 ± 70	197 ± 66	204 ± 72	**1.1E-04**
LDH		483 ± 292	550 ± 263	484 ± 378	471 ± 182	499 ± 163	0.42
BM lymphocytes (%)		61 ± 23	49 ± 25	66 ± 21	58 ± 23	59 ± 21	**4.9E-05**
Chromosomal abnormality	None	165/714 (23%)	9/39 (23%)	60/331 (18%)	90/308 (29%)	6/36 (17%)	**8.0E-03**
	Del[11q]	100/714 (14%)	3/39 (8%)	46/331 (14%)	46/308 (15%)	5/36 (14%)	0.73
	Trisomy 12	110/714 (15%)	10/39 (25%)	72/331 (22%)	23/308 (7%)	5/36 (14%)	**9.7E-07**
	Del[13q]	274/714 (38%)	13/39 (33%)	113/ 331 (34%)	133/ 308 (43%)	15/36 (42%)	0.11
	Del[17p]	69/714 (10%)	4/39 (10%)	43/331 (13%)	17/308 (6%)	5/36 (14%)	**0.01**
	Chromosome 9	16/714 (2%)	3/39 (8%)	10/331 (3%)	2/308 (1%)	1/36 (3%)	**0.01**
IGHV status	Mutated	297/579 (51%)	15/24 (63%)	115/ 273 (42%)	149/ 253 (59%)	18/29 (62%)	**4.9E-04**
SF3B1	Mutated	35/212 (17%)	3/13 (23%)	23/127 (18%)	7/63 (11%)	2/9 (22%)	0.46
ZAP-70	Positive	191/378 (51%)	12/26 (46%)	106/ 166 (64%)	66/166 (40%)	7/20 (35%)	**6.4E-05**

Table 1 continued

CD38 (mean ± SD)		23.9 ± 27.1	21.7 ± 26.5	27.9 ± 28.3	20.2 ± 25.2	27.3 ± 33.3	**0.04**
CD19 (mean ± SD)		81.5 ± 18.6	62.4 ± 20.3	82.8 ± 12.7	81.2 ± 13.8	76.8 ± 22.5	**2.4E-09**
CD20 (mean ± SD)		77.8 ± 20.2	81 ± 23.7	77.7 ± 20.9	77.7 ± 18.8	76.4 ± 22.8	0.82
CD22 (mean ± SD)		63.3 ± 39.4	76.9 ± 35.6	69.7 ± 37.5	54.4 ± 40.1	72.2 ± 38.4	**6.5E-06**
CD23 (mean ± SD)		87.1 ± 17.8	79.3 ± 27.6	88.5 ± 16.5	86.5 ± 17.4	87.2 ± 19.8	**0.048**
CD79b (mean ± SD)		42.8 ± 37.9	46.7 ± 39	48.4 ± 44	36.8 ± 30	42.6 ± 34.3	**3.4E-03**
First treatment	Yes	322/798 (40%)	15/41 (37%)	154/ 368 (42%)	141/ 351 (40%)	12/38 (32%)	0.63
	TTFT (mean ± SD years)	4.9 ± 5.0	3.0 ± 2.8	4.2 ± 4.6	6.0 ± 5.4	5.0 ± 5.3	**1.4E-06**
	Response	165/218 (76%)	7/7 (100%)	82/108 (76%)	73/97 (75%)	3/6 (50%)	**0.23**
Second treatment	Yes	92/320 (29%)	7/15 (47%)	49/153 (32%)	33/140 (24%)	3/12 (25%)	0.16
	TTST (mean ± SD years)	4.3 ± 3.2	2.5 ± 2.1	3.8 ± 2.8	5.0 ± 3.6	3.5 ± 3.2	**1.1E-03**
Vital status	Alive	708/798 (87%)	29/41 (71%)	321/ 368 (87%)	322/ 351 (92%)	36/38 (95%)	**9.8E-04**
	OS (mean ± SD years)	7.2 ± 5.8	4.7 ± 4.3	6.4 ± 5.5	8.5 ± 6.1	6.4 ± 5.8	**4.0E-07**

p-values in bold represents significant difference among the clusters.

patients between those with and those without del[11q] and observed a similar trend (Supplementary Fig. 2A, B). Additionally, we performed the survival analysis with ProSig stratification per treatment for patients with and without del[11q] separately. Similar outcome trends were observed based on signatures in both chemotherapy and BTKi treated patients regardless of del [11q] (Supplementary Fig. 2B–F).

Cluster membership independently predicts overall survival and time to treatment

The four HME and HMM derived ProSig behave biologically and prognostically different. To query whether ProSig membership is an independent predictor for outcome in CLL, we performed Cox regression analysis for OS and TTFT. In univariate analysis, gender, IGHV mutational status, ZAP-70 status, chromosomal abnormalities, Rai staging, and ProSig were predictive but Binet staging, SF3B1, del[13q], chromosomal 9 abnormality were not. In the multivariate model, for the 277 cases with all variables available, ProSig1 membership was the only independent predictor of survival (Table 2). Proteomic signature did not predict in uni- or multivariable analysis of TTFT.

