


Multiparametric analysis of CD8⁺ T cell compartment phenotype in chronic lymphocytic leukemia reveals a signature associated with progression toward therapy

Pauline Gonnord ^a, Manon Costa^b, Arnaud Abreu^{c,d,e}, Michael Peres^f, Loïc Ysebaert^{g,h}, Sébastien Gadatⁱ, and Salvatore Valitutti^{a,e}

^aCentre de Recherches en Cancérologie de Toulouse (CRCT), INSERM U1037, «Equipe labellisée Ligue Nationale contre le cancer 2018», Université de Toulouse III-Paul Sabatier, Toulouse, France; ^bInstitut de Mathématiques de Toulouse, UMR 5219, Université de Toulouse, CNRS, UPS IMT, Toulouse, France; ^cInstitut Roche, Boulogne-Billancourt, France; ^dUniversité de Strasbourg, CNRS, ICube, Strasbourg, France; ^eDepartment of Pathology, Institut Universitaire du Cancer-Oncopole de Toulouse, Toulouse, France; ^fLaboratoire d'Immunologie, CHU de Toulouse, France and Centre de Recherches en Cancérologie de Toulouse (CRCT), INSERM UMR1037, Toulouse, France; ^gDépartement d'Hématologie, Institut Universitaire du Cancer-Oncopole de Toulouse, Toulouse, France; ^hCentre de Recherches en Cancérologie de Toulouse (CRCT), INSERM U1037, «Equipe Innovations thérapeutiques des lymphomes B», Toulouse, France; ⁱToulouse School of Economics, Université Toulouse 1 Capitole, UMR5604, Institut de Mathématiques, Université Paul sabatier, Toulouse, France

ABSTRACT

CD8⁺ T cells are frontline defenders against cancer and primary targets of current immunotherapies. In CLL, specific functional alterations have been described in circulating CD8⁺ T cells, yet a global view of the CD8⁺ T cell compartment phenotype and of its real impact on disease progression is presently elusive. We developed a multidimensional statistical analysis of CD8⁺ T cell phenotypic marker expression based on whole blood multi-color flow-cytometry. The analysis comprises both unsupervised statistics (hClust and PCA) and supervised classification methods (Random forest, Adaboost algorithm, Decision tree learning and logistic regression) and allows to cluster patients by comparing multiple phenotypic markers expressed by CD8⁺ T cells.

Our results reveal a global CD8⁺ T cell phenotypic signature in CLL patients that is significantly modified when compared to healthy donors. We also uncover a CD8⁺ T cell signature characteristic of patients evolving toward therapy within 6 months after phenotyping. The unbiased, not predetermined and multimodal approach highlights a prominent role of the memory compartment in the prognostic signature. The analysis also reveals that imbalance of the central/effector memory compartment in CD8⁺ T cells can occur irrespectively of the elapsed time after diagnosis.

Taken together our results indicate that, in CLL patients, CD8⁺ T cell phenotype is imprinted by disease clinical progression and reveal that CD8⁺ T cell memory compartment alteration is not only a hallmark of CLL disease but also a signature of disease evolution toward the need for therapy.

ARTICLE HISTORY

Received 5 November 2018
Revised 10 December 2018
Accepted 8 January 2019

KEYWORDS

CD8⁺ T cells;
multidimensional
phenotyping; chronic
lymphocytic leukemia;
phenotypic signature;
supervised learning

Introduction

Results obtained in mice and humans established the notion that CD8⁺ T cells, and in particular cytotoxic T lymphocytes (CTL), are key components of the antitumor immune-surveillance. Accordingly, an increased CD8⁺ T cell infiltrate correlates with a better prognosis in various cancers.¹ In line with these observations, therapeutic protocols designed to potentiate CTL responses against tumor cells are currently at the frontline of cancer clinical research.^{2,3} A better understanding of CD8⁺ T cells functional phenotype in cancer patients is becoming increasingly important.

According to the immuno-editing model, the selective pressure of the immune system promotes tumor progression by selecting tumor variants that are fit to survive in an immunocompetent host.⁴ We hypothesize that a global remodeling of the CD8⁺ T cell compartment functional phenotype (beyond T cell exhaustion) in a process mirroring immuno-

editing, could highlight the development of a new equilibrium at the whole organism scale occurring during disease progression. Thus, monitoring the CD8⁺ T cell compartment phenotype might reveal the sculpturing of this compartment by the tumor and might provide tools to classify patients according to their disease evolution and need for therapy.

Chronic lymphocytic leukemia (CLL), a common adult leukemia characterized by the clonal expansion of B lymphocytes in the peripheral blood, lymphoid organs and bone marrow represents an excellent model to test such an hypothesis.⁵ Indeed, in this indolent disease, in which patients can live for years without needing treatment, cellular partners such as CD8⁺ T cells and tumor B cells can interact within the three main tumor compartments (blood, bone marrow and lymph nodes) over prolonged time periods.⁶ Moreover, defined CD8⁺ T cell functional deficiencies have been described in CLL patients, including defective lytic

synapse formation with tumor B cells and limited cytotoxic function.^{7,8}

Although a clear consensus exists on the point that several functional alterations occur in CD8⁺ T cells in CLL patients,⁹ a global view of the CD8⁺ T cell phenotypes reflecting their potential functional status is presently elusive.

To investigate possible global CD8⁺ T cell phenotypic remodeling in CLL patients, we undertook an unbiased approach for multi-dimensional characterization of CD8⁺ T cell phenotypic signature. We centered our study at the patient level so that we could compare multiple marker expression in multiple patients at the same time. For this, we implemented approaches for statistical multi-dimensional analysis of multicolor flow cytometry data sets.

Our results show that CD8⁺ T cell phenotype is altered in CLL patients when compared to healthy donors and that major alterations are embedded within a limited number of functional markers. The analysis also reveals a CD8⁺ T cell phenotypic signature in CLL patients that reflects disease progression toward therapy and is mainly due to imbalance in the memory compartment. Interestingly, memory compartment alteration appears to be an intrinsic feature of aggressive disease rather than the result of chronic immune system activation in CLL patients.

Results

Analysis of individual CD8⁺ T cell phenotypic marker expression reveals the necessity of using dimensionality reduction techniques

We initially compared the expression of a panel of 29 phenotypic markers from a cohort of CLL patients ($n = 31$) and a cohort of healthy donors ($n = 23$) (see [Table 1](#)- clinical information with Binet stage and IGVH mutation and [Table 2](#)-marker description). We focused our study on whole blood to preserve tumor micro-environment and to ensure that CD8⁺ T cells keep imprinting of their recent interactions within tumor niches. Moreover, the observation that the expression of several markers can be altered by cell isolation/freezing procedures supports the validity our choice of a whole blood-based analysis ([Supplementary Figure 1A](#) and ^{10,11}).

The multidimensional raw expression data of phenotypic markers for all individuals included in the study is presented in [Supplementary Figure 2](#) and is summarized in the heat map of [Supplementary Figure 3A](#). We first compared the mean expression levels from CLL patients and healthy donors for each marker. Wilcoxon tests showed that 58% of the markers (17 out of 29) exhibited a significantly different expression in CLL patients and healthy donors ([Supplementary Figure 3B](#)). This observation suggested that taking into account a combination of various markers could be instrumental to better characterize CD8⁺ T cells in CLL patients as compared to healthy donors. We also analyzed the correlation between markers two by two and constructed correlation plots for the two cohorts using pairwise Spearman correlation coefficients ([Supplementary Figure 3C](#)). This analysis showed that the nature and the intensity of marker expression correlation were different in CLL patients when compared to healthy donors.

Together, the above results provided a first indication that the CD8⁺ T cell compartment is molded by the disease. However, the high dimensionality of the data sets prompted us to use multi-dimensional analysis and dimension reduction techniques to have an integrated view of global CD8⁺ T cell remodeling.

Unsupervised multidimensional analysis of functionally diverse phenotypic markers allows CLL patient and healthy donor clustering

We focused our analysis at the patient population level. We thus considered each patient as one data point with coordinates in 29 dimensions. We initially compared CLL patients with healthy donors to define the appropriate method to discriminate individuals, since difference between healthy individuals and CLL patients is an obvious read-out.

First, we used hierarchical clustering algorithm (hClust) to define whether the considered markers allowed clustering of similar individuals at the multi-dimensional level. Based on the 29 marker expression on CD8⁺ T cells, hClust generated a dendrogram separating the individuals in two main clusters ([Figure 1\(a\)](#)): one cluster comprised mainly CLL patients (seven errors), and the other cluster contained mainly healthy donors (four errors). Thus, hClust separated the healthy donors from the CLL patients with 79.6% accuracy (see confusion matrix in [Supplementary Figure 4](#)). These observations confirmed that the markers were reliable to highlight major CD8⁺ T cell compartment differences in our data set and were powerful enough to cluster similar individuals by only analyzing CD8⁺ T cells phenotypes.

In parallel, we performed Principal Component Analysis (PCA) to highlight the markers that were driving the clustering of individuals included in the study. We considered only the first 2 dimensions created by PCA that were driving most of the variation in the data set (31.2%) ([Figure 1\(b\)](#) and [Supplementary Figure 4B](#)) and used them to plot the “hClust generated” clusters. We observed that the “CLL cluster” and the “Healthy cluster” were separated mostly according to dimension 1 of PCA. Interestingly, the markers correlating the most with this first dimension, and thus responsible for the difference between the individuals, are indicators of relevant biological functions of CD8⁺ T cells such as: migration and adhesion (CXCR4, CD11a, CCR7, CD58), lytic function (GzB, GzA, perforin), cell activation and differentiation (CD57, CD127, CD45RA, CD45RO, CD27) ([Figure 1\(c\)](#)). While adhesion molecule and lytic molecule expression correlated positively with dimension 1, chemokine receptor and activation/differentiation molecule expression negatively correlated with dimension 1 ([Figure 1\(b,c\)](#)). We also observed that, four markers (CCR7, CD27 CD45RA and CD45RO) that are commonly used to define naive, central memory (CM), effector memory (EM) and effector (EMRA) CD8⁺ T cells were present within the most correlating markers. We thus combined these four markers in a multi-step gating strategy ([Table 2](#)) to evaluate the impact that the various CD8⁺ T cell subsets (naive, effector, memory, etc.) have on the discrimination of CLL patients from healthy donors since alterations in

Table 1. CLL patients and healthy donors included the study. FCR (fludarabine, cyclophosphamide, and rituximab), Rbenda (rituximab and bendamustine), ND (Not determined).

