



(51) International Patent Classification:

G01N 33/50 (2006.01) G01N 33/574 (2006.01)
G01N 33/569 (2006.01)

(21) International Application Number:

PCT/EP2020/050553

(22) International Filing Date:

10 January 2020 (10.01.2020)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

19305038.2 11 January 2019 (11.01.2019) EP

(71) Applicants: **INSERM (INSTITUT NATIONAL DE LA SANTÉ ET DE LA RECHERCHE MÉDICALE)** [FR/FR]; 101, rue de Tolbiac, 75013 Paris (FR). **UNIVERSITÉ PAUL SABATIER TOULOUSE III** [FR/FR]; 118 route de Narbonne, 31400 Toulouse (FR). **CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS)** [FR/FR]; 3, Rue Michel Ange, 75016 Paris (FR). **CENTRE HOSPITALIER UNIVERSITAIRE DE TOULOUSE** [FR/FR]; 2, rue Viguierie Hôtel-Dieu Pont Neuf, 31300 Toulouse (FR). **UNIVERSITÉ**

TOULOUSE 1 CAPITOLE [FR/FR]; 2 rue du Doyen Gabriel Marty, 31000 Toulouse (FR).

(72) Inventors: **GONNORD, Pauline**; 3, rue du port, 17650 St Denis d'Oleron (FR). **VALITUTTI, Salvatore**; UMR 1037, CRCT, 2 avenue Hubert Curien, CS 53717, 31037 Toulouse (FR). **YSEBAERT, Loïc**; INSERM U1037 – CRCT, 2 avenue Hubert Curien, CS 53717, 31037 Toulouse Cedex 1 (FR). **GADAT, Sébastien**; Toulouse School of Economics, Université Toulouse 1 Capitole, 1, Esplanade de l'Université, 31080 Toulouse Cedex 06 (FR). **COSTA, Manon**; Institut de Mathématique de Toulouse, Université Paul Sabatier, 118 Route de Narbonne, 31400 Toulouse (FR).

(74) Agent: **INSERM TRANSFERT**; 7 rue Watt, 75013 Paris (FR).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA,

(54) Title: METHOD TO PREDICT THE NEED FOR THERAPY OF PATIENTS SUFFERING FROM A CANCER

A

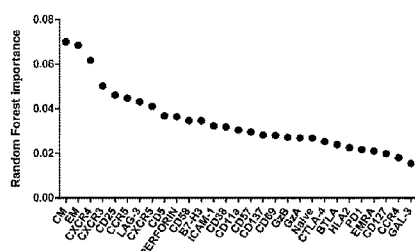


Figure 1A

B

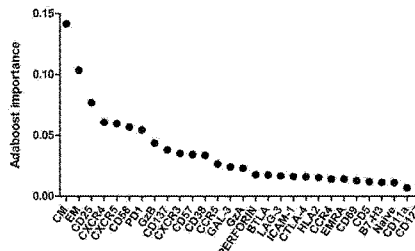


Figure 1B

(57) Abstract: The present invention relates to the prediction of the need for therapy of patients suffering from a cancer. Here, the inventors undertook an unbiased approach for multidimensional characterization of CD8⁺ T cell phenotypic signature to investigate possible global CD8⁺ T cell phenotypic remodeling in CLL patients. They 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. They identify a phenotypic signature combining at least 2 unrelated markers that describes the evolution towards therapy within 6 months after phenotyping in fresh whole or frozen blood samples, which was never done before and which allow the stratification of patients. 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. The invention relates to a method for identifying phenotypic signatures within CD8⁺ lymphocytes to predict the need for therapy of a patient suffering from a cancer comprising several steps.

SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR,
TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

METHOD TO PREDICT THE NEED FOR THERAPY OF PATIENTS SUFFERING FROM A CANCER

FIELD OF THE INVENTION:

The invention relates to a method for identifying phenotypical signatures within CD8⁺ lymphocytes to predict the need for therapy of a patient suffering from a cancer comprising several steps.

BACKGROUND OF THE INVENTION:

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 a 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.

5 SUMMARY OF THE INVENTION:

Herein, the inventors undertook an unbiased approach for multi-dimensional characterization of CD8⁺ T cell phenotypic signature to investigate possible global CD8⁺ T cell phenotypic remodeling in CLL patients. They centered their study at the patient level so that multiple marker expression could be compared in multiple patients at the same time by
10 implementing approaches for statistical multi-dimensional analysis of multicolor flow cytometry data sets.

There 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
15 patients that reflects disease progression towards therapy and is mainly due to imbalance in the memory compartment. They identify a phenotypic signature combining at least 2 unrelated markers that describes the evolution towards therapy within 6 months after phenotyping in fresh whole or frozen blood samples, which was never done before and which allow the stratification of patients.

20 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.

Thus, the invention relates to a method for identifying phenotypical signatures within CD8⁺ lymphocytes to predict the need for therapy of a patient suffering from a cancer comprising several steps. Particularly, the invention is defined by its claims

25

DETAILED DESCRIPTION OF THE INVENTION:

A first aspect of the invention relates to a method for identifying phenotypical signatures within CD8⁺ lymphocytes to predict the need for therapy of a patient suffering from a cancer comprising:

30 i) Determining, in a sample obtained from the patient, the expression level of specific markers of CD8⁺ T lymphocytes;

ii) Comparing the expression level of the markers determined at step i) with the background reference values obtained from unstained samples and calculate the percentage of cells expressing the markers above the background of the unstained sample.

iii) Create a database with all the normalized marker expression values obtained on N patients for which the disease progression towards need for therapy within 4 to 8 months is known (learning cohort);

iv) Perform a supervised analysis of the obtained database to identify the most relevant markers expressed by CD8⁺ T lymphocytes to explain the evolution towards therapy.

When the markers are identified, the calculation of a logistic regression and the calculation of the optimal threshold can be done to distinguish the patients that will evolve towards therapy.

Thus, in another embodiment, the invention also relates to a method for identifying phenotypical signatures within CD8⁺ lymphocytes to predict the need for therapy of a patient suffering from a cancer comprising:

i) Determining, in a sample obtained from the patient, the expression level of specific markers of CD8⁺ T lymphocytes;

ii) Comparing the expression level of the markers determined at step i) with the background reference values obtained from unstained samples and calculate the percentage of cells expressing the markers above the background of the unstained sample.

iii) Create a database with all the normalized marker expression values obtained on N patients for which the disease progression towards need for therapy within 4 to 8 months is known (learning cohort);

iv) Perform a supervised analysis of the obtained database to identify the most relevant markers expressed by CD8⁺ T lymphocytes to explain the evolution towards therapy;

v) Calculate the logistic regression that describes the probability function of the treatment status of the patients as a function of the expression levels of the selected markers from step iv) and that goes from 0 (the patient is not treated) to 1 (the patient is treated);

vi) Calculate the optimal threshold of this function to distinguish the patients that will evolve towards therapy for which the score (S1), defined as the result of the logistic regression function, is above the optimal threshold versus the patients that will not evolve towards therapy for which the score (S1) is below the optimal threshold.

Then, a step of validation of the accuracy of prediction of the logistic regression can be done.

Thus, in another embodiment, the invention also relates to a method for identifying phenotypical signatures within CD8⁺ lymphocytes to predict the need for therapy of a patient suffering from a cancer comprising:

- 5 i) Determining, in a sample obtained from the patient, the expression level of specific markers of CD8⁺ T lymphocytes;
- ii) Comparing the expression level of the markers determined at step i) with the background reference values obtained from unstained samples and calculate the percentage of cells expressing the markers above the background of the unstained sample.
- 10 iii) Create a database with all the normalized marker expression values obtained on N patients for which the disease progression towards need for therapy within 4 to 8 months is known (learning cohort);
- iv) Perform a supervised analysis of the obtained database to identify the most relevant markers expressed by CD8⁺ T lymphocytes to explain the evolution towards therapy;
- 15 v) Calculate the logistic regression that describes the probability function of the treatment status of the patients as a function of the expression levels of the selected markers from step iv) and that goes from 0 (the patient is not treated) to 1 (the patient is treated);
- vi) Calculate the optimal threshold of this function to distinguish the patients that will evolve towards therapy for which the score (S1), defined as the result of the logistic regression function, is above the optimal threshold versus the patients that will not evolve towards therapy
- 20 for which the score (S1) is below the optimal threshold;
- vii) Validate the accuracy of prediction of the logistic regression using a prediction cohort and calculate accuracy and F1 measure of the prediction by comparing prediction and real life values.