Identification of an individual protein that are clinical predictors of outcome in CLL

We next looked for epigenetics markers differentially expressed between ProSig. First, we focused on individual protein expressions that were differentially expressed relatively to $CD19^{+}$ among the signatures. ProSig1 was for example represented by low MEN1, KMT2A, and SETD1A, but high HDAC6 and EZH2 (Fig. 4A–E). ProSig4 was the only cluster that had relatively low ASH2L and DNMT1 relative to $CD19^{+}$ normalization. ProSig1 and 4 an had opposite prognostic impact and we, therefore, sought to explain what biology might underlie these signatures. ProSig1 and 4 were mostly discriminated by HMM expression (Fig. 1, dark green constellation) of H3K4me1-2, H3K27me3, and H3K36me3 (Fig. 4I–K) suggesting potential importance. Across all four ProSig, both HMM expression and prognosis increased sequentially from ProSig 1 to 2, 3, and 4 (Fig. 4I–K), with the exception of H3K9me2, which was lowest in ProSig4 (Fig. 4L). To identify which proteins have significant predictive value in the whole CLL cohort based on their continuous expression, we created a Cox proportional hazard model (Supplementary Table 1). Among these, not H3K9me2, but H3K36me3, and H3K27me3 emerged as predictors for OS, TTFT, and TTST. Recently we observed that loss of H3K27me3 independently predicts inferior OS in adult AML patients [27] we, therefore, looked for a similar prognostic effect in CLL. Overall, H3K27me3 predicted shorter OS (HR = 0.33, 95% CI = 0.22–0.48, $p < 0.001$), TTFT (HR = 0.62, 95% CI = 0.49–0.78, $p < 0.001$) and TTST (HR = 0.35, 95% CI = 0.23–0.56, $p < 0.001$) (Supplementary Table 1). We then correlated H3K27me3 levels with clinical features, finding no association with age, gender, race, del[11q], normal cytogenetics, chromosome 9 abnormalities (data not shown), but did find that levels were lower in the traditionally adverse features of unmutated IGHV ($p = 0.0042$), positive ZAP-70 ($p = 0.0019$), del[17p] ($p = 0.031$), trisomy 12 ($p = 0.0023$), and higher in del[13q] ($p = 0.0054$) (Supplementary Fig. 3). On multivariate analysis, the impact of H3K27me3 (low/high based on median H3K27me3 expression [~ 1.37]) on TTFT was significant when accounting for gender, cytogenetic abnormalities (none, del [11q] and trisomy 12), IGHV, and ZAP-70 status and Rai risk group (p t= 0.004) and on TTST when accounting for IGHV and Rai risk group ($p = 0.002$) (Supplementary Table 2).

Characterizing functionally dysregulated protein networks by epigenetic signatures

We used Cytoscape for visualizing protein networks among all 384 antibodies utilized to identify differences in correlated protein networks (Fig. 5). This enables a more wholistic visualization of how the core HMM/HME (large nodes) proteins, as well as other correlated proteins (small nodes), change between the four signatures. This reveals that there are four protein constellations with correlated changes across the signatures. First, PAK1 expression levels were highest in ProSig1 and progressively decreased with increasing signature number. Many known PAK1 connections were present: PAK4, PDGFRB, PLK1, PTK2, PXN, RAF1, SRC, and SYK. Several previously unknown PAK1 connections were also revealed by the RPPA including negative correlations with H3K36me3, H3K4me1-3, BCL2 and BCL2L11, HDAC1-2, and CBX7 (Supplementary Fig. 4A). Opposite of PAK1, CBX7 increased by signature. The third set of KAT2A, ASH2L, and H3K36me3 comigrate across the protein networks. All three are low in ProSig1, but ASH2L and KAT2A are interconnectedly high in ProSig2, along with phosphorylated CDK1B on tyrosine 198 (CDK1B.pT198) and PRKAA1 on tyrosine 172 (PRKAA1.2.pT172) and MSI2. The fourth set of interconnected proteins BRD4, WTAP,

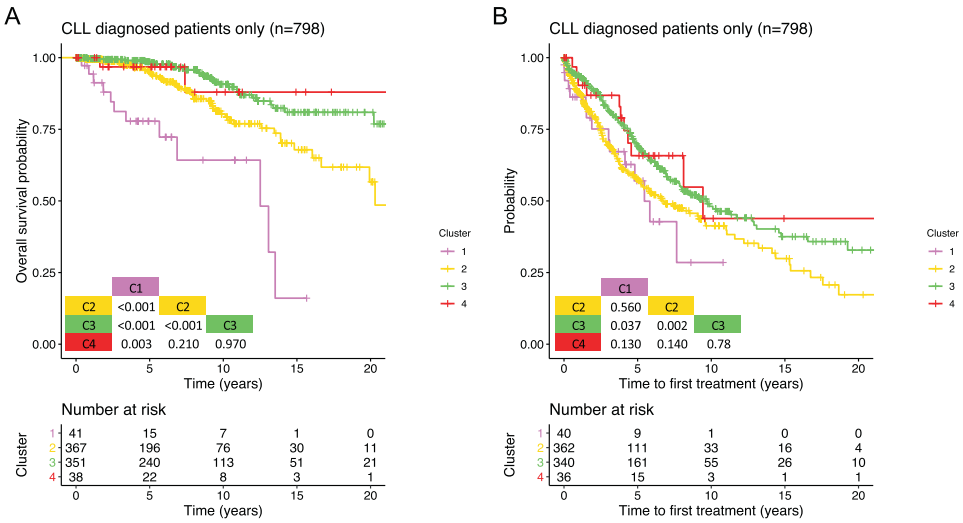

Fig. 2 Kaplan–Meier analysis of overall survival (OS) and time to first treatment (TTFT) per protein signature cluster for CLL patients. A OS and **B** TTFT of 773 CLL patients with available outcome data (773/798 = 97%) stratified by cluster membership (ProSig1 in pink; ProSig 2 in yellow; ProSig 3 in green and ProSig 4 in red). Data was shown for a follow-up period of 20 (OS) and (TTFT) years. Inset tables list *p*-values for all possible comparisons using the log-rank pairwise comparison test. The number of patients included at the different 5-year intervals is shown at the bottom.