CLL patients											Healthy donors						
patient ID	Sex	Age	Lymphocyte count(10e9/L)	IGHV mutational status	Binet Stage	Del 13q	Del 11q(A/TM)	Trisomy 12	Del 17p (TP53)	Evolution toward Treatment 6 months after phenotyping	CMV sero-status	Time to diagnosis (years)	patient ID	Sex	Age	CMV sero-status	
																	% del 13q
1	CLL16	M	64.5	63.7	C	YES	YES	NO	NO	FCR	+	0.0	1	H26	M	56	-
2	CLL17	F	67.8	52.2	A	YES	NO	YES	NO	None	ND	ND	2	H36	M	65	-
3	CLL19	M	66.4	11.8	B	NO	NO	YES	NO	Rbenda	+	0.0	3	H37	M	37	ND
4	CLL20	M	65.6	108.6	A	YES	NO	NO	NO	None	ND	12.0	4	H38	M	57	-
5	CLL21	F	65.5	131.7	A	NO	NO	NO	YES	None	ND	17.0	5	H39	M	48	ND
6	CLL22	M	84.2	51.2	B	YES	YES	NO	NO	None	ND	3.9	6	H40	M	45	ND
7	CLL23	M	74.5	196.4	C	NO	YES	NO	NO	None	ND	2.0	7	H44	M	35	ND
8	CLL29	F	71.3	7.2	A	ND	ND	ND	ND	None	-	0.0	8	H46	M	50	-
9	CLL32	M	55.0	12.8	A	ND	ND	ND	ND	None	ND	ND	9	H47	M	37	+
10	CLL42	M	75.4	128.8	A	YES	NO	NO	NO	None	ND	5.5	10	H48	M	37	ND
11	CLL43	M	75.8	25	A	NO	NO	NO	NO	None	ND	6.5	11	H49	M	50	-
12	CLL53	F	50.7	140	A	NO	YES	NO	NO	None	ND	ND	12	H50	F	47	ND
13	CLL59	M	39.0	182.5	B	NO	NO	NO	YES	None	-	1.0	13	H51	M	45	ND
14	CLL61	M	76.9	113.4	B	NO	NO	NO	NO	Rbenda	+	8.9	14	H52	M	49	ND
15	CLL62	F	72.1	90.4	A	YES	NO	NO	NO	None	-	ND	15	H53	M	42	ND
16	CLL63	M	67.3	25	C	NO	NO	NO	NO	Rbenda	ND	0.1	16	H54	M	60	ND
17	CLL64	F	68.7	67	B	NO	NO	NO	NO	None	ND	6.3	17	H55	M	64	ND
18	CLL65	M	67.6	37.6	A	NO	NO	NO	NO	None	-	2.5	18	H56	M	53	ND
19	CLL66	F	77.2	80.9	A	NO	NO	NO	YES	None	+	4.5	19	H57	M	55	ND
20	CLL67	F	71.2	131	A	NO	NO	NO	NO	None	-	5.9	20	H58	F	60	ND
21	CLL68	F	77.6	114	A	NO	NO	NO	NO	None	ND	2.2	21	H59	F	64	ND
22	CLL72	M	51.6	27.6	A	YES	NO	NO	NO	None	ND	2.5	22	H60	F	66	ND
23	CLL73	M	61.2	63	B	NO	YES	NO	NO	None	ND	10.3	23	H61	F	66	ND
24	CLL74	M	73.8	38.7	A	NO	NO	YES	NO	None	ND	3.0					
25	CLL75	F	69.2	111.5	A	YES	NO	NO	NO	None	ND	6.0					
26	CLL76	M	84.2	44.5	B	NO	NO	YES	NO	Ibrutinib	ND	2.7					
27	CLL79	M	62.9	20.8	B	NO	NO	NO	NO	None	ND	3.0					
28	CLL80	F	65.1	72.9	B	NO	NO	NO	NO	None	ND	4.7					
29	CLL82	M	72.9	102.8	C	NO	NO	NO	NO	FCR	ND	ND					
30	CLL83	F	67.0	28.4	C	NO	NO	NO	YES	Ibrutinib	ND	3.0					
31	CLL84	F	66.7	63.6	B	NO	YES	NO	NO	None	+	8.2					
		M/F	Mean			% del 13q	% del 11q	% trisomy 12	% del 17p	% evolution					M/F	Mean	
		1.38	68			25.8 (8/31)	19.3 (6/31)	12.9 (4/31)	12.9 (4/31)	22.5 (7/31)					3.60	52	

Table 2. List of markers and parameters extracted from flow cytometry data and used in the study.

	Marker	Parameters extracted	Gating parameters
1	B7-H3	% of CD8	
2	BTLA	% of CD8	
3	CCR4	% of CD8	
4	CCR5	% of CD8	
5	CCR7	% of CD8	
6	CD127	% of CD8	
7	CD137	% of CD8	
8	CD25	% of CD8	
9	CD27	% of CD8	
10	CD38	% of CD8	
11	CD45RA	% of CD8	
12	CD45RO	% of CD8	
13	CD5	% of CD8	
14	CD54	% of CD8	
15	CD57	% of CD8	
16	CD58	% of CD8	
17	CD69	% of CD8	
18	CTLA-4	% of CD8	
19	CXCR3	% of CD8	
20	CXCR4	% of CD8	
21	CXCR5	% of CD8	
22	Gal-3	% of CD8	
23	GzA	% of CD8	
24	GzB	% of CD8	
25	HLA-II	% of CD8	
26	LAG-3	% of CD8	
27	PD1	% of CD8	
28	PERFORIN	% of CD8	
29	CD11A	% of CD8	CD11A ^{high}
30	Naive	% of CD8	CD45 RA ⁺ , CD45RO ⁻ , CCR7 ⁺ , CD27 ⁺
31	EMRA	% of CD8	CD45 RA ⁺ , CD45RO ⁻ , CCR7 ⁻ , CD27 ⁻
32	EM	% of CD8	CD45 RA ⁻ , CD45RO ⁺ , CCR7 ⁺ , CD27 ⁻
33	CM	% of CD8	CD45 RA ⁻ , CD45RO ⁺ , CCR7 ⁺ , CD27 ⁺

CD8⁺ T cell differentiation subsets have been described in CLL.¹² When the differentiation subsets were introduced into the clustering analysis (instead of the markers individually) the accuracy increased to 81.5%.

To test whether the observed imprinting of CD8⁺ T cells from CLL patients was correlated with functional modifications, we analyzed the *ex vivo* effector capabilities of CD8⁺ T cells. We observed that the average amount of IFN γ produced per cell was lower in CLL patients compared to healthy donors even though the percentage of cells producing IFN γ was more important in CLL patients (Supplementary Figure 5A). Moreover, the cytotoxicity of CD8⁺ T cells toward conventional targets or autologous tumor B cells was reduced (Supplementary Figure 5B) despite high levels of lytic molecules expression (Supplementary Figure 2). In agreement with previously reported data,^{7,8} these observations suggest that although exhibiting an activated phenotype CLL CD8⁺ T cells are functionally deficient.

Taken together these results show that non-supervised analysis of multiple and biologically non-related CD8⁺ T cell markers can efficiently discriminate CLL patients from healthy donors. These results imply that the CD8⁺ T cell compartment of CLL patients is molded by the disease and suggest that the CD8⁺ T cell imprinting is affecting markers of biological activation.

Clustering of healthy donors and CLL patients is not explained by age differences and CMV infection

Since some discriminating markers between CLL patients and healthy donors are markers of activation and differentiation,

known to be influenced by age,¹³ and since CLL is a disease associated with aging, we investigated whether the “*patient/healthy donor clusters*” we observed were due to age differences. For that, we performed hClust/PCA analysis by considering samples of individuals from two smaller cohorts (CLL and healthy) with a narrow age-matching (50–67 y for CLL patients and 50–66 y for healthy donors). We observed that the accuracy of clustering was comparable to that obtained with the previous analysis (82.1%) and that markers correlating the most with dimension 1 (responsible for CLL patient/healthy donor discrimination) were essentially not changed (Figure 2(a–c)).

CMV infection has been associated with CLL, and CMV specific expansion of CD8⁺ T cells in CLL patients has been reported to be more pronounced than in age-matched healthy individuals.¹⁴ We thus wondered whether CMV imprinting of CD8⁺ T cells could explain CLL patient/healthy donor clustering. Since we did not have access to the CMV sero-status information for all the individuals, we investigated reduced groups of individuals (CLL and healthy, for which we had access to CMV sero-status information) by hClust/PCA analysis. hClust clustered patients with no error (accuracy = 100%) even though several CLL patients were CMV⁻ and one healthy donor was CMV⁺ (Figure 2(d–f)).

These results indicate that, even though we cannot exclude some influences of age and CMV infection on CD8⁺ T cell remodeling, disease imprinting on CD8⁺ T cells appears to be the main driver of CLL/healthy donor clustering.