According to the invention, the specific markers of CD8⁺ T lymphocytes are CM, EM, 25 CXCR4, CXCR3, CD25, CCR5, LAG-3, CXCR5, CD5, Perforin, CD58, B7-H3, ICAM-1, CD38, CD11a, CD57, CD137, CD69, GzB, GzA, Naïve, CTLA-4, BTLA, HLA2, PD1, EMRA, CD127, CCR4, GAL-3.

According to the invention, the terms “N” denotes at least 10 patients.

According to the inventions, the term “N” denotes at least of 10, 11, 12, 13, 14, 15, 16, 30 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more patients.

According to the invention, the supervised analysis can be performed for example by using Random Forest algorithm, Adaboost algorithm, Decision tree algorithm (CART), Linear

Discriminant analysis, Topological data analysis, Support Vector Machine algorithm or Neural Networks.

In a particular embodiment, the supervised analysis can be performed by using Random Forest algorithm, Adaboost algorithm or Decision tree algorithm.

5 In preferred embodiment, the supervised analysis can be performed by using Random Forest algorithm.

In a particular embodiment, the time wherein the disease will progress towards need for therapy can be of 6 months.

10 As used herein, the term "Random Forests algorithm" or "RF" has its general meaning in the art and refers to classification algorithm. Random Forest is a decision-tree-based classifier that is constructed using an algorithm originally developed by Leo Breiman (Breiman L, "Random forests," Machine Learning 2001, 45:5-32).

15 As used herein, the term "Adaboost algorithm" has its general meaning in the art and refers to classification algorithm using an algorithm originally developed by Freund and Schapire (Freund Y. and Schapire R.E. "A decision-theoretic generalization of on-line learning and an application to boosting". Journal of Computer and System Sciences, 55(1):119-139, 1997).

20 As used herein, the term "support vector machine (SVM)" has its general meaning in the art and refers to a universal learning machine useful as a statistical tool for classification and using an algorithm developed by Cortes and Vapnik (Cortes C. and Vapnik V.N. "Support-vector networks" Machine Learning 1995, 20(3):273-297).

25 As used herein the term "S1" denotes the score of a given patient that is calculated with the markers expression levels using the formula of the logistic regression previously determined from the learning cohort. This score represent the probability of a given patient to evolve towards therapy within 4 to 8 months, particularly 6 months based on the phenotype of its CD8⁺ T lymphocyte compartment and particularly based on specific CD8⁺ T lymphocytes markers.

30 As used herein, the term "learning cohort" denotes the cohort of patients that were phenotyped using the X markers from which the relevant markers of the need for therapy are learned using supervised learning algorithms and from which the logistic regression (probability function) is identified and fitted thanks to the phenotyping data.

As used herein the term "accuracy" denotes the statistical calculation of the fidelity of prediction of a method with reality. The accuracy can be calculated from the confusion matrix using the formula: (True positive + True negative)/(True positive + False Positive + False negative + True negative). True positive corresponds to the number of patients that are predicted

positive for the considered outcome and are in reality positive. False positive corresponds to the number of patients that are predicted positive but are in reality negative. False negative corresponds to the number of patients that are predicted negative but are in reality positive. True negative corresponds to the number of patients that are predicted negative and are in reality negative.

As used herein, the term “F1” denotes another statistical calculation of the fidelity of prediction of a method with reality. The F1 measure can be calculated using the formula: $2 * ((\text{precision} * \text{recall}) / (\text{precision} + \text{recall}))$. Precision corresponds to: $(\text{True positive}) / (\text{True positive} + \text{False positive})$. Recall corresponds to: $(\text{True positive}) / (\text{True positive} + \text{False negative})$.

As used herein, the term “optimal threshold” correspond to the threshold value of the S1 score that will split the patients in 2 groups: S1 above the threshold will corresponds to the group of patients that will evolve towards therapy within 4 to 8 months and particularly 6 months, S1 below the threshold will corresponds to the group of patients that will not need therapy. This threshold value of the S1 score is qualified of optimal because it is calculated to obtain the value of S1 that will give the best accuracy of prediction with the learning cohort. In this threshold calculation, the number of False positive and False negative are weighted differentially in order to minimize the amount of False negative (Patients that are predicted as not needing therapy but needing therapy in reality and thus won't be treated when they should have been). This weight is given by applying a penalty (also called cost) to False negative that is twice as much important as the penalty of False positive.

As used herein, the term “background reference value” denotes the minimal expression of a marker in CD8⁺ T lymphocyte.

Another aspect of invention relates to a method for predicting the need for therapy of a patient suffering from a cancer comprising: i) determining, in a sample obtained from the patient, the expression level of the markers of step iv) of the method described above ; ii) calculating the S1 score of the patient, using the expression levels from step i), according to the logistic regression defined from a cancer learning cohort according to step v) of the method described above ; iii) comparing the S1 score obtained for a given patient, to the optimal threshold value defined from the cancer learning cohort ; iv) providing a prognosis of evolution toward need for therapy within 4 to 8 months when the S1 score determined at step ii) is higher than the optimal threshold value, or providing a prognosis of not needing therapy within 4 to 8 months when the S1 score is lower than the optimal threshold.

According to the invention, the cancer may be selected in the group consisting of adrenal cortical cancer, anal cancer, bile duct cancer, bladder cancer, bone cancer, brain and central nervous system cancer, breast cancer, Castleman disease, cervical cancer, colorectal cancer, endometrial cancer, esophagus cancer, gallbladder cancer, gastrointestinal carcinoid tumors, Hodgkin's disease, non-Hodgkin's lymphoma, Kaposi's sarcoma, kidney cancer, laryngeal and hypopharyngeal cancer, liver cancer, lung cancer, mesothelioma, plasmacytoma, nasal cavity and paranasal sinus cancer, nasopharyngeal cancer, neuroblastoma, oral cavity and oropharyngeal cancer, ovarian cancer, pancreatic cancer, penile cancer, pituitary cancer, prostate cancer, retinoblastoma, rhabdomyosarcoma, salivary gland cancer, skin cancer, stomach cancer, testicular cancer, thymus cancer, thyroid cancer, vaginal cancer, vulvar cancer, and uterine cancer.

In a particular embodiment, the cancer is a chronic lymphocytic leukemia (CLL).

According to the invention, the markers of the invention are interne or surface markers. According to the invention, "surface markers" denotes proteins express on the surface of CD8⁺ T lymphocytes. According to the invention, "interne markers" denotes proteins express inside CD8⁺ T lymphocytes.

In a particular embodiment, the marker are surface markers.

In another aspect, the invention also relates to a method for predicting the need for therapy of a patient suffering from a cancer comprising: i) determining, in a sample obtained from the patient, the expression level of markers expressed by the CD8⁺ T lymphocytes; ii) calculating the CD8⁺ T cell compartment S1 score of the patient, using the expression levels from step i), according to the logistic regression defined from a cancer learning cohort; iii) comparing the S1 score obtained for a given patient, to the optimal threshold value defined from the cancer learning cohort ; iv) providing a prognosis of evolution toward therapy within 4 to 8 months when the S1 score determined at step ii) is higher than the optimal threshold value, or providing a prognosis of not needing therapy within 4 to 8 months when the S1 score is lower than the optimal threshold.