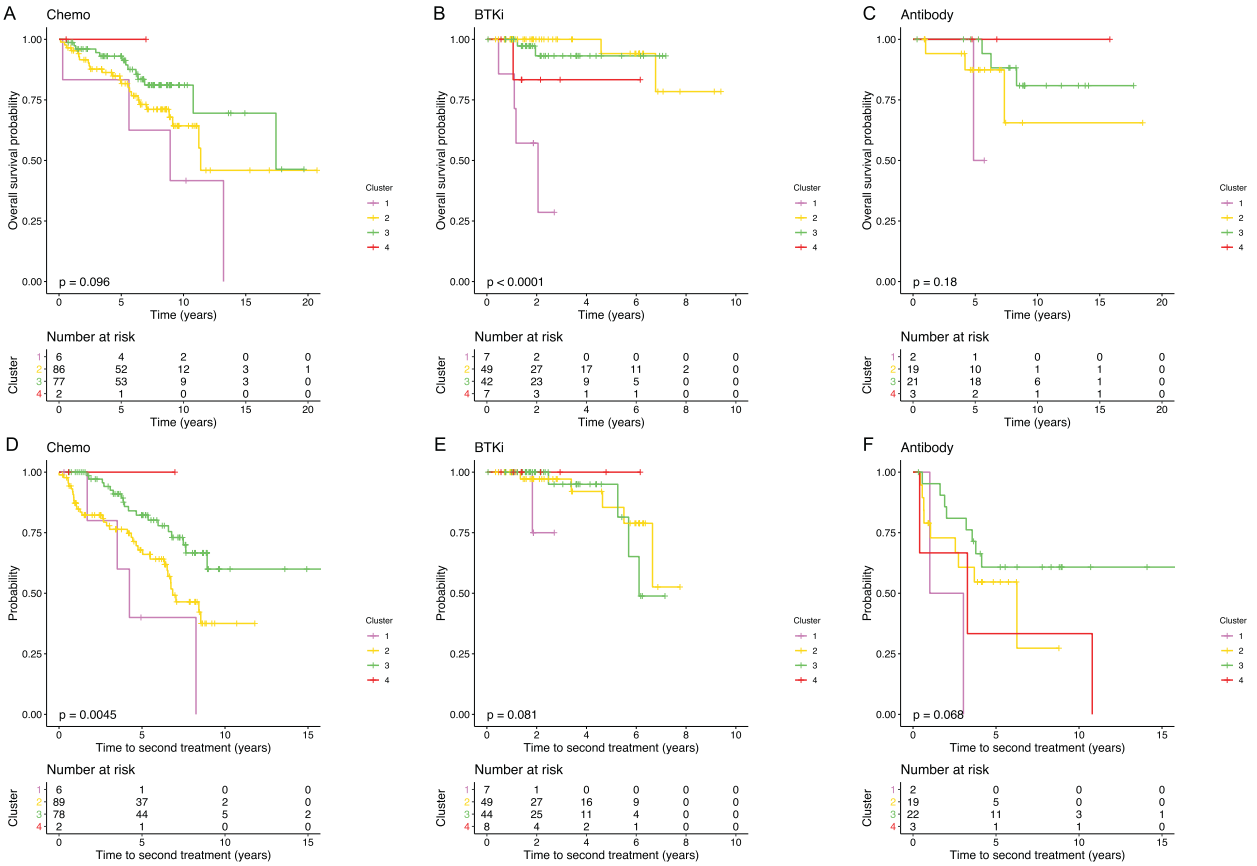

Fig. 3 Overall survival (OS) after first treatment and time to second treatment (TTST) durations after different treatment regimes according to cluster membership in CLL patients only. OS and TTST per cluster after **A** and **D** chemotherapy-based therapies (including FCR), **B** and **E** BTK inhibitors based treatment, and **C** and **F** antibody and targeted based therapies (excluding FCR). Cluster membership: ProSig1 in pink, ProSig 2 in yellow, ProSig 3 in green, and ProSig 4 in red. The number of patients included at the different 5-year intervals is shown at the bottom. *P*-value is for the four-way comparison.

Table 2. Univariate and multivariate Cox regression analysis of overall survival (OS) and time to first treatment (TTFT) in CLL patients.