Alteration of CD8⁺ T cell memory compartment correlates with need for therapy as revealed by supervised statistical methods

Since CLL is an indolent disease and some patients can live for years without therapy, predicting the potential need of treatment before uncontrolled tumor progression is of major interest. Since we described a CLL CD8⁺ T cell phenotypic imprinting that is strong enough to cluster CLL patients and healthy donors, we asked whether this signature could also classify patients on the basis of their need for therapy. We selected progression toward therapy as a readout rather than established prognostic markers since the decision to treat is a turning point of the disease that could be associated with observable phenotypic changes among CD8⁺ T cells (See Table 1 for clinical information). We used a similar strategy of hClust/PCA analysis to generate clusters of patients. The optimal number of clusters proposed by hClust was two and we observed that the markers that were correlating with dimension 1 of PCA were for a large part similar to the ones responsible of the CLL/Healthy discrimination. However, we could not observe a significant clustering of the patients according to need for therapy (Figure 3(a–c)). Significant clustering of patients according to Binet stage or IGVH mutational status was also not observed (Supplementary Figure 8A and Table 1). This observation could have two major explanations: (1) CD8⁺ T cell compartment is not shaped by clinical progression toward need for therapy; (2) unsupervised statistical methods might not

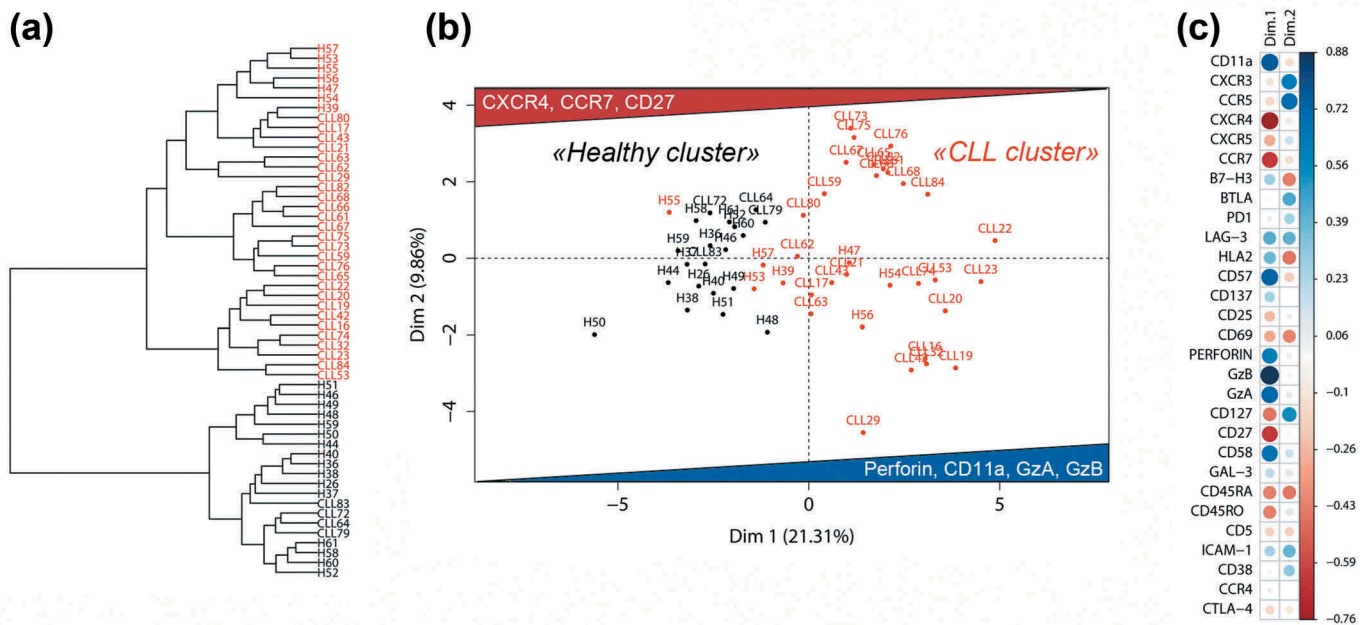


Figure 1. Clustering of CLL patients and healthy donors using unsupervised multidimensional analysis of functionally diverse phenotypic markers. (a) Dendrogram based on 29 marker expression on CD8⁺ T cells of CLL patients and healthy donor cohorts, generated by hierarchical clustering on Euclidian distances between the marker expression values. One group containing mostly CLL patients is colored in red, and the other group containing mostly healthy donors is colored in black. (b) Two-dimensional representation of PCA analysis. The whole data set is reduced using PCA analysis and the patients are plotted in the first two dimensions generated by PCA using the same color code as in Figure 1A. The blue triangle indicates the orientation of the expression of markers positively correlating with dimension 1; the red triangle indicates the orientation of the expression of markers negatively correlating with dimension 1. Examples of markers correlating positively and negatively according to correlation plot of Figure 1C are indicated in the triangles. (c) Correlation coefficients of each marker with the PCA dimension 1 and 2. Correlation coefficients are described by dot color for the nature of the correlation (blue for positive correlation, red for negative correlation, see scale beside the panel) and dot size for amplitude of correlation.

be able to unveil subtle phenotypic differences. To address this point, we used supervised learning algorithms to investigate whether a significant phenotypic signature was associated with need for therapy.

We first applied Random Forest (RF) algorithm which learns from profiles of marker expression how to create decision trees that select a combination of relevant markers allowing separation of individuals according to a defined criterion. We generated an RF model to uncover the phenotypic markers that would allow discrimination of CLL patients treated within the 6 months period after phenotyping. We conducted a repeated cross-validation scheme of the RF algorithm from which we extracted an average accuracy of the prediction on the validation set and an average importance of the markers (Figure 3(d)). The average accuracy of need for treatment prediction (out of 1000 repetitions) was 73.37% demonstrating the existence of a specific phenotype of CD8⁺ T cells associated with patient need for therapy. The three first markers of the RF analysis (CM, EM, and CXCR4) were effective to distinguish the patients that evolve toward therapy from the one who do not (Figure 3(e) and Supplementary Figure 6A). Interestingly, the expression profile of the differentiation markers associated with evolution toward therapy (increased representation of EM cells and decreased representation of CM cells among CD8⁺ T cells Supplementary Figure 6A) was also observed, to a lesser extent, in patients with more advanced

Binet stage (B and C) and patients with unmutated IGVH genes (Supplementary Figure 8B).

To confirm the existence of this specific phenotypic signature of CD8⁺ T cells in patients evolving toward therapy, we tested two additional learning algorithms from which we could get a feature hierarchy after learning. With Adaboost algorithm (Supplementary Figure 6B), we found again differentiation status subsets (EM, CM) among the five most important markers to predict the need for therapy. Those two subsets were also the most important markers in the Decision tree learning (Supplementary Figure 6C).

In conclusion, our supervised analysis highlights a phenotypic signature correlating with evolution toward need for therapy.

CD8⁺ T cell compartment signature associated with need for therapy allows to score CLL patients on the basis of their CD8⁺ T cell compartment

To extend the validity and relevance of the uncovered phenotypic signature, we assessed whether the markers identified in the signature could be used to score patients on the basis of their CD8⁺ T cell compartment and to evaluate the “statistical chance” for a patient’s disease to evolve to a stage requiring treatment during the 6 months following phenotypic analysis.

We thus computed a logistic regression using the most relevant markers of the RF model predicting the need for treatment 6 months after phenotyping. The calculated score