Particularly, the inventors have used the method described above to determine a set of marker specific of the chronic lymphocytic leukemia (CLL) and to determine the need for therapy of a specific patient.

Thus, according to the invention, when the method is specific of CLL, the markers expressed by CD8 + T cells whose expression level is determined in step i) are CM and EM.

In one embodiment, when the method is specific of CLL, the markers expressed by CD8 + T cells whose expression level is determined in step i) are CM, EM and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26 or 27 markers selected among CXCR4, CXCR3, CD25, CCR5, LAG-3, CXCR5, CD5, Perforin, CD58, B7-H3, ICAM-1, CD38, CD11a, CD57, CD137, CD69, GzB, GzA, Naïve, CTLA-4, BTLA, HLA2, PD1, EMRA, CD127, CCR4, GAL-3.

In one embodiment, when the method is specific of CLL, the markers expressed by CD8 + T cells whose expression level is determined in step i) are CM, EM and CXCR4.

In one embodiment, when the method is specific of CLL, the markers expressed by CD8 + T cells whose expression level is determined in step i) are CM, EM, CXCR4 and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 markers selected among CXCR3, CD25, CCR5, LAG-3, CXCR5, CD5, Perforin, CD58, B7-H3, ICAM-1, CD38, CD11a, CD57, CD137, CD69, GzB, GzA, Naïve, CTLA-4, BTLA, HLA2, PD1, EMRA, CD127, CCR4, GAL-3.

In one embodiment, when the method is specific of CLL, the markers expressed by CD8 + T cells whose expression level is determined in step i) are CM, EM, CXCR4 and CXCR3.

In one embodiment, when the method is specific of CLL, the markers expressed by CD8 + T cells whose expression level is determined in step i) are CM, EM, CXCR4 and CD25.

In one embodiment, when the method is specific of CLL, the markers expressed by CD8 + T cells whose expression level is determined in step i) are CM, EM, CXCR4 and CCR5.

In one embodiment, when the method is specific of CLL, the markers expressed by CD8 + T cells whose expression level is determined in step i) are CM, EM, CXCR4, CXCR3 and CD25.

In one embodiment, when the method is specific of CLL, the markers expressed by CD8 + T cells whose expression level is determined in step i) are CM, EM, CXCR4, CXCR3 and CCR5.

In one embodiment, when the method is specific of CLL, the markers expressed by CD8 + T cells whose expression level is determined in step i) are CM, EM, CXCR4, CCR5 and CD25.

In one embodiment, when the method is specific of CLL, the markers expressed by CD8 + T cells whose expression level is determined in step i) are CM, EM, CXCR4, CXCR3, CD25 and CCR5.

Thus, the invention also relates to a method for predicting the need for therapy of a patient suffering from a CLL comprising: i) determining, in a sample obtained from the patient, the expression level of the markers CM and EM expressed by the CD8⁺ T lymphocytes; ii) calculating the CD8⁺ T cell compartment S1 score of the patient, using the expression levels from step i), according to the logistic regression defined from a cancer learning cohort; iii) comparing the S1 score obtained for a given patient, to the optimal threshold value defined from the cancer learning cohort ; iv) providing a prognosis of evolution toward therapy within 4 to 8 months when the S1 score determined at step ii) is higher than the optimal threshold value, or providing a prognosis of not needing therapy within 4 to 8 months when the S1 score is lower than the optimal threshold.

In a particular embodiment, the time wherein the disease will progress towards need for therapy can be of 6 months.

In one embodiment, a step of v) communicating the result to the patient may be added to the method for predicting the need for therapy of a patient suffering from a chronic lymphocytic leukemia (CLL).

Thus, in another aspect the invention also relates to a method for predicting the need for therapy of a patient suffering from a CLL comprising: i) determining, in a sample obtained from the patient, the expression level of the markers CM, EM and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26 or 27 markers selected in the group consisting in CXCR4, CXCR3, CD25, CCR5, LAG-3, CXCR5, CD5, Perforin, CD58, B7-H3, ICAM-1, CD38, CD11a, CD57, CD137, CD69, GzB, GzA, Naïve, CTLA-4, BTLA, HLA2, PD1, EMRA, CD127, CCR4 and GAL-3 expressed by the CD8⁺ T lymphocytes; ii) calculating the CD8⁺ T cell compartment S1 score of the patient, using the expression levels from step i), according to the logistic regression defined from a cancer learning cohort; iii) comparing the S1 score obtained for a given patient, to the optimal threshold value defined from the cancer learning cohort ; iv) providing a prognosis of evolution toward therapy within 4 to 8 months when the S1 score determined at step ii) is higher than the optimal threshold value, or providing a prognosis of not needing therapy within 4 to 8 months when the S1 score is lower than the optimal threshold value.

In another aspect the invention also relates to a method for predicting the need for therapy of a patient suffering from a CLL comprising: i) determining, in a sample obtained from the patient, the expression level of the markers CM, EM and CXCR4 expressed by the CD8⁺ T lymphocytes; ii) calculating the CD8⁺ T cell compartment S1 score of the patient, using the expression levels from step i), according to the logistic regression defined from a cancer

learning cohort; iii) comparing the S1 score obtained for a given patient, to the optimal threshold value defined from the cancer learning cohort ; iv) providing a prognosis of evolution toward therapy within 4 to 8 months when the S1 score determined at step ii) is higher than the optimal threshold value, or providing a prognosis of not needing therapy within 4 to 8 months when the S1 score is lower than the optimal threshold value

In another aspect the invention also relates to a method for predicting the need for therapy of a patient suffering from a CLL comprising: i) determining, in a sample obtained from the patient, the expression level of the markers CM, EM, CXCR4 and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 markers selected in the group consisting in CXCR3, CD25, CCR5, LAG-3, CXCR5, CD5, Perforin, CD58, B7-H3, ICAM-1, CD38, CD11a, CD57, CD137, CD69, GzB, GzA, Naïve, CTLA-4, BTLA, HLA2, PD1, EMRA, CD127, CCR4 and GAL-3 expressed by the CD8⁺ T lymphocytes; ii) calculating the CD8⁺ T cell compartment S1 score of the patient, using the expression levels from step i), according to the logistic regression defined from a cancer learning cohort; iii) comparing the S1 score obtained for a given patient, to the optimal threshold value defined from the cancer learning cohort ; iv) providing a prognosis of evolution toward therapy within 4 to 8 months when the S1 score determined at step ii) is higher than the optimal threshold value, or providing a prognosis of not needing therapy within 4 to 8 months when the score is lower than the optimal threshold value.

As used herein, the term “need for therapy” denotes the evolution of patients suffering from CLL to an advanced stage which require a treatment during the six months following the phenotypic analysis (i.e. the determination of the expression level of the markers according to the invention).

In other word, the invention relates to a method for predicting the need for therapy of a patient suffering from a chronic lymphocytic leukemia (CLL) comprising: i) determining, in a sample obtained from the patient, the expression level of CD8⁺ T lymphocytes expressing the markers CM and EM ii) calculating the CD8⁺ T cell compartment S1 score of the patient; iii) comparing the S1 score obtained for a given patient, to the optimal threshold value defined from the cancer learning cohort, iv) providing a prognosis of evolution toward therapy within 4 to 8 months when the S1 score determined at step ii) is higher than the optimal threshold value, or providing a prognosis of not needing therapy within 4 to 8 months when the S1 score is lower than the optimal threshold.

In one embodiment, the measurement of the expression level of at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26 or 27 markers selected

in the group consisting of CXCR4, CXCR3, CD25, CCR5, LAG-3, CXCR5, CD5, Perforin, CD58, B7-H3, ICAM-1, CD38, CD11a, CD57, CD137, CD69, GzB, GzA, Naïve, CTLA-4, BTLA, HLA2, PD1, EMRA, CD127, CCR4 and GAL-3 expressed by the CD8⁺ T lymphocytes may be added to the method for predicting the need for therapy of a patient suffering from a chronic lymphocytic leukemia.