		Univariate analysis			Multivariate analysis			Univariate analysis			Multivariate analysis		
Variable		**OS ($n = 798$, events $= 90$)**			**OS ($n = 277$, events $= 32$)**			**TTFT ($n = 778$, events $= 312$)**			**TTFT ($n = 277$, events $= 144$)**		
		HR	**95% CI**	**P-value**	**HR**	**95% CI**	**P-value**	**HR**	**95% CI**	**p**	**HR**	**95% CI**	**p**
Proteomic signature	1	1			1			1					
	2	0.28	0.15–0.53	**1.0E-04**	0.05	0.01–0.19	**7.4E-06**	0.84	1.15–1.44	0.53			
	3	0.12	0.06–0.24	**1.7E-09**	0.10	0.03–0.37	**5.2E-04**	0.59	0.34–1.00	0.05			
	4	0.12	0.03–0.54	**5.6E-03**	0.09	0.01–0.97	**0.048**	0.53	0.62–1.16	0.11			
Gender	Female	1			1			1					
	Male	1.71	1.09–2.769	**0.02**	1.66	0.68–4.03	0.26	1.11	0.88–1.40	0.37			
Chromosomal abnormalities	Any	1			1			1					
	None	0.50	0.329–0.88	**0.02**	0.89	0.31–2.60	0.84	0.90	0.68–1.19	0.46			
Del[11q]	Negative	1			1			1			1		
	Positive	1.76	1.01–3.05	**0.04**	0.72	0.22–2.43	0.60	2.03	1.50–2.73	**3.6E-06**	1.31	0.80–2.16	0.28
Trisomy 12	Negative	1			1			1			1		
	Positive	2.24	1.33–3.76	**2.3E-03**	1.95	0.70–5.44	0.20	1.55	1.14–2.11	**0.01**	1.15	0.70–1.90	0.59
Del[13q]	Negative	1						1			1		
	Positive	0.66	0.41–1.05	0.08				0.69	0.53 - 0.88	**3.6E-03**	0.88	0.58–1.33	0.54
Del[17p]	Negative	1			1			1					
	Positive	1.90	1.09–3.33	**0.02**	1.68	0.50–5.57	0.40	1.00	0.68–1.47	1.00			
Chromosome 9 abnormality	No	1						1					
	Yes	1.94	0.61–6.14	0.26				0.82	0.36–1.83	0.62			
IGHV status	Mutated	1			1			1			1		
	Unmutated	2.41	1.38–4.22	**2.0E-03**	1.96	0.75–5.11	0.17	2.67	2.02–3.52	**3.9E-12**	1.65	1.12–2.44	**0.01**
ZAP-70	Negative	1			1			1			1		
	Positive	2.84	1.60–5.07	**3.9E-04**	2.36	0.91–6.10	0.08	2.01	1.50–2.69	**3.0E-06**	1.38	0.93–2.03	0.11
SF3B1	Mutated	1						1					
	Unmutated	0.41	1.10–1.75	0.23				0.72	0.32–1.59	0.41			
Binet stage	A	1						1			1.00		
	B	1.22	0.65–2.29	0.55				2.66	1.94–3.63	**8.9E-10**	1.56	0.83–2.94	0.16
	C	1.58	0.97–2.58	0.07				1.49	1.15–1.92	**2.6E-03**	0.53	0.07–4.01	0.54
Rai Stage	0	1			1			1			1		
	I	1.17	0.67–2.04	0.58	0.62	0.19–2.02	0.43	1.76	1.31–2.37	**2.1E-04**	1.16	0.69–1.96	0.58
	II	1.68	0.74–3.79	0.21	0.78	0.17–3.48	0.74	3.14	2.06–4.78	**1.1E-07**	1.76	0.81–3.80	0.15
	III	1.39	0.65–2.96	0.39	0.71	0.22–2.28	0.57	1.51	1.04–2.21	**0.03**	1.93	0.26–14.51	0.52
	IV	2.23	1.16–4.28	**0.02**	1.26	0.47–3.35	0.65	2.14	1.49–3.09	**4.1E-05**	3.21	0.40–25.69	0.27

p-values in bold represents significant difference.

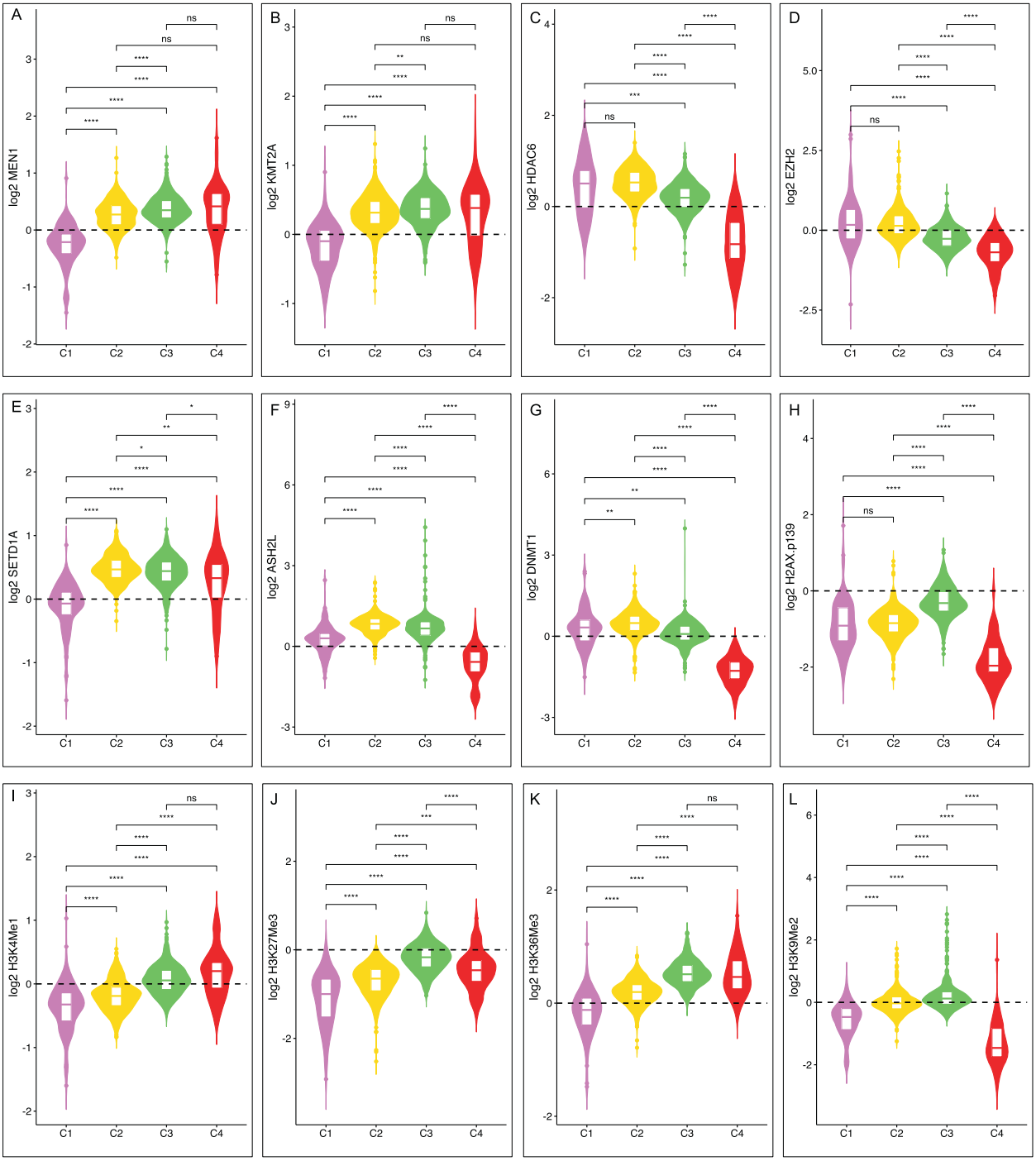

Fig. 4 Individual protein expression according to proteomic signature. A MEN1, **B** KMT2A, **C** HDAC6, **D** EZH2, **E** SETD1A, **F** ASH2L, **G** DNMT1, **H** H2AX.pSer139, **I** H3K4me1, **J** H3K27me3, **K** H3K36me3, and **L** H3K9me2 according to proteomic signature membership relative to $CD19^{+}$ normalization (dashed line). The inset box shows the 25–75% boundaries and the horizontal line in the middle shows the median expression. Statistical comparisons between each signature are shown above, with *p*-values listed as ns not significant, ***$p < 0.001$ and ****$p =$ 1e-04.