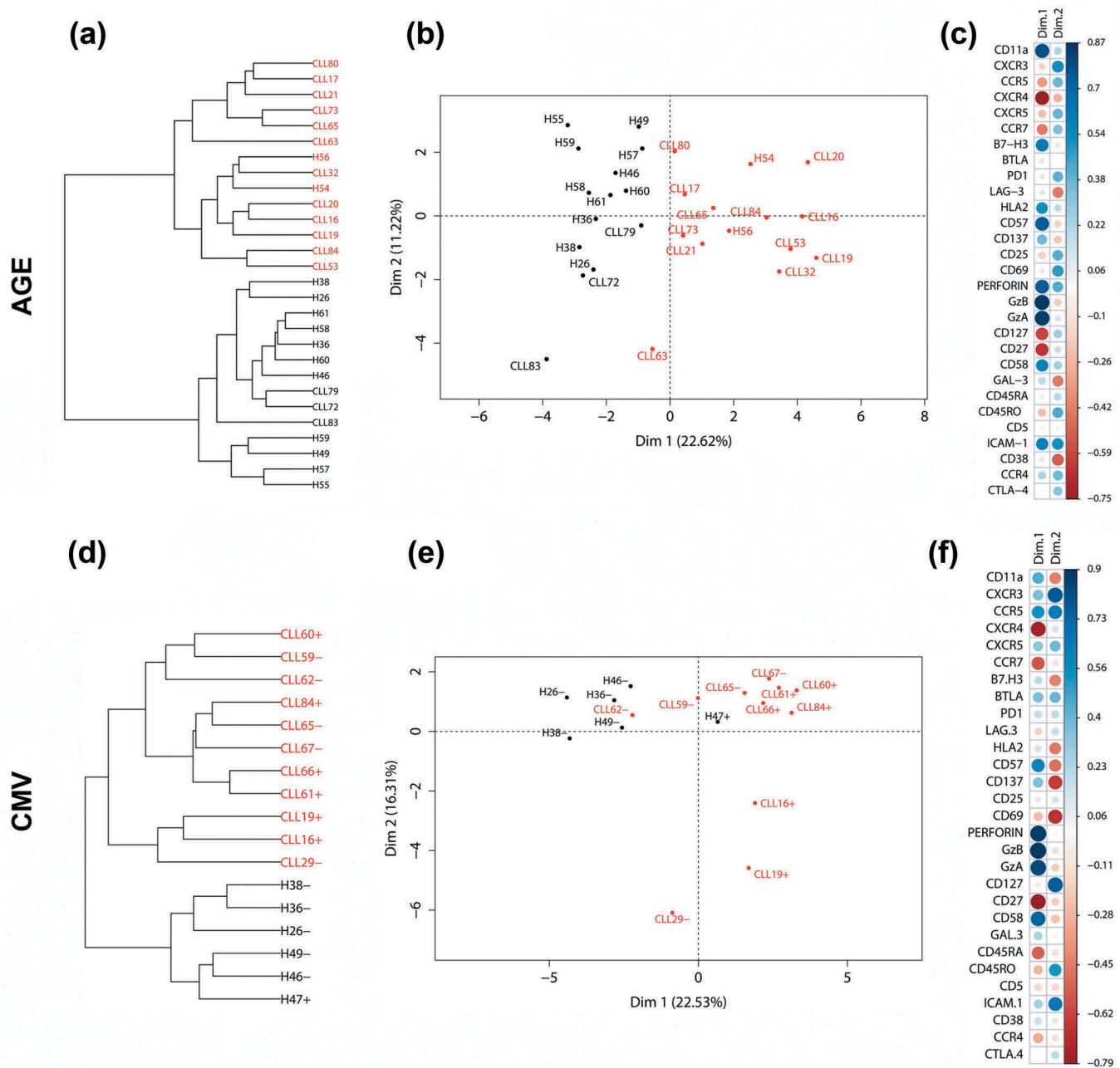


Figure 2. Clustering of healthy donors and CLL patients is not explained by age differences and CMV infection. PCA/hClust analyses based on 29 marker expression on CD8⁺ T cells of “age-matched” CLL patients and healthy donors (a-c) and selected CLL patients and healthy donor with known CMV sero-status (d-f). (a and d) Dendrograms generated by hierarchical clustering on Euclidian distances between the marker expression values on CD8⁺ T cells. One group containing mostly CLL patients is colored in red, and the other group containing mostly healthy donors is colored in black. B and E- Two-dimensional representation of PCA analysis. The whole data set is reduced using PCA analysis and the patients are plotted in the first two dimensions generated by PCA using the same color code as in Figure 2A. (c and f) Correlation coefficients of each marker with the PCA dimensions 1 and 2. Correlation coefficients are described by dot color for the nature of the correlation (blue for positive correlation, red for negative correlation, see scale beside the panels) and dot size for amplitude of correlation.

represents the probability of being in one of the two states (“treated” or “untreated”) according to the phenotypic marker expression values and the coefficients applied in the regression (Figure 4(a)). We used a cross-validation scheme where we split the cohort into three groups. We sequentially used two groups as learning cohorts to calculate the coefficients applied to each variable and one group as testing cohort to calculate the score on remaining patients and assess the validity of the model. Then, based on the learning data, we adjusted an optimal threshold with ROCR package¹⁵ to

optimize the number of “false positive” and “true negative”. The individuals whose score was above the cutoff were predicted as “treated” and “untreated” otherwise. An example of scores calculated for learning and testing patients is presented in Figure 4(b). We calculated the average accuracy and the F-measure of prediction by our model using increasing numbers of markers following the hierarchy of the RF model. We also screened different groups of patients to see whether groups’ composition was affecting the precision of accuracy and F-measure (Figure 4(c,d)). We observed that increasing

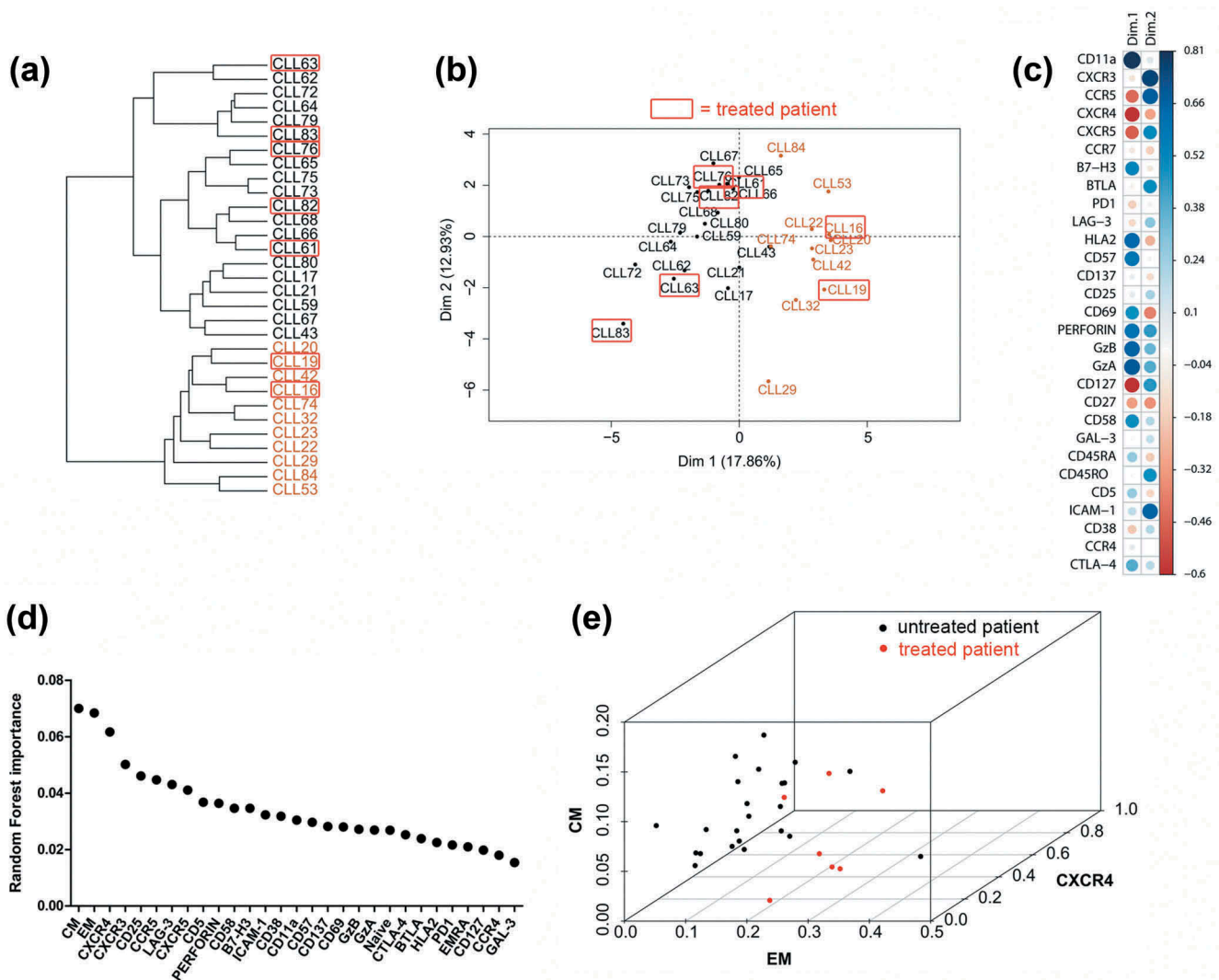


Figure 3. Supervised learning of phenotypic imprinting of CD8⁺ T cells associated with need for therapy confirms the importance of the memory compartment. (a) Dendrogram based on 29 marker expression on CD8⁺ T cells of CLL patients, generated by hierarchical clustering on Euclidian distances between the marker expression values. The two groups of patients proposed by hClust are colored in black and brown. (b) Two-dimensional representation of PCA analysis. The whole data set is reduced using PCA analysis and the patients are plotted in the first two dimensions generated by PCA using the same color code as in Figure 3A. Treated patients are indicated by red boxes. (c) Correlation coefficients of each marker with the PCA dimension 1 and 2. Correlation coefficients are described by dot color for the nature of the correlation (blue for positive correlation, red for negative correlation, see scale beside the panels) and dot size for amplitude of correlation. (d) Parameters correlating with “need for therapy” as ranked by Random Forest analysis. The parameters are ranked according to normalized Gini index of their importance (Random Forest importance). (e) 3-D representation of the patients (untreated: black dot, treated 6 months after phenotyping: red dot) according to CM, EM and CXCR4 expression values.

the number of markers taken into account in the logistic regression did not improve the accuracy or the F-measure suggesting that the first 2 markers (CM and EM) have a strong influence on the evolution toward therapy.

As a control, we calculated the score in healthy donors by applying the regression learned on CLL patients and observed that all donors were predicted “untreated” when using the first 2 markers (CM and EM) (Supplementary Figure 7A).

To test whether the observed signature of CD8⁺ T cells from CLL patients that will evolve toward therapy correlated with functional modifications, we also analyzed the *ex vivo* cytokine production capability of CD8⁺ T cells in the different groups of patients. We observed that the average percentage of cytokine-producing cells and the amount produced per cell (IFN γ , TNF α , IL-2 and MIP-1 β) was not statistically different

in CLL patients who evolve toward treatment versus the ones who do not (Supplementary Figure 7B).

In conclusion, these results identify a phenotypic signature of CD8⁺ T cells in CLL patients that evolve toward therapy that reflects tumor sculpturing of CD8⁺ T cells. They highlight a combination of surface markers (CM, EM, CXCR4) that can be used to score CLL patients probability of disease progression.

A frozen validation cohort confirms the existence of CD8⁺ T cell phenotypic imprinting

We next used a validation cohort of frozen PBMC from CLL patients (untreated at the time of sample collection) to have access to a larger cohort of patients with available clinical follow-up data (Table 3).