Thus in another aspect, the invention relates to a method for predicting the need for therapy of a patient suffering from a chronic lymphocytic leukemia (CLL) comprising: i) determining, in a sample obtained from the patient, the expression level of CD8⁺ T lymphocytes expressing the markers CM, EM and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26 or 27 markers selected in the group consisting in CXCR4, CXCR3, CD25, CCR5, LAG-3, CXCR5, CD5, Perforin, CD58, B7-H3, ICAM-1, CD38, CD11a, CD57, CD137, CD69, GzB, GzA, Naïve, CTLA-4, BTLA, HLA2, PD1, EMRA, CD127, CCR4 and GAL-3; ii) calculating the CD8⁺ T cell compartment (S1) score of the patient; iii) comparing the S1 score obtained for a given patient, to the optimal threshold value defined from the cancer learning cohort and iv) providing a prognosis of evolution toward therapy within 4 to 8 months when the S1 score determined at step ii) is higher than the optimal threshold value, or providing a prognosis of not needing therapy within 4 to 8 months when the S1 score is lower than the optimal threshold.

The invention also relates to a method for predicting the need for therapy of a patient suffering from a chronic lymphocytic leukemia (CLL) comprising: i) determining, in a sample obtained from the patient, the expression level of CD8⁺ T lymphocytes expressing the markers CM, EM and CXCR4 ii) calculating the CD8⁺ T cell compartment S1 score of the patient; iii) comparing the S1 score obtained for a given patient, to the optimal threshold value defined from the cancer learning cohort and iv) providing a prognosis of evolution toward therapy within 4 to 8 months when the S1 score determined at step ii) is higher than the optimal threshold value, or providing a prognosis of not needing therapy within 4 to 8 months when the S1 score is lower than the optimal threshold.

In one embodiment, the measurement of the expression level of at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 markers selected in the group consisting of CXCR3, CD25, CCR5, LAG-3, CXCR5, CD5, Perforin, CD58, B7-H3, ICAM-1, CD38, CD11a, CD57, CD137, CD69, GzB, GzA, Naïve, CTLA-4, BTLA, HLA2, PD1, EMRA, CD127, CCR4 and GAL-3 expressed by the CD8⁺ T lymphocytes may be added to the method for predicting the need for therapy of a patient suffering from a chronic lymphocytic leukemia.

Thus in another aspect, the invention relates to a method for predicting the need for therapy of a patient suffering from a chronic lymphocytic leukemia (CLL) comprising: i) determining, in a sample obtained from the patient, the expression level of CD8⁺ T lymphocytes expressing the markers CM, EM, CXCR4 and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 markers selected in the group consisting in CXCR3, CD25, CCR5, LAG-3, CXCR5, CD5, Perforin, CD58, B7-H3, ICAM-1, CD38, CD11a, CD57, CD137, CD69, GzB, GzA, Naïve, CTLA-4, BTLA, HLA2, PD1, EMRA, CD127, CCR4 and GAL-3; ii) calculating the CD8⁺ T cell compartment (S1) score of the patient; iii) comparing the S1 score obtained for a given patient, to the optimal threshold value defined from the cancer learning cohort and iv) providing a prognosis of evolution toward therapy within 4 to 8 months when the S1 score determined at step ii) is higher than the optimal threshold value, or providing a prognosis of not needing therapy within 4 to 8 months when the S1 score is lower than the optimal threshold.

Another aspect of the invention relates to a method for predicting the survival time of a patient suffering from a chronic lymphocytic leukemia (CLL) comprising: i) determining, in a sample obtained from the patient, the expression level of CD8⁺ T lymphocytes expressing the markers CM and EM; ii) calculating the CD8⁺ T cell compartment (S1) score of the patient; iii) comparing the expression level of the markers determined at step i) to the optimal threshold value defined from the cancer learning cohort and iv) providing a good prognosis when the S1 score determined at step ii) is higher than the optimal threshold value, or providing a bad prognosis when the S1 score is lower than the optimal threshold value.

In one embodiment, the measurement of the expression level of at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26 or 27 markers in the group consisting of CXCR4, CXCR3, CD25, CCR5, LAG-3, CXCR5, CD5, Perforin, CD58, B7-H3, ICAM-1, CD38, CD11a, CD57, CD137, CD69, GzB, GzA, Naïve, CTLA-4, BTLA, HLA2, PD1, EMRA, CD127, CCR4 and GAL-3 may be added to the method for predicting the survival time of a patient suffering from a chronic lymphocytic leukemia.

Thus, in another aspect, the invention relates to a method for predicting the survival time of a patient suffering from a chronic lymphocytic leukemia comprising i) determining in a sample obtained from the patient the expression level of the markers CM, EM and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26 or 27 markers selected in the group consisting in CXCR4, CXCR3, CD25, CCR5, LAG-3, CXCR5, CD5, Perforin, CD58, B7-H3, ICAM-1, CD38, CD11a, CD57, CD137, CD69, GzB, GzA, Naïve,

CTLA-4, BTLA, HLA2, PD1, EMRA, CD127, CCR4 and GAL-3 ii) calculating the CD8⁺ T cell compartment (S1) score of the patient iii) comparing the expression level of the markers determined at step i) to the optimal threshold value defined from the cancer learning cohort and iv) providing a good prognosis when the S1 score determined at step ii) is higher than the optimal threshold value, or providing a bad prognosis when the S1 score is lower than the optimal threshold value.

Another aspect of the invention relates to a method for predicting the survival time of a patient suffering from a chronic lymphocytic leukemia (CLL) comprising: i) determining, in a sample obtained from the patient, the expression level of CD8⁺ T lymphocytes expressing the markers CM, EM and CXCR4 ii) calculating the CD8⁺ T cell compartment (S1) score of the patient iii) comparing the expression level of the markers determined at step i) to the optimal threshold value defined from the cancer learning cohort and iv) providing a good prognosis when the S1 score determined at step ii) is higher than the optimal threshold value, or providing a bad prognosis when the S1 score is lower than the optimal threshold value.

In one embodiment, the measurement of the expression level of at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 markers in the group consisting of CXCR3, CD25, CCR5, LAG-3, CXCR5, CD5, Perforin, CD58, B7-H3, ICAM-1, CD38, CD11a, CD57, CD137, CD69, GzB, GzA, Naïve, CTLA-4, BTLA, HLA2, PD1, EMRA, CD127, CCR4 and GAL-3 may be added to the method for predicting the survival time of a patient suffering from a chronic lymphocytic leukemia.

Thus, in another aspect, the invention relates to a method for predicting the survival time of a patient suffering from a chronic lymphocytic leukemia comprising i) determining in a sample obtained from the patient the expression level of the markers CM, EM, CXCR4 and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 markers selected in the group consisting in CXCR3, CD25, CCR5, LAG-3, CXCR5, CD5, Perforin, CD58, B7-H3, ICAM-1, CD38, CD11a, CD57, CD137, CD69, GzB, GzA, Naïve, CTLA-4, BTLA, HLA2, PD1, EMRA, CD127, CCR4 and GAL-3 ii) calculating the CD8⁺ T cell compartment (S1) score of the patient iii) comparing the expression level of the markers determined at step i) to the optimal threshold value defined from the cancer learning cohort and iv) providing a good prognosis when the S1 score determined at step i) is higher than the optimal threshold value, or providing a bad prognosis when the S1 score is lower than the optimal threshold value.

Another aspect, the invention relates to a method for predicting the survival time of a patient suffering from a chronic lymphocytic leukemia comprising i) determining in a sample

obtained from the patient the expression level of the markers CM, EM, CXCR4, CXCR3, CD25 and CCR5 ii) calculating the CD8⁺ T cell compartment S1 score of the patient iii) comparing the expression level of the markers determined at step i) to the optimal threshold value defined from the cancer learning cohort and iv) providing a good prognosis when the score determined at step i) is higher than the optimal threshold value, or providing a bad prognosis when the score is lower than the optimal threshold value.