MEN1, and KMT2A, associated with LEF1 and RELA, was low in ProSig1, but this entire cluster was upregulated in ProSig2. Patients in ProSig2 were further characterized by high PAK1, ASH2L, KAT2A, BRD4, EZH2, DNMT1, and HDAC6. Based on correlation data, high ASH2L equals lower PI3K/Akt and SRC signaling but increased BCL2, WNT, and mTOR (Supplementary Fig. 4B). EZH2 was positively correlated with proteins involved in MAPK signaling, transcriptional activation, WNT, cell cycle, metabolism, NOTCH, chemokine, and ribosomal activity signaling (Supplementary Fig. 4C). To demonstrate changes across the clusters, we created a video that demonstrates the protein network as shown in Fig. 5 in transformation (Supplementary Video).

DISCUSSION

Predicting which CLL patients will require therapy and when it will be necessary remains challenging, especially, as modern therapy

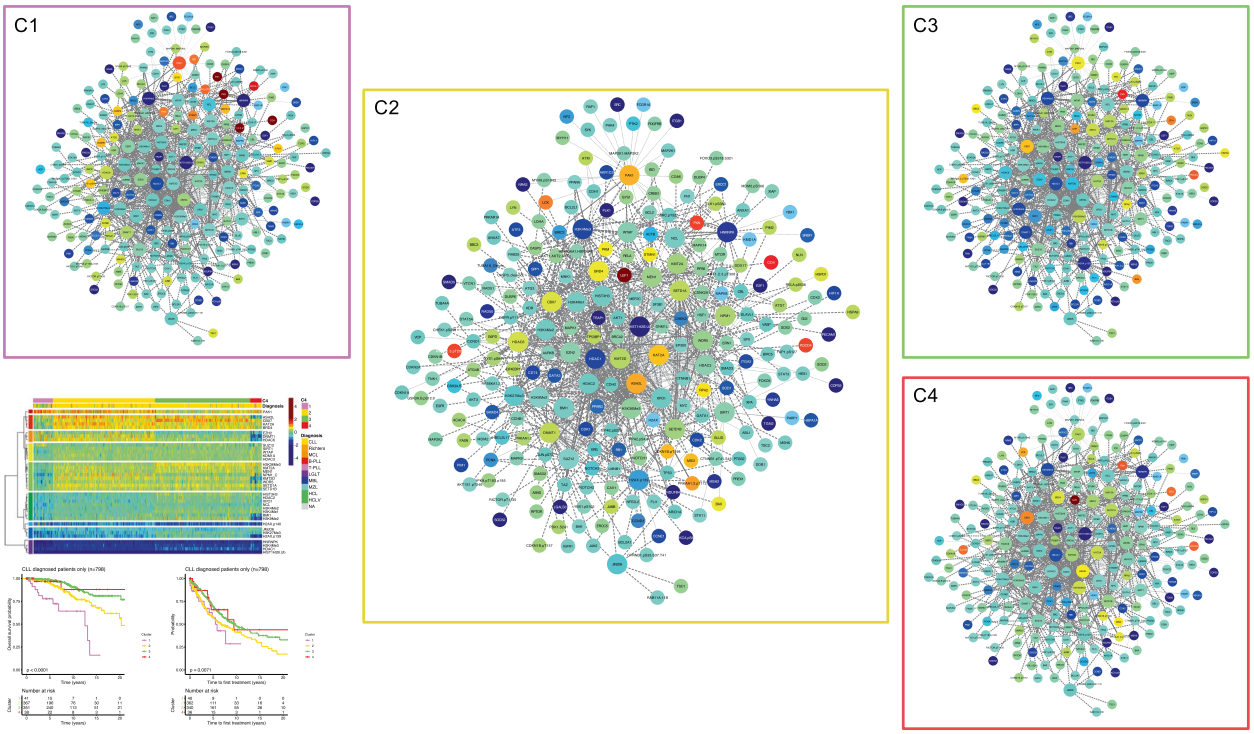

Fig. 5 Known protein–protein associations of the 384 proteins on the CLL RPPA, including 37 HME (big nodes). Node colors represent protein expression value relative to the protein expression in the CD19^{+} normal cells per cluster. Large nodes are the 37 HMM and HME proteins in this study, smaller nodes are the other 347 proteins studied on this array. Color intensity represents expression level and is on the same Log2 scale as in Fig. 1. Dashed lines reflect interactions identified from existing online databases (i.e. KEGG, String) or relationships identified in this dataset. Please also see the "Supplemental Video" for a larger format dynamic visualization of how protein expression changes between the protein signatures.

has rendered several historically useful prognostic markers as irrelevant. In this study, we identified four epigenetically distinct proteins signatures in CLL and CLL-like diagnoses that prognosticate TTFT, TTST, and OS with old and new therapies. ProSig1, characterized by the lowest levels of histone methylation, contained a high proportion of CLL-like diseases (46%) mixed with CLL (54%) and predicted the worst clinical behavior for the CLL patients independent of known risk factors and regardless of therapy choice. The majority of CLL cases were in ProSig2 and ProSig3, discriminated by expression of HMM and ASH2L, KAT2A, EZH2, DNMT1, and HDAC6. Those with lower HMM (ProSig2) had worse outcomes than those in ProSig3 with higher HMM, i.e. H3K27me3, H3K4me1-2, H2AX.pSer139, H3K36me3 to chemotherapy and antibody treatment, but this was ameliorated with BTK inhibition. ProSig4 with low ASH2L, EZH2, DNMT1, HDAC6, KMT2D, WDR5, BMI1, KMT2D, relatively high-HMM but low H3K9me2 did prognostically well, with chemotherapy or BTKi therapy, but not with antibody-based treatment.