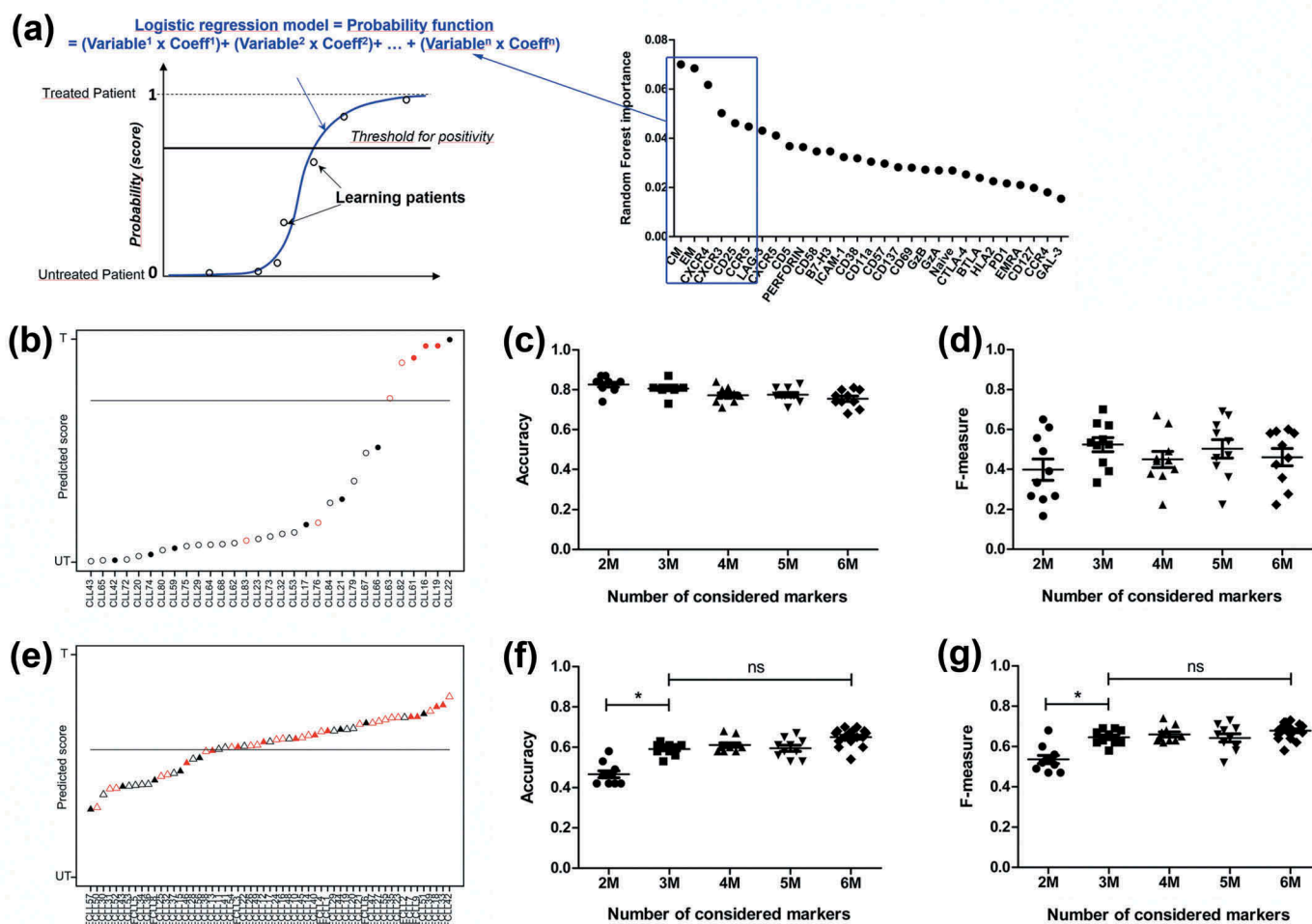


Figure 4. CD8⁺ T cell compartment signature associated with need for therapy allows to score CLL patients on the basis of their CD8⁺ T cell compartment. (a) Example of graphical representation of a typical logistic regression model as used in Figure 4B,E. Random Forest analysis of Figure 3(d) is reminded to indicate which markers will be taken into account to create the logistic regression. (b) Graphical example of calculated scores of patients using a logistic regression model constructed with two markers (CM, EM). The cohort was split into three groups to conduct a three-fold validation scheme (Only one fold is presented here – see R file for visualization of all repetitions). The patients that were used to learn the regression and calculate the coefficient are represented as open circle and the patients that were used to apply the regression and calculate scores are represented as close circle. Patients that evolved toward therapy within 6 months after phenotyping are plotted in red while all other patients that did not need therapy are plotted in black. The optimized threshold was calculated using the ROCR package. The accuracy of this particular example is 0.9 (90%) and F-measure is 0.86. (c) Accuracies of logistic regression models predictions. Different logistic regressions were generated using 2,3,4,5 or 6 markers according to RF analysis. For each model, the mean accuracy of the threefold validation scheme was calculated and plotted as one dot. The same process was repeated 10 times after changing the groups of patients (two-tailed Mann–Whitney test * $p < 0.05$, ns = non-significant, black line = mean, error bars = S.D.). (d) F-measure of logistic regression models predictions. Different logistic regressions were generated using 2,3,4,5 or 6 markers according to RF analysis. For each model, the mean F-measure of the threefold validation scheme was calculated and plotted as one dot. The same process was repeated 10 times after changing the groups of patients (two-tailed Mann–Whitney test * $p < 0.05$, ns = non-significant, black line = mean, error bars = S.D.). (e) Graphical example of calculated scores of patients from the validation cohort using a logistic regression model constructed with three markers (CM, EM and CXCR4). The cohort was split into three groups to conduct a threefold validation scheme (Only one fold is presented here – see R file for visualization of all repetitions). The patients that were used to learn the regression and calculate the coefficient are represented as open circle and the patients that were used to apply the regression and calculate scores are represented as close circle. Patients that evolved toward therapy within 6 months after phenotyping are plotted in red while all other patients that did not need therapy are plotted in black. The optimized threshold was calculated using the ROCR package. The accuracy of this particular example is 0.74 (74%) and F-measure is 0.78. (c) Accuracies of logistic regression models predictions of the validation cohort. Different logistic regressions were generated using 2,3,4,5 or 6 markers as in Figure 4C. For each model, the mean accuracy of the threefold validation scheme was calculated and plotted as one dot. The same process was repeated 10 times after changing the groups of patients (two-tailed Mann–Whitney test * $p < 0.05$, ns = non-significant, black line = mean, error bars = S.D.). (d) F-measure of logistic regression models predictions of the validation cohort. Different logistic regressions were generated using 2,3,4,5 or 6 markers as in Figure 4D. For each model, the mean F-measure of the threefold validation scheme was calculated and plotted as one dot. The same process was repeated 10 times after changing the groups of patients (two-tailed Mann–Whitney test * $p < 0.05$, ns = non-significant, black line = mean, error bars = S.D.).

As mentioned before, we observed that the expression of several phenotypic markers is altered by cell isolation/freezing procedures (Supplementary Figure 1A). Nevertheless, the use of frozen samples is convenient for several reasons: (a) frozen samples are much more easily available and shareable than fresh samples; (b) they can be processed in a more automated fashion; (c) they can be used in retrospective studies. We

thus investigated whether our CD8⁺ T cell signature and scoring system might be still valid on an additional frozen sample CLL cohort in spite of the possible alteration of some marker expression.

CD8⁺ T cell scores of these patients were computed using the same logistic regression method we described on the fresh cohort. An example of scores calculated for learning and

Table 3. Patients included in the frozen sample validation cohort.

REF NUMBER	AGE	TREATED AT 6 MONTHS
FCLL1	65.8	Yes
FCLL2	66.4	Yes
FCLL3	65.5	No
FCLL4	68.7	No
FCLL5	64.2	Yes
FCLL6	54.8	Yes
FCLL7	51.1	No
FCLL8	60.3	Yes
FCLL9	72.2	Yes
FCLL10	58.5	Yes
FCLL11	69.0	No
FCLL12	65.7	Yes
FCLL13	75.6	Yes
FCLL14	68.0	No
FCLL15	64.7	No
FCLL16	65.9	Yes
FCLL17	60.7	No
FCLL18	72.8	Yes
FCLL19	66.5	Yes
FCLL20	63.3	Yes
FCLL21	77.7	Yes
FCLL22	74.8	No
FCLL23	67.3	No
FCLL24	37.7	Yes
FCLL25	56.4	Yes
FCLL26	63.7	No
FCLL27	71.0	Yes
FCLL28	69.7	No
FCLL29	56.9	No
FCLL30	54.2	Yes
FCLL31	65.7	Yes
FCLL32	61.1	Yes
FCLL33	59.1	No
FCLL34	69.0	Yes
FCLL35	67.5	Yes
FCLL36	61.7	Yes
FCLL37	67.5	Yes
FCLL38	60.5	No
FCLL39	64.7	Yes
FCLL40	41.9	No
FCLL41	64.6	No
FCLL42	54.1	No
FCLL43	60.1	Yes
FCLL44	64.7	No
FCLL45	57.2	No
FCLL46	73.0	No
FCLL47	41.6	Yes
FCLL48	56.2	No
FCLL49	62.5	No
FCLL50	35.7	No
FCLL51	65.0	Yes
FCLL52	55.4	Yes
FCLL53	68.6	Yes
FCLL54	64.7	Yes
FCLL55	56.5	Yes
FCLL56	48.4	No
FCLL57	61.0	Yes

testing patients is presented in [Figure 4\(e\)](#). We again calculated the average accuracy and F-measure of prediction of our model for different cross-validation groups and using increasing numbers of markers following the hierarchy of the RF analysis ([Figure 4\(f–g\)](#)). We observed that increasing the number of markers taken into account in the logistic regression to three markers (CM, EM, CXCR4) (but not above) did improve the accuracy or the F-measure.

These observations confirm that a signature of three relevant markers (CM, EM, CXCR4) can be used to predict the need for therapy of CLL patients based on the phenotype of their CD8⁺ T cell compartment and that our approach might be extended to frozen cohorts.

CD8⁺ T cell memory compartment alteration can be detected early after disease diagnosis

Having observed an imprinting of the CD8⁺ T cell compartment associated with disease progression, we wondered whether observed phenotypic alterations were resulting from chronic immune system stimulation or were rather an intrinsic characteristic of an aggressive form of the disease.