In another aspect, the invention relates to a method for predicting the need for therapy of a patient suffering from a cancer comprising : i) determining, in a sample obtained from the patient, the expression level of the markers expressed by the CD8⁺ T lymphocytes, ii) implementing a classification algorithm on data comprising the quantified plurality of markers expressed by the CD8⁺ T lymphocytes so as to obtain an algorithm output, iii) determining the probability that the patient will need for therapy from the algorithm output of step ii).

In one embodiment, when the method is specific of CLL, the markers expressed by CD8⁺ T cells whose expression level is determined in step i) are CM, EM and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26 or 27 markers selected above CXCR4, CXCR3, CD25, CCR5, LAG-3, CXCR5, CD5, Perforin, CD58, B7-H3, ICAM-1, CD38, CD11a, CD57, CD137, CD69, GzB, GzA, Naïve, CTLA-4, BTLA, HLA2, PD1, EMRA, CD127, CCR4, GAL-3.

In one embodiment, the classification algorithm implemented in step ii) is Random Forest algorithm, Adaboost algorithm, Decision tree algorithm (CART), Linear Discriminant analysis, Topological data analysis, Support Vector Machine algorithm or Neural Networks.

Another aspect of the invention relates to a method to stratify a patient suffering from a chronic lymphocytic leukemia (CLL) according to its treatment need comprising determining, in a sample obtained from the patient, the expression level of the markers CM, EM.

Another aspect of the invention relates to a method to stratify a patient suffering from a chronic lymphocytic leukemia (CLL) according to its treatment need comprising determining, in a sample obtained from the patient, the expression level of the markers CM, EM and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26 or 27 markers selected in the group consisting in CXCR4 CXCR3, CD25, CCR5, LAG-3, CXCR5, CD5, Perforin, CD58, B7-H3, ICAM-1, CD38, CD11a, CD57, CD137, CD69, GzB, GzA, Naïve, CTLA-4, BTLA, HLA2, PD1, EMRA, CD127, CCR4 and GAL-3.

Another aspect of the invention relates to a method to stratify a patient suffering from a chronic lymphocytic leukemia (CLL) according to its treatment need comprising determining, in a sample obtained from the patient, the expression level of the markers CM, EM and CXCR4.

Another aspect of the invention relates to a method to stratify a patient suffering from a chronic lymphocytic leukemia (CLL) according to its treatment need comprising determining, in a sample obtained from the patient, the expression level of the markers CM, EM, CXCR4 and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26 or 27 markers selected in the group consisting in CXCR3, CD25, CCR5, LAG-3, CXCR5, CD5, Perforin, CD58, B7-H3, ICAM-1, CD38, CD11a, CD57, CD137, CD69, GzB, GzA, Naïve, CTLA-4, BTLA, HLA2, PD1, EMRA, CD127, CCR4 and GAL-3.

As used herein, the terms “to stratify a patient suffering from a chronic lymphocytic leukemia according to its treatment need” denotes the possibility to determine a numerical value (score) that will reflect the disease progression stage toward treatment of the patient suffering from a CLL and thus determine the need for therapy of said patient. CLL staging systems already exist (Binet or Rai classification, Scarfo L et al. *Crit Rev Oncol Hematol*, 2016) but the present invention relates to treatment need evaluation. If one consider the disease progression scale from diagnosis to time to first treatment, the present invention stratify the patients within this time frame.

As used herein, the term “survival time” denotes the expected duration of time until death of a patient suffering from cancer or CLL.

As used herein, the term “Good Prognosis” denotes a significantly more favourable probability of survival after patient treatment in the group of patients defined as “good prognosis” compared with the group of patients defined as “bad prognosis”. According to the invention, the term “Good Prognosis” also denotes a significantly more favourable probability of not needing treatment to survive in the group of patients defined as “good prognosis” compared with the group of patients defined as “bad prognosis”.

As used herein and according to all aspects of the invention, the term “sample” denotes, blood, fresh whole blood, peripheral-blood or peripheral blood mononuclear cell (PBMC).

In one embodiment, the sample and more particularly the PBMC, is not frozen.

In one embodiment, the sample, and more particularly the PBMC, is frozen.

As used herein, the term "patient" refers to an individual who is being managed for a cancer like a CLL disease and who is susceptible to develop a cancer like a CLL at any stage of the disease like at more advanced Binet stage (B and C) (Scarfo L et al. *Crit Rev Oncol Hematol*, 2016).

According to the invention and the CLL cancer, the terms “markers (or biomarkers) of the invention” correspond to the proteins, or group of proteins CM, EM, CXCR4, CXCR3, CD25, CCR5, LAG-3, CXCR5, CD5, Perforin, CD58, B7-H3, ICAM-1, CD38, CD11a, CD57, CD137, CD69, GzB, GzA, Naïve, CTLA-4, BTLA, HLA2, PD1, EMRA, CD127, CCR4 and GAL-3.

As used herein the term “CM” for “Central Memory” or CM marker denotes CD8⁺ T lymphocytes which express at their surfaces CD45RO, CCR7 and CD27 (CD45RO⁺, CCR7⁺ and CD27⁺) but not CD45RA (CD45RA⁻),

As used herein the term “EM” for “Effector Memory” or EM marker denotes CD8⁺ T lymphocytes which express at their surface CD45RO (CD45RO⁺) but not CD45RA, CCR7 and CD27 (CD45RA⁻, CCR7⁻ and CD27⁻).

As used herein the term “Naïve” or Naïve marker denotes CD8⁺ T lymphocytes which express at their surfaces CD45RA, CCR7 and CD27 (CD45RA⁺, CCR7⁺ and CD27⁺) but not CD45RO (CD45RO⁻).

As used herein the term “EMRA” for “Effector Memory RA” or EMRA marker denotes CD8⁺ T lymphocytes which express at their surface CD45RA (CD45RA⁺) but not CD45RO, CCR7 and CD27 (CD45RO⁻, CCR7⁻ and CD27⁻).

As used herein the term “CD45RA” refers to a member of the protein tyrosine phosphatase (PTP) family. PTPs are known to be signaling molecules that regulate a variety of cellular processes including cell growth, differentiation, mitotic cycle, and oncogenic transformation. The sequence of said protein can be found under the Uniprot accession number P08575. CD45RA denotes a specific isoform of CD45 corresponding to a splicing variant containing exon 4 but not exon 5 and 6. The CD45RA antibody used to detect this isoform can detect glycosylated membrane proteins carrying exon 4 when expressed alone or together with exon 5 and 6.

As used herein the term “CD45RO” refers to a member of the protein tyrosine phosphatase (PTP) family. PTPs are known to be signaling molecules that regulate a variety of cellular processes including cell growth, differentiation, mitotic cycle, and oncogenic transformation. The sequence of said protein can be found under the Uniprot accession number P08575. CD45RO denotes a specific isoform of CD45 corresponding to a splicing variant from which exon 4, 5 and 6 were removed. The CD45RO antibody used to detect this isoform can only detect glycosylated membrane proteins that do not express the amino acid portion corresponding to exon 4, 5 and 6.

As used herein the term "CXCR4", for "C-X-C chemokine receptor type 4" refers to Gα protein-coupled receptor in the CXC chemokine receptor family specific for stromal-derived-factor-1 (SDF-1 also called CXCL12), a molecule endowed with potent chemotactic activity for lymphocytes. CXCR4 is one of several chemokine receptors that HIV can use to infect CD4⁺ T cells. HIV isolates that use CXCR4 are traditionally known as T-cell tropic isolates. The sequence of said protein can be found under the Uniprot accession number P61073. According to the invention, the term "CXCR4 marker" denote CD8⁺ T lymphocytes which express at their surface CXCR4.