The proteomic patterns of ProSig1 with markedly lower expression of SETD1A, KMT2A (also known as MLL1), MEN1 (menin), and other HMM provide insight into the biology underlying the poor clinical outcome of these patients. Both SETD1A and KMT2A are H3K4 methyltransferases and SETD1A deletion has been associated with a global reduction of H3K4me1-3 [28], a finding also observed in our study. SETD1A is required for B cell differentiation from progenitor to precursor B cell in the BM and maturation in the spleen and KMT2A and MEN1 are critical during B-cell maturation in the BM [29]. In the BM, SETD1A deficiency led to decreased H3K4me3 deposition and disrupted gene transcription of critical B-cell development regulators (e.g. Pax5 and Rag1-2, important for Ig heavy-chain rearrangement) [30, 31]. SETD1A knock-out mice evince a phenotype with enlarged spleen and leukocytopenia, concordant with our findings of lowest lymphocyte number and high Binet staging in ProSig1 patients. These results suggest that the ~ 5% of CLL patients with extreme loss of SETD1A, KMT2A, MEN1, and H3K4me3 may have disrupted B-cell maturation in the progenitor to precursor cell phase transition, and that this is an independent predictor for poor outcome regardless of the treatment given. Although ProSig1 had worst outcomes per treatment compared to the other signatures, patients may do better with chemo (median OS-post-Tx of 8.9 years) than with BTK inhibitors (median OS-post-Tx of 2 years).

Another key protein in early B-cell development is the histone methyltransferase EZH2 [32]. Previously, upregulated EZH2 mRNA levels have been identified as a pro-survival marker in CLL cells, indicating cell growth and poor outcome [12, 13]. In our study ProSig3 and Prosig4 with lower EZH2 expression had outstanding long-term survival, while EZH2 was overexpressed in poorer prognostic ProSig1 and ProSig2 patients. Additionally, lower levels of EZH2 have earlier been related to mutated IGHV CLL [12], a finding confirmed here as low EZH2 ProSig1 displayed a high percentage of mutated cases (62%, vs. 51% overall). Notably, EZH2 is part of the PRC2 complex that catalyzes H3K27 trimethylation, but we observed these proteins to be inversely correlated ($r = -0.39$, $p < 2.23\text{-}16$). Instead, H3K27me3 levels significantly associated with: HDAC1 ($r = 0.56$, $p < 2.2\text{e-}16$), HDAC2 ($r = 0.42$, $p < 2.2\text{e-}16$) and CBX7 ($r = 0.58$, $p < 2.2\text{e-}16$), a PcG-repressor complex-1 (PRC1) component that recognizes and binds to trimethylated H3K27 [33] We observe an identical strong positive correlation between EZH2 and H3K27me3 ($r = 0.49$, $p < 2.2\text{ep16}$) in a separate pediatric T-cell acute lymphoid leukemia RPPA dataset ($n = 358$ cases, unpublished) supporting this interrelation. These data suggest that in CLL patients EZH2 loss positively correlates with the outcome, but does not impact global H3K27me3 levels, which rather correlated with PRC1 components including CBX7.

Moreover, EZH2 correlated with proteins affecting transcription and translation indicating higher proliferative potential. DNMT1, which also correlated with EZH2 and HDAC6, correlated with KRAS-mediated hypermethylation of tumor suppressors, ultimately promoting growth.

ProSig4 demonstrated a global picture suggesting a loss of histone lysine methyltransferase activity in this signature with lower protein levels of KMT2D, DNMT1, ASH2L, HDAC6, and H3K9me2. Interestingly, selective HDAC6 inhibition has hitherto shown therapeutic efficacy in pre-clinical CLL models [34]. Maharaj et al. validated HDAC6 overexpression in CLL, similar to our dataset. They also showed that genetically silencing HDAC6 or pharmacological inhibition has led to downregulated BCR signaling, causing a proliferation defect and apoptosis. In our cohort, ProSig4 patients with lower HDAC6 expression were long-term survivors while ProSig2 with the highest HDAC6 expression had poorer outcomes, possibly identifying these patients as ideal candidates for HDAC6 inhibition.