To address this question, we investigated whether the observed memory compartment signatures might be related to the elapsed time since diagnosis. We thus defined four groups of patients based on the time elapsed between diagnosis and phenotyping (0–2 y, 2–5 y, 5–10 y and >10 y) and color-coded the patients accordingly. We plotted CLL patients according to their CM/EM expression values since these markers were strongly influencing CD8⁺ T cell imprinting using the time color-code. Our results show that we could not observe a natural classification of patients according to the time elapsed since CLL diagnosis ([Figure 5](#)). The elapsed time after diagnosis did not correlate with the CM/EM signatures of individual patients irrespectively of whether they were undergoing therapy or not.

These results provide evidences that CD8⁺ T cell phenotypic imprinting of the memory compartment in CLL disease are not linearly correlated with time exposure of CD8⁺ T cells to tumor CLL cells. They imply that the observed memory compartment alteration is an intrinsic signature of disease aggressiveness.

Discussion

In the present work, we investigated the remodeling of the CD8⁺ T cell compartment in CLL and the impact that CD8⁺ T cell phenotypic alterations might have on disease progression.

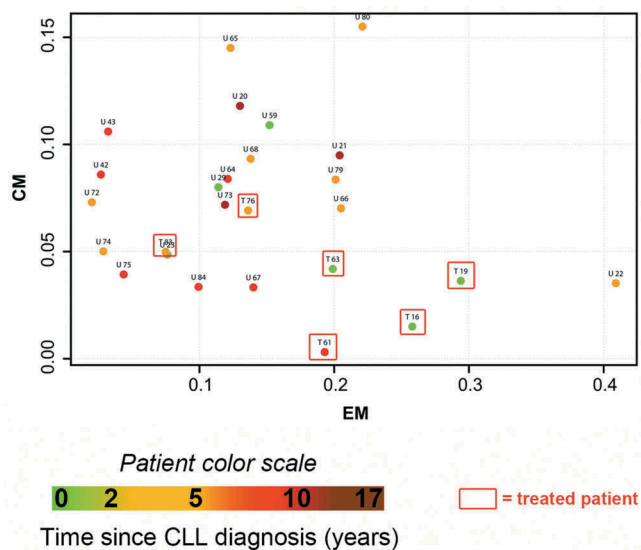


Figure 5. CD8⁺ T cell memory compartment alteration is uncoupled from elapsed time since diagnosis. Dot plot representation of patients according to their raw value expression of EM and CM markers. The patients are color-coded according to their time since CLL diagnosis (see legend below). Red rectangles indicate patients that were treated 6 months after phenotyping.

Statistical tools reveal a CD8⁺ T cell compartment-specific signature distinguishing CLL patients from healthy donors. In addition, supervised learning reveals a signature of “need for therapy” among patients that can be used to score disease progression toward therapy in individual patients. Moreover, alteration in CD8⁺ T cell memory compartment can occur irrespectively of the elapsed time after diagnosis.

A peculiar characteristic of our study is that statistical analysis was based on data obtained by phenotyping fresh samples that did not undergo any manipulation before staining to preserve both (i) the imprinting of the recent interactions of CD8⁺ T cells within tumor niches and (ii) molecules expression that can be affected by PBMC preparation and freezing^{10,11} and Supplementary Figure 1. Analyzing cell phenotypes of unprocessed cells by flow cytometry has the advantage of providing integrated pictures of the protein expression by each cell, taking into account genetic and epigenetic regulations.

The initial cohorts included in our study might appear of limited dimensions. Yet, our results are based on fresh blood samples collected over a 2-y period and that have been analyzed by deep multiplexed phenotyping with a multidimensionality comparable to that achieved by mass cytometry.¹⁶ Moreover, our experimental approach has the advantage over mass cytometry of being readily reproducible in clinical routine. In addition, although the use of frozen cell sample is less indicated for our multiplexed analysis, we were able to confirm the validity of our statistical procedure on a larger independent cohort of frozen CLL patients' peripheral blood lymphocyte samples.

The unsupervised analysis tools (hClust/PCA) allowed us to non-subjectively identify the CD8⁺ T cell markers whose expression is mainly altered in the tumor environment thanks to a patient cluster centered approach and provide a global view of CD8⁺ T cells phenotype. Importantly, our analysis revealed that clustering of healthy donors and CLL patients cannot only be explained by age differences or CMV infection, indicating that the disease itself (and not additional co-morbidity factors) is responsible of the observed global remodeling of the CD8⁺ T cell phenotype.

From a methodological point of view, the multidimensional statistical approach we employed represent a crucial tool to highlight the markers that are critical and exclude the non-informative ones in a non-subjective manner. This procedure might allow, in other studies based on multiplexed approaches aiming at characterizing CD8⁺ T cells signatures, to create sub-groups of markers that better allow classification of individuals¹⁷ such as multiplexed analysis of tissue samples analyzing activation signatures of tumor-infiltrating T cells in solid tumors.^{18,19}

Unsupervised hClust/PCA analysis has been reported as being adequate to highlight groups of patients with clinical relevance in other tumor models.^{20,21} Yet, they did not prove sensitive enough to reveal clear phenotype changes indicating evolution toward therapy in our disease model. Conversely, we find that supervised learning techniques such as RF, Adaboost and Decision Tree^{22–24} efficiently unveil the CD8⁺ T cell phenotypic profiles associated with disease progression. Moreover, we validated the reliability of the markers

characterizing the CD8⁺ T cell phenotypic signature of disease progression by a scoring system using a logistic regression model. The accuracy and F-measure of the model were validated not only in our fresh whole blood cohort, but also in an additional independent validation cohort despite analysis of frozen samples might have some limitations. This last observation might be explained by the fact that the difference in the expression of markers important for the signature (CM, EM, CXCR4) and the rest of the markers might have more weight than the difference of marker expression in fresh versus frozen samples.

The pattern revealed by supervised learning approaches and used for score calculation contain markers of CD8⁺ T cell activation state and migration potential (CM, EM, and CXCR4). Interestingly, these markers are compatible with numerous previous observations of CD8⁺ T cell phenotypic dysregulation in CLL patients, thus validating our automated approach.^{6,12,25–29} This signature also contains chemokine receptor CXCR4 that we find highly expressed in patients who will be treated. CXCR4 plays a central role in CLL³⁰ since it regulates the localization of B-CLL cells in the different tumor compartments and is coupled to BTK, a tyrosine kinase target of Ibrutinib.^{31,32} CXCR4 expression is higher in peripheral blood resting tumor CLL cells than in proliferative “recent” emigrants from LN or bone marrow.^{33,34} It is tempting to speculate that similarly to tumor CLL cells, CD8⁺ T cells that have high CXCR4 expression are confined in the blood and thus cannot interact with tumor CLL cells in other tumor niches. This hypothesis would explain why a higher expression of CXCR4 on CD8⁺ T cells is predictive of evolution toward therapy. Intriguingly, our results show that the signature of evolution toward therapy does not contain immune checkpoint markers. In CLL, their expression by CD8⁺ T cells is rather controversial. Some authors found PD1 upregulation on CD8⁺ T cell surface in CLL patients^{12,35–37} while others reported no or barely significant upregulation of PD1.^{29,38,39} The discrepancy among different reports could be due to the age of patients at phenotyping⁴⁰ or to disease stage.^{26,29} It should be noted that all the patients we considered are untreated at the time of phenotyping, while this is not always the case in other studies.

We found that alteration of the CD8⁺ T cell memory compartment is a prominent component of the “need for therapy” signature in CLL patients. The fact the EM/CM markers are part of the CD8⁺ T cell signature is certainly not novel,^{12,28,39,41,42} but the unbiased and multimodal approach by which EM and CM markers emerged from deconvolution of a large number of surface markers as the combined parameters defining patients that will evolve toward therapy validates both (i) our novel way “to weight” independent markers and (ii) the biological relevance of these markers to stratify CLL patients.

We observed a general trend of decreased representation of CD8⁺ T cells that have an unexperienced phenotype and an increased representation of CD8⁺ T cells that have an “effector type phenotype” in CLL patients versus healthy donors. This tendency is even more prominent in the memory compartment and even characterize the patients that will evolve toward therapy. Interestingly, our functional experiments

revealed that the *ex vivo* effector potential of CD8⁺ T cells from the CLL patients that will evolve toward therapy does not correlate with the increase of “effector type phenotype” suggesting that these cells are functionally deficient^{7,8} with no scaling in the deficiency.

It is tempting to speculate that our results might have implication for chimeric antigen receptor (CAR) T cell therapy in CLL. Our observation that having less differentiated memory CD8⁺ T cells (CM) is associated with lower chance to evolve toward therapy, suggests that the CD8⁺ CM T cell subset might be preferred for adoptive cell therapy or CART therapy, in line with previously reported data.^{43,44}

Moreover, our patient classification approach, rarely seen in immunophenotyping studies, allows investigating both the relation of memory markers with disease evolution toward therapy over time and the impact that the duration of the disease might have on memory compartment imbalance. By this way, we can define the evolution of phenotypic signatures despite the absence of same patient samples over time. In other words, we show that, at the patient population level, alteration in CD8⁺ T cell memory compartment can occur irrespectively of the elapsed time after diagnosis. This observation has two implications. First, it suggests that this phenotypic change is a component of unfavorable disease evolution rather than being the result of chronic stimulation of the immune system. Second, it suggests that periodic examination of patient circulating CD8⁺ T cell memory compartment might be used as a tool to uncover immune remodeling going on at the organism scale that could alert about disease evolution before clinical symptoms apparition.

All in all our results reveal a tumor-related imprinting of the CD8⁺ compartment in CLL patients and prompt to re-think immuno-editing as a bidirectional phenomenon in which immune system and tumor cells progressively sculpt each other. The possibility to read subtle changes engrafted into the CD8⁺ T cells phenotype by tumor clinical progression might pave the road to new immunomonitoring approaches aiming at scoring disease progression by assessing the new equilibrium established between tumor and immune system at whole organism scale.