As used herein the term "CD27", refers to a receptor which is a member of the TNF-receptor superfamily. This receptor is required for generation and long-term maintenance of T cell immunity. It binds to ligand CD70, and plays a key role in regulating B-cell activation and immunoglobulin synthesis. The sequence of said protein can be found under the Uniprot accession number P26842. According to the invention, the term "CD27 marker" denote CD8⁺ T lymphocytes which express at their surface CD27.

As used herein the term "CXCR3", for "C-X-C chemokine receptor type 3" refers to a Gα protein-coupled receptor in the CXC chemokine receptor family. CXCR3 is expressed primarily on activated T lymphocytes and NK cells, and some epithelial cells and regulates leukocytes trafficking. CXCR3 and CCR5 are preferentially expressed on Th1 cells, whereas Th2 cells favor the expression of CCR3 and CCR4. The sequence of said protein can be found under the Uniprot accession number P49682. According to the invention, the term "CXCR3 marker" denote CD8⁺ T lymphocytes which express at their surface CXCR3.

As used herein the term "CCR7", for "C-C chemokine receptor type 7" refers to a G protein-coupled receptor in the CC chemokine receptor family. This receptor was identified as a gene induced by the Epstein-Barr virus (EBV), and is thought to be a mediator of EBV effects on B lymphocytes. This receptor is expressed in various lymphoid tissues, activates B and T lymphocytes and regulates their trafficking. CCR7 has been shown to stimulate dendritic cell maturation. The sequence of said protein can be found under the Uniprot accession number P32248. According to the invention, the term "CCR7 marker" denote CD8⁺ T lymphocytes which express at their surfaces CCR7.

As used herein the term "CD58", refers to a cell adhesion molecule expressed on Antigen Presenting Cells (APC), particularly macrophages but also on T lymphocytes. It binds to CD2 (LFA-2) on T cells and is important in strengthening the adhesion between the T cells and Professional Antigen Presenting Cells. The sequence of said protein can be found under the

Uniprot accession number P19256. According to the invention, the term “CD58 marker” denote CD8⁺ T lymphocytes which express at their surfaces CD58.

As used herein the term "CCR5", for “C-C chemokine receptor type 5” refers to a G protein-coupled receptor in the CC chemokine receptor family mostly expressed in the immune system cells where it acts as a receptor for chemokines. This is the process by which T cells are attracted to specific tissue and organ targets. The sequence of said protein can be found under the Uniprot accession number P51681. According to the invention, the term “CCR5 marker” denote CD8⁺ T lymphocytes which express at their surfaces CCR5.

As used herein the term "Perforin", refers to a pore forming cytolytic protein found in the granules of cytotoxic T lymphocytes (CTLs) and NK cells. Upon degranulation, perforin binds to the target cell's plasma membrane, and oligomerises in a Ca²⁺ dependent manner to form pores on the target cell. The sequence of said protein can be found under the Uniprot accession number P14222. According to the invention, the term “Perforin marker” denote CD8⁺ T lymphocytes which express at their surfaces the Perforin.

As used herein, the term “CXCR5”, for “C-X-C chemokine receptor type 5” refers to a Gα protein-coupled receptor in the CXC chemokine receptor family. According to the invention, the term “CXCR5 marker” denote CD8⁺ T lymphocytes which express at their surface CXCR5.

According to the invention, the term “CD127 marker” denote CD8⁺ T lymphocytes which express at their surfaces CD127, also known as interleukin-7 receptor-α.

According to the invention, the term “CD137 marker” denote CD8⁺ T lymphocytes which express at their surfaces CD137, also known as tumor necrosis factor receptor superfamily member 9.

According to the invention, the term “CD25 marker” denote CD8⁺ T lymphocytes which express at their surfaces CD25, also known as interleukin-2-receptor subunit alpha.

As used herein the term “B7-H3”, also known as CD276, refers to an immune checkpoint molecule, member of the B7 and CD28 families. B7-H3 participate in the regulation of T-cell mediated immune response. According to the invention, the term “B7-H3 marker” denote CD8⁺ T lymphocytes which express at their surfaces B7-H3

As used herein the term “CTLA-4”, for “Cytotoxic T-lymphocyte-Associated Protein 4” refers to a CD28 homologue expressed on T cells which regulates the amplitude of the early stages of T-cell activation. According to the invention, the term “CTLA-4 marker” denotes CD8⁺ T lymphocytes which express at their surfaces CTLA-4.

According to the invention, the term “CD38 marker” denote CD8⁺ T lymphocytes which express at their surfaces CD38, also known as cyclic ADP ribose hydrolase 1.

According to the invention, the term “Gal-3 marker” denotes CD8⁺ T lymphocytes which express at their surfaces galectin 3 (Gal-3).

5 According to the invention, the term “GzA marker” denotes CD8⁺ T lymphocytes which express at their surfaces granzyme A (GzA).

According to the invention, the term “GzB marker” denotes CD8⁺ T lymphocytes which express at their surfaces granzyme B (GzB).

10 According to the invention, the term “HLA-II marker” denotes CD8⁺ T lymphocytes which express at their surfaces human leukocyte antigen II (HLA-II).

As used herein the term “LAG-3”, for “Lymphocyte-activation gene 3”, refers to a cell surface protein expressed on activated T cells, NK cells, B cells and plasmacytoid dendritic cells. According to the invention, the term “LAG-3 marker” denotes CD8⁺ T lymphocytes which express at their surfaces LAG-3.

15 According to the invention, the term “CD57 marker” denotes CD8⁺ T lymphocytes which express at their surfaces CD57, also known as “Beta-1-3-Glucuronyltransferase 1”.

According to the invention, the term “CD69 marker” denotes CD8⁺ T lymphocytes which express at their surfaces CD69, also known as early T-cell activation antigen P60.

20 According to the invention, the term “CD11a marker” denotes CD8⁺ T lymphocytes which express at their surfaces CD11a, also known as integrin alpha-L.

According to the invention, the term “ICAM-1 marker” denotes CD8⁺ T lymphocytes which express at their surfaces ICAM-1, also known as Intercellular Adhesion Molecule 1.

According to the invention, the term “CD5 marker” denotes CD8⁺ T lymphocytes which express at their surfaces CD5, also known as lymphocyte antigen T1.

25 As used herein the term “PD1”, for Programmed cell death protein 1” refers to a cell surface receptor that belongs to the immunoglobulin family. The interaction between PD1 and its two ligands regulates the induction and maintenance of peripheral tolerance and protect tissues from autoimmune attack. According to the invention, the term “PD1 marker” denotes CD8⁺ T lymphocytes which express at their surfaces PD1.

30 As used herein the term “BTLA”, for “B- and T- lymphocyte attenuator” refers to a lymphocyte inhibitory receptor which inhibits lymphocytes during immune response. According to the invention, the term “BTLA marker” denotes CD8⁺ T lymphocytes which express at their surfaces BTLA.

As used herein the term "CCR4", for "C-C chemokine receptor type 4" refers to a G protein-coupled receptor in the CC chemokine receptor family. According to the invention, the term "CCR4 marker" denote CD8⁺ T lymphocytes which express at their surfaces CCR4.

5 According to the invention, an optional step of statistical classifier or any machine learning algorithm can be added at the end of the methods of the invention. For example a step of classification of the patients with the k-neighbours algorithm (see for example Fix, E. et al 1951), principal component analysis (see for example Pearson K. 1901 "On Lines and Planes of Closest Fit to Systems of Points in Space". *Philosophical Magazine*. **2** (11): 559–572), support vector machine (see for example Cortes C. et al 1995. "Support-vector networks". *Machine Learning*. **20** (3): 273–297), neural networks (see for example Rosenblatt, F. 1958. "The Perceptron: A Probabilistic Model For Information Storage And Organization In The Brain". *Psychological Review*. **65** (6): 386–408) or any classifier, can be added.

15 According to the invention, the expression level of the markers listed above may also be measured by measuring the protein expression level of these markers and can be performed by a variety of techniques well known in the art.