The majority of CLL patients (90%) were in ProSig2 and ProSig3. These signatures differed based on clinical features with a higher percentage of trisomy 12, positive ZAP-70, and unmutated IGHV in ProSig2 compared to 3. Moreover, ProSig2 had significantly shorter median OS and TTST after chemotherapy-based treatment. Compared to ProSig3, these patients had low expression of the HMM, including H3K27me3, H3K36me3, and H2AX.pSer139. In AML, we have recently reported that global loss of H3K27me3 independently predicts poor prognosis in all and p53 mutated patients [16]. Similarly, we identified the H3K27me3 level in CLL as an independent clinical predictor for TTFT and TTST, regardless of IGHV status. Expressions of EZH2, DNMT1, HDAC6, and H4K27me3, but not ASH2L, also discriminates between bulk CLL signatures ProSig2 and ProSig3. EZH2 as part of PRC2 modulate DNA methylation via DNMT1 recruitment. DNMT1 is a methyltransferase that adds methyl groups to gene promotors that lead to epigenetic gene silencing. In myeloid malignancies, overexpression of DNMT1 has been linked to altered hypermethylated tumor suppressor genomic regions [35] HDAC6 is not directly interconnected to EZH2 or DNMT1 but has been linked via heat shock factor 90 (hsp90), which is a critical molecular chaperone for maintaining EZH2 stability. Pan-HDAC inhibition via Panobinostat treatment leads to decreased levels of HDAC6, subsequent heat shock factor 90 (hsp90) acetylation, and EZH2 depletion suggesting HDAC6 plays part in regulating hsp90 and indirectly EZH2 protein stability and DNMT1 recruitment [35, 36]. ProSig2 patients also overexpressed ASH2L which we found was associated with poor prognosis in adult AML [15]. Although ProSig2 patients had lower OS-post-Tx after chemotherapy (77% after 10 years), 95% survived 10 year OS-post-Tx when treated with BTK inhibitor-based regimes. We hypothesize that patients with the ProSig2 signature may be eligible candidates for ibrutinib with or without HDAC6 inhibitors, as synergy between these has already been reported in follicular lymphoma [37]. Based on the biological processes influenced by EZH2, DNMT1, and HDAC6, ProSig3 (and ProSig4) represent a more indolent CLL phenotype.

Over the past two decades, most studies of epigenetics in CLL focused on DNA methylation profiling, concluding that distinct methylome patterns that symbolize stable molecular marks exist in the CLL cell of origin, but lacked functional impact [10, 38]. The novelty of this study is shown by the simultaneous analysis of HME and HMM at the proteomic level. We show here for the first time that histone and chromatin-modifying proteins form patterns in CLL and that these correlate with patient characteristics and therapy-specific outcomes. Using RPPA, we observed that EZH2, DNMT1, and HDAC6, which individually yield prognostic value in our CLL cohort, were significantly positively correlated with each other. When levels of these three proteins and ASH2L are lower, CLL patients had outstanding overall survival rates and a median TTFT of more than 9 years. In contrast, in ProSig2, characterized by simultaneous overexpression of EZH2, DNMT1, and HDAC6, a poorer outcome was observed after chemotherapy, but that this is mitigated by BTK inhibition. Thus, the integrated proteomic signatures identify patients that are more (ProSig2) and less likely (ProSig1) to benefit from modern CLL therapy. We conclude that an analysis of epigenetic protein biomarkers is valuable for prognostication and treatment selection in CLL and complements prior genomic and methylomic CLL classifications.

REFERENCES

1. Rozman C, Montserrat E. Chronic lymphocytic leukemia. N. Engl J Med. 1995;333:1052–7.
2. Jain N, O'Brien S. Initial treatment of CLL: integrating biology and functional status. Blood. 2015;126:463–70.
3. Wierda WG, O'Brien S, Wang X, Faderl S, Ferrajoli A, Do KA, et al. Multivariable model for time to first treatment in patients with chronic lymphocytic leukemia. J Clin Oncol: Off J Am Soc Clin Oncol. 2011;29:4088–95.
4. Condoluci A, Terzi di Bergamo L, Langerbeins P, Hoechstetter MA, Herling CD, De Paoli L, et al. International prognostic score for asymptomatic early-stage chronic lymphocytic leukemia. Blood. 2020;135:1859–69.
5. Kanduri M, Cahill N, Göransson H, Enström C, Ryan F, Isaksson A, et al. Differential genome-wide array-based methylation profiles in prognostic subsets of chronic lymphocytic leukemia. Blood. 2010;115:296–305.
6. Pei L, Choi JH, Liu J, Lee EJ, McCarthy B, Wilson JM, et al. Genome-wide DNA methylation analysis reveals novel epigenetic changes in chronic lymphocytic leukemia. Epigenetics. 2012;7:567–78.
7. Landau DA, Clement K, Ziller MJ, Boyle P, Fan J, Gu H, et al. Locally disordered methylation forms the basis of intratumor methylome variation in chronic lymphocytic leukemia. Cancer cell. 2014;26:813–25.
8. Cahill N, Kanduri M, Mansouri L, Sundstrom C, Rosenquist R, Ryan F, et al. 450K-array analysis of chronic lymphocytic leukemia cells reveals global DNA methylation to be relatively stable over time and similar in resting and proliferative compartments. Leukemia. 2013;27:150–8.
9. Kulis M, Heath S, Bibikova M, Queirós AC, Navarro A, Clot G, et al. Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. Nat Genet. 2012;44:1236–42.
10. Queiros AC, Villamor N, Clot G, Martinez-Trillos A, Kulis M, Navarro A, et al. A B-cell epigenetic signature defines three biologic subgroups of chronic lymphocytic leukemia with clinical impact. Leukemia. 2015;29:598–605.
11. Smith EN, Ghia EM, DeBoever CM, Rassenti LZ, Jepsen K, Yoon K, et al. Genetic and epigenetic profiling of CLL disease progression reveals limited somatic evolution and suggests a relationship to memory-cell development. Blood Cancer J. 2015;5:e303.
12. Papakonstantinou N, Ntoufa S, Chartomatsidou E, Kotta K, Agathangelidis A, Giassafaki L, et al. The histone methyltransferase EZH2 as a novel prosurvival factor in clinically aggressive chronic lymphocytic leukemia. Oncotarget. 2016;7:35946–59.
13. Rabello DdA, Lucena-Araujo AR, Alves-Silva JCR, Souza dEVBA, de Vasconcellos MC, de Oliveira FM, et al. Overexpression of EZH2 associates with a poor prognosis in chronic lymphocytic leukemia. Blood Cells Mol Dis. 2015;54:97–102.
14. Mansouri L, Wierzbinska JA, Plass C, Rosenquist R. Epigenetic deregulation in chronic lymphocytic leukemia: Clinical and biological impact. Semin Cancer Biol. 2018;51:1–11.
15. van Dijk AD, Hu CW, de Bont, ESJM, Qiu Y, Hoff FW, Yoo SY, et al. Histone modification patterns using RPPA-based profiling predict outcome in acute myeloid leukemia patients. Proteomics. 2018;18:e1700379.
16. van Dijk AD, Hoff FW, Qiu Y, de Bont ES, Bruggeman SWM, Horton TM, et al. Loss of H3K27 methylation identifies poor outcome in adult-onset acute myeloid leukemia. Blood. 2020;136:24.
17. Tibes R, Qiu Y, Lu Y, Hennessy B, Andreeff M, Mills GB, et al. Reverse phase protein array: validation of a novel proteomic technology and utility for analysis of primary leukemia specimens and hematopoietic stem cells. Mol cancer therapeutics. 2006;5:2512–21.
18. Kornblau SM, Coombes KR. Use of reverse phase protein microarrays to study protein expression in leukemia: technical and methodological lessons learned. Methods Mol Biol. 2011;785:141–55.
19. Kornblau SM, Tibes R, Qiu YH, Chen W, Kantarjian HM, Andreeff M, et al. Functional proteomic profiling of AML predicts response and survival. Blood. 2009;113:154–64.
20. Hu CW, Kornblau SM, Slater JH, Qutub AA. Progeny clustering: a method to Identify Biological Phenotypes. Sci Rep. 2015. https://doi.org/10.1038/srep12894.
21. Hartigan JA, Wong MAA. *K*-Means clustering algorithm. J R Stat Soc: Ser C (Appl Stat). 1979;28:100–8.