Methods

Patients

All patients were referred for CLL (before any therapy) according to IWCLL criteria, between 2015 and 2017, in the Hematology Department of the Institut Universitaire du Cancer de Toulouse-Oncopole. Peripheral blood samples from untreated CLL patients (n = 31) were collected and processed following standard ethical procedures (according to the Declaration of Helsinki), after obtained written informed consent and referenced at the HIMIP laboratory (Collection des hémopathies de l'INSERM Midi-Pyrénées). According to the French law, HIMIP has been declared to the Ministry of Higher Education and Research (DC 2008-307 collection 1) and obtained a transfer agreement (AC 2008-29) after approbation by an ethical committee (Comité de Protection des Personnes Sud-Ouest et Outremer II). Clinical and biological annotations of the samples have been declared to

the CNIL (Comité National Informatique et Libertés, i.e. Data processing and Liberties National Committee). All patient clinical data are described in Table 1. Healthy specimens (n = 23) used for control conditions were obtained from fresh blood samples (Etablissement français du sang Midi-Pyrénées, Purpan University Hospital, Toulouse, France).

Flow cytometry staining protocol

We designed antibody panels for flow cytometry analysis to monitor 29 parameters describing the main biological functions of CD8⁺ T cells, such as cytotoxicity, migration, adhesion, activation, differentiation and expression of checkpoint molecules (Table 2 – marker description). For every patient, the blood sample was split into 8 tubes containing marker-specific antibodies pooled by 5 combined to gating antibodies (CD3, CD4, CD8 and CD19 antibodies). A control sample corresponding to gating antibodies mixed with isotype control antibodies was also recorded for each patient. We extracted the percentage of positive cells above the isotype control in the CD8⁺ T cell population (defined as CD3⁺ CD19⁻ CD8⁺ cells) as described in Supplementary Figure 1B for all markers of interest. We decided to focus on percentage of positive cells after comparing percentage of positive cells and mean fluorescence intensity in a dedicated analysis (see Supplementary methods and Supplementary Figure 1C).

Whole blood samples were directly mixed with fluorochrome-coupled antibodies (see complete list in Table 4) for at least 1 h at 4°C. Red blood cells were then lysed and cell samples were washed twice with FACS buffer (PBS, 1% Fetal

Table 4. List of the antibody specificities, clones, fluorochromes and suppliers used in the study.

Marker	Clone	Fluorochrome	Supplier
B7-H3 (CD276)	DCN.70	PE	Biologend
BTLA (CD272)	J168-540	BV421	BD Biosciences
CCR4	1G1	PECY7	BD Biosciences
CCR5	2D7/CCR5	BV421	BD Biosciences
CCR7	150503	BV421	BD Biosciences
CD11A	HI111	PE	BD Biosciences
CD127	HIL-7R-M21	V450	BD Biosciences
CD137	4B4-1	BV421	BD Biosciences
CD19	HIB19	PECF594	BD Biosciences
CD25	M-A251	PECY5	BD Biosciences
CD27	L128	APC	BD Biosciences
CD3	UCHT1	V500	BD Biosciences
CD38	HIT2	FITC	BD Biosciences
CD4	SK3	A700	Biologend
CD45RA	HI100	PECY7	Biologend
CD45RO	UCHL1	PE	BD Biosciences
CD5	UCHT2	PECY7	Biologend
CD54	HA58	PE	BD Biosciences
CD57	NK-1	FITC	BD Biosciences
CD58	1C3	FITC	BD Biosciences
CD69	FN50	PE	BD Biosciences
CD8	RPA-T8	BV786	BD Biosciences
CTLA-4 (CD152)	L3D10	PECY7	Biologend
CXCR3	1C6/CXCR3	AF488	BD Biosciences
CXCR4	12G5	PECY5	BD Biosciences
CXCR5	51505	PE	R&D systems
Gal-3	M3/38	AF647	Biologend
GzA	CD09	PB	Biologend
GzB	GB11	AF700	BD Biosciences
HLA-II	Tu39	FITC	BD Biosciences
LAG-3	REA351	APC	Miltenyi
LAMP1 (CD107a)	H4A3	PECY7	BD Biosciences
PD1	EH12.1	PECY7	BD Biosciences
PERFORIN	dG9	AF647	Biologend

calf serum, 1% Human serum, 0.01% Na Azide) before fixation in 2% Paraformaldehyde. Samples were then washed twice before permeabilization in FACS buffer with 0.1% Saponin. Cell suspensions were then stained with antibodies directed against intracellular proteins for 30 min at room temperature.

Data acquisition

Fluorescence distributions were acquired by flow cytometry on a BD FORTESSE cytometer. Fluorescence compensations, gating and selection of cells of interest (CD19⁻, CD3⁺, CD4⁻, CD8⁺, alive) were performed using FlowJo software (Tree Star, Inc, Ashland, Ore) and fluorescence data files corresponding to CD8⁺ T cells only were exported as csv files.

Statistics, data and code availability

Detailed statistical methods used throughout the study are available as supplementary information. We used R software for most statistical analyses⁴⁵ and python software for supervised learning.⁴⁶ The data sets generated and analyzed in this study, together with code generated in R and python software, are available as supplementary material.

Acknowledgments

We thank Sylvain Cussat-Blanc for discussion and critical reading of the manuscript. Eric Espinosa and Vera Pancaldi for discussion. We thank the CPTP-INSERM UMR 1043 flow cytometry core facility. We thank Sebastien Déjean for help with statistical analysis.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Funding

This work was supported by the Institut National du Cancer under contracts INCa PBLIO11-130 and INCa/DGOS 2012-054; Region Occitanie under contracts RCLÉ R14007BB, No 12052802, and RBIO R15070BB, No 14054342; the Laboratoire d'Excellence Toulouse Cancer (TOUCAN) under contract ANR11-LABX; Fondation Toulouse Cancer Santé under contract 2014CS044 and Ligue Nationale contre le Cancer (Equipe labellisée 2018). P.G. was supported by INCa, Region Occitanie and Fondation Toulouse Cancer Santé. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

ORCID

Pauline Gonnord  <http://orcid.org/0000-0001-8712-605X>

References

- Fridman WH, Pages F, Sautes-Fridman C, Galon J. The immune contexture in human tumours: impact on clinical outcome. *Nat Rev Cancer*. 2012;12:298–306. doi:10.1038/nrc3245.
- Kalos M, June CH. Adoptive T cell transfer for cancer immunotherapy in the era of synthetic biology. *Immunity*. 2013;39:49–60. doi:10.1016/j.immuni.2013.07.002.
- Chen DS, Mellman I. Oncology meets immunology: the cancer-immunity cycle. *Immunity*. 2013;39:1–10. doi:10.1016/j.immuni.2013.07.012.
- Dunn GP, Bruce AT, Ikeda H, Old LJ, Schreiber RD. Cancer immunoeediting: from immunosurveillance to tumor escape. *Nat Immunol*. 2002;3:991–998. doi:10.1038/ni1102-991.
- Scarfo L, Ferreri AJ, Ghia P. Chronic lymphocytic leukaemia. *Crit Rev Oncol Hematol*. 2016;104:169–182. doi:10.1016/j.critrevonc.2016.06.003.
- Mellstedt H, Choudhury A. T and B cells in B-chronic lymphocytic leukaemia: Faust, Mephistopheles and the pact with the Devil. *Cancer Immunol Immunother*. 2006;55:210–220. doi:10.1007/s00262-005-0675-4.
- Ramsay AG, Johnson AJ, Lee AM, Gorgun G, Le Dieu R, Blum W, Byrd JC, Gribben JG. Chronic lymphocytic leukemia T cells show impaired immunological synapse formation that can be reversed with an immunomodulating drug. *J Clin Invest*. 2008;118:2427–2437. doi:10.1172/JCI35017.
- Kabanova A, Sanseviero F, Candi V, Gamberucci A, Gozzetti A, Campoccia G, Bocchia M, Baldari CT. Human cytotoxic T lymphocytes form dysfunctional immune synapses with B cells characterized by non-polarized lytic granule release. *Cell Rep*. 2016;15:9–18. doi:10.1016/j.celrep.2016.02.084.
- Riches JC, Ramsay AG, Gribben JG. T-cell function in chronic lymphocytic leukaemia. *Semin Cancer Biol*. 2010;20:431–438. doi:10.1016/j.semcancer.2010.09.006.
- Romeu MA, Mestre M, Gonzalez L, Valls A, Verdager J, Corominas M, Bas J, Massip E, Buendia E. Lymphocyte immunophenotyping by flow cytometry in normal adults. Comparison of fresh whole blood lysis technique, ficoll-paque separation and cryopreservation. *J Immunol Methods*. 1992;154:7–10.
- Ng AA, Lee BT, Teo TS, Poidinger M, Connolly JE. Optimal cellular preservation for high dimensional flow cytometric analysis of multicentre trials. *J Immunol Methods*. 2012;385:79–89. doi:10.1016/j.jim.2012.08.010.
- Nunes C, Wong R, Mason M, Fegan C, Man S, Pepper C. Expansion of a CD8(+)/PD-1(+) replicative senescence phenotype in early stage CLL patients is associated with inverted CD4:CD8 ratios and disease progression. *Clin Cancer Res*. 2012;18:678–687. doi:10.1158/1078-0432.CCR-11-2630.
- Nikolich-Zugich J. The twilight of immunity: emerging concepts in aging of the immune system. *Nat Immunol*. 2018;19:10–19. doi:10.1038/s41590-017-0006-x.
- Mackus WJ, Frakking FN, Grummels A, Gamadia LE, De Bree GJ, Hamann D, Van Lier RAW, Van Oers MHJ. Expansion of CMV-specific CD8+CD45RA+CD27- T cells in B-cell chronic lymphocytic leukemia. *Blood*. 2003;102:1057–1063. doi:10.1182/blood-2003-01-0182.
- Sing T, Sander O, Beerwinkler N, Lengauer T. ROCRC: visualizing classifier performance in R. *Bioinformatics*. 2005;21:3940–3941. doi:10.1093/bioinformatics/bti623.
- Newell EW, Sigal N, Bendall SC, Nolan GP, Davis MM. Cytometry by time-of-flight shows combinatorial cytokine expression and virus-specific cell niches within a continuum of CD8+ T cell phenotypes. *Immunity*. 2012;36:142–152. doi:10.1016/j.immuni.2012.01.002.
- Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. *Nature*. 2017;541:321–330. doi:10.1038/nature21349.
- Gorris MAJ, Halilovic A, Rabold K, van Duffelen A, Wickramasinghe IN, Verweij D, Wortel IMN, Textor JC, de Vries IJM, Figdor CG. Eight-color multiplex immunohistochemistry for simultaneous detection of multiple immune checkpoint molecules within the tumor microenvironment. *J Immunol*. 2018;200:347–354. doi:10.4049/jimmunol.1701262.
- Tsujikawa T, Kumar S, Borkar RN, Azimi V, Thibault G, Chang YH, Balter A, Kawashima R, Choe G, Sauer D, et al. Quantitative multiplex immunohistochemistry reveals myeloid-inflamed tumor-immune complexity associated with poor prognosis. *Cell Rep*. 2017;19:203–217. doi:10.1016/j.celrep.2017.03.037.