Typically protein expression level may be measured for example by capillary electrophoresis-mass spectroscopy technique (CE-MS), flow cytometry, mass cytometry or
20 ELISA performed on the sample.

In the present application, the "level of protein" or the "protein level expression" means the quantity or concentration of said protein. In particular embodiment, the protein is expressed at the cell surface for markers whose function is linked to their correct plasma membrane expression or total expression for markers whose function is not limited to membrane
25 expression. In still another embodiment, the "level of protein" means the quantitative measurement of the proteins expression relative to a negative control.

Such methods comprise contacting a sample with a binding partner capable of selectively interacting with proteins present in the sample. The binding partner is generally an antibody that may be polyclonal or monoclonal, preferably monoclonal.

30 The presence of the protein can be detected using standard electrophoretic and immunodiagnostic techniques, including immunoassays such as competition, direct reaction, or sandwich type assays. Such assays include, but are not limited to, Western blots; agglutination tests; enzyme-labeled and mediated immunoassays, such as ELISAs; biotin/avidin type assays;

radioimmunoassays; immunoelectrophoresis; immunoprecipitation, capillary electrophoresis-mass spectroscopy technique (CE-MS).etc. The reactions generally include revealing labels such as fluorescent, chemiluminescent, radioactive, enzymatic labels or dye molecules, or other methods for detecting the formation of a complex between the antigen and the antibody or antibodies reacted therewith.

The aforementioned assays generally involve separation of unbound protein in a liquid phase from a solid phase support to which antigen-antibody complexes are bound. Solid supports which can be used in the practice of the invention include substrates such as nitrocellulose (e. g., in membrane or microtiter well form); polyvinylchloride (e. g., sheets or microtiter wells); polystyrene latex (e.g., beads or microtiter plates); polyvinylidene fluoride; diazotized paper; nylon membranes; activated beads, magnetically responsive beads, and the like.

More particularly, an ELISA method can be used, wherein the wells of a microtiter plate are coated with a set of antibodies against the proteins to be tested. A sample containing or suspected of containing the marker protein is then added to the coated wells. After a period of incubation sufficient to allow the formation of antibody-antigen complexes, the plate(s) can be washed to remove unbound moieties and a detectably labeled secondary binding molecule is added. The secondary binding molecule is allowed to react with any captured sample marker protein, the plate is washed and the presence of the secondary binding molecule is detected using methods well known in the art.

Particularly, a mass spectrometry-based quantification methods may be used. Mass spectrometry-based quantification methods may be performed using either labelled or unlabelled approaches [DeSouza and Siu, 2012]. Mass spectrometry-based quantification methods may be performed using chemical labeling, metabolic labeling or proteolytic labeling. Mass spectrometry-based quantification methods may be performed using mass spectrometry label free quantification, a quantification based on extracted ion chromatogram (EIC) and then profile alignment to determine differential level of polypeptides.

Particularly, a mass spectrometry-based quantification method particularly useful can be the use of targeted mass spectrometry methods as selected reaction monitoring (SRM), multiple reaction monitoring (MRM), parallel reaction monitoring (PRM), data independent acquisition (DIA) and sequential window acquisition of all theoretical mass spectra (SWATH) [Moving target Zeliadt N 2014 The Scientist; Liebler Zimmerman Biochemistry 2013 targeted quantitation of proteins by mass spectrometry; Gallien Domon 2015 Detection and quantification of proteins in clinical samples using high resolution mass spectrometry. Methods

v81 p15-23 ; Sajic, Liu, Aebersold, 2015 Using data-independent, high-resolution mass spectrometry in protein biomarker research: perspectives and clinical applications. Proteomics Clin Appl v9 p 307-21].

Particularly, the mass spectrometry-based quantification method can be the mass
5 cytometry also known as cytometry by time of flight (CYTOF) (Bandura DR, Analytical chemistry, 2009).

Particularly, the mass spectrometry-based quantification is used to do peptide and/or protein profiling can be use with matrix-assisted laser desorption/ionisation time of flight (MALDI-TOF), surface-enhanced laser desorption/ionization time of flight (SELDI-TOF;
10 CLINPROT) and MALDI Biotyper apparatus [Solassol, Jacot, Lhermitte, Bouille, Maudelonde, Mangé 2006 Clinical proteomics and mass spectrometry profiling for cancer detection. Journal: Expert Review of Proteomics V3, I3, p311-320 ; FDA K130831].

Methods of the invention may comprise a step consisting of comparing the proteins and fragments concentration in circulating cells with a control value. As used herein, "concentration
15 of protein" refers to an amount or a concentration of a transcription product, for instance the proteins of the invention. Typically, a level of a protein can be expressed as nanograms per microgram of tissue or nanograms per milliliter of a culture medium, for example. Alternatively, relative units can be employed to describe a concentration. In a particular embodiment, "concentration of proteins" may refer to fragments of the proteins of the invention.

20 In a particular embodiment, the detection of the level of the markers of the invention can be performed by flow cytometry. When this method is used, the method consists of determining the percentages of CD8+ lymphocytes expressing the markers of the invention.

Expression levels may be expressed as absolute expression level or normalized expression level of the marker.

25 Optimal threshold values can be determined as explained below and as explained in the examples part and by statistical methods well known in the state of art.

According to the invention, the "background" as used in the context of the invention may be predetermined by carrying out a method comprising the steps of:

a) measuring the expression level of a given marker in a sample of interest using a
30 specific detection reagent at the cellular level;

b) make the same measurement experiment on the same sample but without the specific detection reagent to obtain the "background" value of the measurement at the cellular level;

c) calculate among the individual cell measurements obtained from the sample for said protein, the percentage of cells that have an expression level that is above the “background” value for which the expression level is said to be “positive” and the percentage of cells that have an expression level that is below the “background” value for which the expression level is said to be “negative”.

A second aspect of the invention relates to a chemotherapeutic compound for use in the treatment of a cancer in a patient with need for therapy as described above.

In another embodiment the invention relates to a chemotherapeutic compound for use in the treatment of a cancer in a patient with a bad prognosis as described above.

In another embodiment, the invention relates to a method for treating a cancer in a patient with need for therapy as described above comprising administering to said subject in need of a chemotherapeutic compound.

In another embodiment, the invention relates to a method for treating a cancer in a patient with a bad prognosis as described above comprising administering to said subject with need of a chemotherapeutic compound.

In a particular embodiment, the cancer is a CLL.

According to the invention, chemotherapeutic compounds may be selected in the group consisting in: fludarabine, gemcitabine, capecitabine, methotrexate, taxol, taxotere, mercaptopurine, thioguanine, hydroxyurea, cytarabine, cyclophosphamide, ifosfamide, nitrosoureas, platinum complexes such as cisplatin, carboplatin and oxaliplatin, mitomycin, dacarbazine, procarbazine, etoposide, teniposide, campathecins, bleomycin, doxorubicin, idarubicin, daunorubicin, dactinomycin, plicamycin, mitoxantrone, L-asparaginase, epimibicm, 5-fluorouracil, taxanes such as docetaxel and paclitaxel, leucovorin, levamisole, irinotecan, estramustine, etoposide, nitrogen mustards, BCNU, nitrosoureas such as carmustine and lomustine, vinca alkaloids such as vinblastine, vincristine and vinorelbine, imatinib mesylate, hexamethylnelamine, topotecan, kinase inhibitors, phosphatase inhibitors, ATPase inhibitors, tyrosine kinases, protease inhibitors, inhibitors herbimycin A, genistein, erastatin, temozolomide and lavendustin A. In one embodiment, additional anticancer agents may be selected from, but are not limited to, one or a combination of the following class of agents: alkylating agents, plant alkaloids, DNA topoisomerase inhibitors, anti-folates, pyrimidine analogs, purine analogs, DNA antimetabolites, taxanes, podophyllotoxin, hormonal therapies, retinoids, photosensitizers or photodynamic therapies, angiogenesis inhibitors, antimitotic agents,

isoprenylation inhibitors, cell cycle inhibitors, actinomycins, bleomycins, anthracyclines, MDR inhibitors and Ca²⁺ ATPase inhibitors.