22. Therneau PM. Modeling survival data: extending the Cox model. Springer; 2000.
23. Swerdlow SH, WHO Classification of tumors of haematopoietic and lymphoid tissues.4th ed. IARC; 2008.
24. Scarfò L, Ferreri AJM, Ghia P. Chronic lymphocytic leukaemia. Crit Rev Oncol / Hematol. 2016;104:169–82.
25. Kipps TJ, Fraser G, Coutre SE, Brown JR, Barrientos JC, Barr PM, et al. Long-Term Studies Assessing Outcomes of Ibrutinib Therapy in Patients With Del(11q) Chronic Lymphocytic Leukemia. Clin Lymphoma Myeloma Leuk. 2019;19: 715–22.e6.
26. Burger JA, Barr PM, Robak T, Owen C, Ghia P, Tedeschi A, et al. Long-term efficacy and safety of first-line ibrutinib treatment for patients with CLL/SLL: 5 years of follow-up from the phase 3 RESONATE-2 study. Leukemia 2020;34:787–98.
27. van Dijk AD, Hoff FW, Qiu YH, Chandra J, Jabbour E, de Bont, ESJM, et al. Loss of H3K27 methylation identifies poor outcomes in adult-onset acute leukemia. Clin Epigenet. 2021. https://doi.org/10.1186/s13148-021-01011-x.
28. Li Y, Schulz VP, Deng C, Li G, Shen Y, Tusi BK, et al. Setd1a and NURF mediate chromatin dynamics and gene regulation during erythroid lineage commitment and differentiation. Nucleic Acids Res. 2016;44:7173–88.
29. Li BE, Gan T, Ernst P, Meyerson M, Rabbitts TH. Distinct pathways regulated by menin and by MLL1 in hematopoietic stem cells and developing B cells. Blood. 2013;122:2039–46.
30. Tusi BK, Deng C, Salz T, Zeumer L, Li Y, So CWE, et al. Setd1a regulates progenitor B-cell-to-precursor B-cell development through histone H3 lysine 4 trimethylation and Ig heavy-chain rearrangement. FASEB J. 2015;29:1505–15.
31. Yang W, Ernst P. Distinct functions of histone H3, lysine 4 methyltransferases in normal and malignant hematopoiesis. Curr Opin Hematol. 2017;24:322–8.
32. Su IH, Basavaraj A, Krutchinsky AN, Hobert O, Ullrich A, Chait BT, et al. Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. Nat Immunol. 2003;4:124–31.
33. Zhou M. Small-Molecule Modulators of Methyl-Lysine Binding for the CBX7 Chromodomain. Chem Biol. 2015;22:161–8.
34. Maharaj K, Powers JJ, Achille A, Deng S, Fonseca R, Pabon-Saldana M, et al. Silencing of HDAC6 as a therapeutic target in chronic lymphocytic leukemia. Blood Adv. 2018;2:3012–24.
35. Fiskus W, Buckley K, Rao R, Mandawat A, Yang Y, Joshi R, et al. Panobinostat treatment depletes EZH2 and DNMT1 levels and enhances decitabine mediated de-repression of JunB and loss of survival of human acute leukemia cells. Cancer Biol Ther. 2009;8:939–50.
36. Fiskus W, Pranpat M, Balasis M, Herger B, Rao R, Chinnaiyan A, et al. Histone deacetylase inhibitors deplete enhancer of zeste 2 and associated polycomb repressive complex 2 proteins in human acute leukemia cells. Mol cancer therapeutics. 2006;5:3096–104.
37. Lee DH, Kim GW, Kwon SH. The HDAC6-selective inhibitor is effective against non-Hodgkin lymphoma and synergizes with ibrutinib in follicular lymphoma. Mol Carcinog. 2019;58:944–56.
38. Cahill N, Rosenquist R. Uncovering the DNA methylome in chronic lymphocytic leukemia. Epigenetics. 2013;8:138–48.

AUTHOR CONTRIBUTIONS

ADvD, TLG, JWL, ESJMdB, and SMK designed and supervised the study. SMK, JAB, and WW collected data, ADvD, TLG, FWH, ET, KR, PPR, and YHQ performed research. ADvD and SMK analyzed data and wrote the paper. All authors read and approved the final manuscript.

CONFLICT OF INTEREST

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41375-021-01438-4.

Correspondence and requests for materials should be addressed to Anneke D. van Dijk.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.