20. Giraldo NA, Becht E, Vano Y, Petitprez F, Lacroix L, Validire P, Sanchez-Salas R, Ingels A, Oudard S, Moatti A, et al. Tumor-infiltrating and peripheral blood T-cell immunophenotypes predict early relapse in localized clear cell renal cell carcinoma. *Clin Cancer Res.* 2017;23:4416–4428. doi:10.1158/1078-0432.CCR-16-2848.
21. Zucchetto A, Cattarossi I, Nanni P, Zaina E, Prato G, Gilestro M, Marconi D, Bulian P, Rossi FM, Del Vecchio L, et al. Cluster analysis of immunophenotypic data: the example of chronic lymphocytic leukemia. *Immunol Lett.* 2011;134:137–144. doi:10.1016/j.imlet.2010.09.017.
22. Breiman L. Random forests. *Mach Learn.* 2001;45:5–32. doi:10.1023/A:1010933404324.
23. Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. *J Comput Syst Sci.* 1997;55:119–139. doi:10.1006/jcss.1997.1504.
24. Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and regression trees. Monterey (CA): Wadsworth and Brooks; 1984.
25. Van den Hove LE, Vandenbergh P, Van Gool SW, Ceuppens JL, Demuynck H, Verhoef GE, Boogaerts MA. Peripheral blood lymphocyte subset shifts in patients with untreated hematological tumors: evidence for systemic activation of the T cell compartment. *Leuk Res.* 1998;22:175–184.
26. Christopoulos P, Pfeifer D, Bartholome K, Follo M, Timmer J, Fisch P, Veelken H. Definition and characterization of the systemic T-cell dysregulation in untreated indolent B-cell lymphoma and very early CLL. *Blood.* 2011;117:3836–3846. doi:10.1182/blood-2010-07-299321.
27. Scrivener S, Kaminski ER, Demaine A, Prentice AG. Analysis of the expression of critical activation/interaction markers on peripheral blood T cells in B-cell chronic lymphocytic leukaemia: evidence of immune dysregulation. *Br J Haematol.* 2001;112:959–964.
28. Gothert JR, Eisele L, Klein-Hitpass L, Weber S, Zesewitz ML, Sellmann L, Röth A, Pircher H, Dühsen U, Dürig J. Expanded CD8+ T cells of murine and human CLL are driven into a senescent KLRG1+ effector memory phenotype. *Cancer Immunol Immunother.* 2013;62:1697–1709. doi:10.1007/s00262-013-1473-z.
29. Palma M, Gentilcore G, Heimersson K, Mozaffari F, Nasman-Glaser B, Young E, Rosenquist R, Hansson L, Österborg A, Mellstedt H. T cells in chronic lymphocytic leukemia display dysregulated expression of immune checkpoints and activation markers. *Haematologica.* 2017;102:562–572. doi:10.3324/haematol.2016.151100.
30. Burger JA, Gribben JG. The microenvironment in chronic lymphocytic leukemia (CLL) and other B cell malignancies: insight into disease biology and new targeted therapies. *Semin Cancer Biol.* 2014;24:71–81. doi:10.1016/j.semcancer.2013.08.011.
31. de Rooij MF, Kuil A, Geest CR, Eldering E, Chang BY, Buggy JJ, Pals ST, Spaargaren M. The clinically active BTK inhibitor PCI-32765 targets B-cell receptor- and chemokine-controlled adhesion and migration in chronic lymphocytic leukemia. *Blood.* 2012;119:2590–2594. doi:10.1182/blood-2011-11-390989.
32. Chen SS, Chang BY, Chang S, Tong T, Ham S, Sherry B, Burger JA, Rai KR, Chiorazzi N. BTK inhibition results in impaired CXCR4 chemokine receptor surface expression, signaling and function in chronic lymphocytic leukemia. *Leukemia.* 2016;30:833–843. doi:10.1038/leu.2015.316.
33. Calissano C, Damle RN, Marsilio S, Yan XJ, Yancopoulos S, Hayes G, Emson C, Murphy EJ, Hellerstein MK, Sison C, et al. Intracellular complexity in chronic lymphocytic leukemia: fractions enriched in recently born/divided and older/quiescent cells. *Mol Med.* 2011;17:1374–1382. doi:10.2119/molmed.2011.00360.
34. Herishanu Y, Perez-Galan P, Liu D, Biancotto A, Pittaluga S, Vire B, Gibellini F, Njuguna N, Lee E, Stennett L, et al. The lymph node microenvironment promotes B-cell receptor signaling, NF-kappaB activation, and tumor proliferation in chronic lymphocytic leukemia. *Blood.* 2011;117:563–574. doi:10.1182/blood-2010-05-284984.
35. Gassner FJ, Zaborsky N, Catakovic K, Rebhandl S, Huemer M, Egle A, Hartmann TN, Greil R, Geisberger R. Chronic lymphocytic leukaemia induces an exhausted T cell phenotype in the TCL1 transgenic mouse model. *Br J Haematol.* 2015;170:515–522. doi:10.1111/bjh.13467.
36. Wu J, Xu X, Lee EJ, Shull AY, Pei L, Awan F, Wang X, Choi J-H, Deng L, Xin H-B, et al. Phenotypic alteration of CD8+ T cells in chronic lymphocytic leukemia is associated with epigenetic reprogramming. *Oncotarget.* 2016;7:40558–40570. doi:10.18632/oncotarget.9941.
37. Long M, Beckwith K, Do P, Mundy BL, Gordon A, Lehman AM, Maddocks KJ, Cheney C, Jones JA, Flynn JM, et al. Ibrutinib treatment improves T cell number and function in CLL patients. *J Clin Invest.* 2017;127:3052–3064. doi:10.1172/JCI89756.
38. Riches JC, Davies JK, McClanahan F, Fatah R, Iqbal S, Agrawal S, Ramsay AG, Gribben JG. T cells from CLL patients exhibit features of T-cell exhaustion but retain capacity for cytokine production. *Blood.* 2013;121:1612–1621. doi:10.1182/blood-2012-09-457531.
39. Tonino SH, van de Berg PJ, Yong SL, Ten Berge IJ, Kersten MJ, van Lier RA, van Oers MH, Kater AP. Expansion of effector T cells associated with decreased PD-1 expression in patients with indolent B cell lymphomas and chronic lymphocytic leukemia. *Leuk Lymphoma.* 2012;53:1785–1794. doi:10.3109/10428194.2012.673224.
40. McClanahan F, Riches JC, Miller S, Day WP, Kotsiou E, Neuberger D, Croce CM, Capasso M, Gribben JG. Mechanisms of PD-L1/PD-1-mediated CD8 T-cell dysfunction in the context of aging-related immune defects in the Emicro-TCL1 CLL mouse model. *Blood.* 2015;126:212–221. doi:10.1182/blood-2015-02-626754.
41. Tinhofer I, Weiss L, Gassner F, Rubenzer G, Holler C, Greil R. Difference in the relative distribution of CD4+ T-cell subsets in B-CLL with mutated and unmutated immunoglobulin (Ig) VH genes: implication for the course of disease. *J Immunother.* 2009;32:302–309. doi:10.1097/CJI.0b013e318197b5e4.
42. Rissiek A, Schulze C, Bacher U, Schieferdecker A, Thiele B, Jacholkowski A, Flammiger A, Horn C, Haag F, Tiegs G, et al. Multidimensional scaling analysis identifies pathological and prognostically relevant profiles of circulating T-cells in chronic lymphocytic leukemia. *Int J Cancer.* 2014;135:2370–2379. doi:10.1002/ijc.28884.
43. Sommermeyer D, Hudecek M, Kosasih PL, Gogishvili T, Maloney DG, Turtle CJ, Riddell SR. Chimeric antigen receptor-modified T cells derived from defined CD8+ and CD4+ subsets confer superior antitumor reactivity in vivo. *Leukemia.* 2016;30:492–500. doi:10.1038/leu.2015.247.
44. Fraietta JA, Lacey SF, Orlando EJ, Pruteanu-Malinici I, Gohil M, Lundh S, Boesteanu AC, Wang Y, O'Connor RS, Hwang W-T, et al. Determinants of response and resistance to CD19 chimeric antigen receptor (CAR) T cell therapy of chronic lymphocytic leukemia. *Nat Med.* 2018;24:563–571. doi:10.1038/s41591-018-0010-1.
45. R Core Team. R: A language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing; 2017.
46. Rossum G. Python reference manual. Amsterdam, The Netherlands: CWI (Centre for Mathematics and Computer Science); 1995.