In a particular embodiment, chemotherapeutic compounds used to treat CLL can be chlorambucil (alkylating agent) alone or in combination with anti-CD20, fludarabine, pentostatin, and cladribine (2-CdA), CHOP (cyclophosphamide, doxorubicin, vincristine, prednisone), CAP (cyclophosphamide, doxorubicin, prednisone), Bendamustine (4-[5-[bis(2-chloroethyl)amino]-1-methylbenzimidazol-2-yl]butanoic acid, monoclonal antibodies : anti-CD20 (Rituximab, Ofatumumab, Obinutuzumab) or anti-CD52 (Alemtuzumab), molecular agents targeting B cell receptor and chemokine receptor signaling like idelalisib (PI3K-delta isoform specific inhibitor), ibrutinib (Burton tyrosine Kinase inhibitor), acalabrutinib (ACP-196, more selective, irreversible Burton tyrosine Kinase inhibitor), BCL-2 inhibitors like Venetoclax (ABT-199, BH3 mimetic drug) or immunomodulatory drugs like lenalidomide (thalidomide analog).

As used herein, the term "treatment" or "treat" refer to both prophylactic or preventive treatment as well as curative or disease modifying treatment, including treatment of subjects at risk of contracting the disease or suspected to have contracted the disease as well as subjects who are ill or have been diagnosed as suffering from a disease or medical condition, and includes suppression of clinical relapse. The treatment may be administered to a subject having a medical disorder or who ultimately may acquire the disorder, in order to prevent, cure, delay the onset of, reduce the severity of, or ameliorate one or more symptoms of a disorder or recurring disorder, or in order to prolong the survival of a subject beyond that expected in the absence of such treatment. By "therapeutic regimen" is meant the pattern of treatment of an illness, e.g., the pattern of dosing used during therapy. A therapeutic regimen may include an induction regimen and a maintenance regimen. The phrase "induction regimen" or "induction period" refers to a therapeutic regimen (or the portion of a therapeutic regimen) that is used for the initial treatment of a disease. The general goal of an induction regimen is to provide a high level of drug to a subject during the initial period of a treatment regimen. An induction regimen may employ (in part or in whole) a "loading regimen", which may include administering a greater dose of the drug than a physician would employ during a maintenance regimen, administering a drug more frequently than a physician would administer the drug during a maintenance regimen, or both. The phrase "maintenance regimen" or "maintenance period" refers to a therapeutic regimen (or the portion of a therapeutic regimen) that is used for the maintenance of a subject during treatment of an illness, e.g., to keep the subject in remission for long periods of time (months or years). A maintenance regimen may employ continuous

therapy (e.g., administering a drug at a regular intervals, e.g., weekly, monthly, yearly, etc.) or intermittent therapy (e.g., interrupted treatment, intermittent treatment, treatment at relapse, or treatment upon achievement of a particular predetermined criteria [e.g., disease manifestation, etc.]). In a particular embodiment, the patient has already a cancer (for example a CLL).

5

A third aspect of the invention relates to a therapeutic composition comprising a chemotherapeutic compound according to the invention for use in the treatment of CLL in a patient with need for therapy as described above.

10 In another embodiment, the invention relates to a therapeutic composition comprising a chemotherapeutic compound according to the invention for use in the treatment of CLL in patient with a bad prognosis as described above.

Any therapeutic agent of the invention may be combined with pharmaceutically acceptable excipients, and optionally sustained-release matrices, such as biodegradable polymers, to form therapeutic compositions.

15 "Pharmaceutically" or "pharmaceutically acceptable" refers to molecular entities and compositions that do not produce an adverse, allergic or other untoward reaction when administered to a mammal, especially a human, as appropriate. A pharmaceutically acceptable carrier or excipient refers to a non-toxic solid, semi-solid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type.

20 The form of the pharmaceutical compositions, the route of administration, the dosage and the regimen naturally depend upon the condition to be treated, the severity of the illness, the age, weight, and sex of the patient, etc.

The pharmaceutical compositions of the invention can be formulated for a topical, oral, intranasal, parenteral, intraocular, intravenous, intramuscular or subcutaneous administration
25 and the like.

Particularly, the pharmaceutical compositions contain vehicles which are pharmaceutically acceptable for a formulation capable of being injected. These may be in particular isotonic, sterile, saline solutions (monosodium or disodium phosphate, sodium, potassium, calcium or magnesium chloride and the like or mixtures of such salts), or dry,
30 especially freeze-dried compositions which upon addition, depending on the case, of sterilized water or physiological saline, permit the constitution of injectable solutions.

The doses used for the administration can be adapted as a function of various parameters, and in particular as a function of the mode of administration used, of the relevant pathology, or alternatively of the desired duration of treatment.

In addition, other pharmaceutically acceptable forms include, e.g. tablets or other solids for oral administration; time release capsules; and any other form currently can be used.

The invention will be further illustrated by the following figures and examples. However, these examples and figures should not be interpreted in any way as limiting the scope of the present invention.

FIGURES:

Figure 1: Supervised analysis of phenotypic expression data defines the markers that are important in detecting patients who will be treated.

A- 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).

B- Parameters correlating with “need for therapy” as ranked by Adaboost analysis. The parameters are ranked according to normalized Gini index of their importance (Adaboost importance).

C- Parameters correlating with need for therapy as ranked by Decision tree analysis. The parameters are ranked according to normalized Gini index of their importance (Decision tree importance)

Figure 2: 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- Parameters correlating with need for therapy as ranked by random forest analysis. The parameters are ranked according to the Gini index of their importance (Random Forest importance). Dotted lines highlight the six selected parameters for CD8⁺ T cell compartment signature chosen of the basis of the analysis shown in panels B and C.

B- 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 three fold 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).

C- Accuracies of logistic regression models predictions of the frozen sample validation cohort. Different logistic regressions were generated using 2,3,4,5 or 6 markers as in Figure 2B. For each model, the mean accuracy of the three fold 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).

	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 ⁺

5 **Table 1:** List of markers and parameters extracted from flow cytometry data and used in the study.

EXAMPLE:**Material & 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 2008-129) 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). 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 1). 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) for all markers of interest.

Whole blood samples were directly mixed with fluorochrome-coupled antibodies for at least one hour at 4°C. Red blood cells were then lysed and cell samples were washed twice with FACS buffer (PBS, 1% Fetal calf serum, 1% Human serum, 0.01% Na Azide) before fixation in 2% Paraformaldehyde. Samples were then washed twice before permeabilisation 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 FORTESSA 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.

5 Fresh versus frozen sample comparison

For 9 CLL patients, the freshly obtained blood samples were split in two : one half was analyzed right away using our standard protocol for fresh blood samples and the other half was processed for freezing : peripheral Blood Mononuclear Cells (PBMC) were prepared and frozen in Fetal Calf serum with 10% DMSO. Frozen samples were then thawed and stained according
10 to our staining protocol. Fresh and frozen samples were acquired by cytometry and data were processed identically and compared.

Cytotoxic assays

CD8⁺ T cells were isolated using Whole Blood CD8 microbeads (Miltenyi Biotec SAS) from 1-2 ml of fresh patient or healthy donor blood. Normal B cells or B-CLL cells were
15 isolated using RosetteSep human B cell enrichment cocktail (STEMCELL technologies). EBV transformed B cells (JY) were used as conventional target cells. Target cells were either unpulsed or pulsed with a bacterial superantigen cocktail (100 ng/mL TSST-1, SEC 1, SEB, SEA, SEE) for 1 h 30 at 37 °C in RPMI/5% FCS/Hepes and washed. CD8⁺ T cells were conjugated with target cells at different Effector: Target cell ratios for 18 h. To distinguish
20 CD8⁺ T cells from target cells in the analysis, CD8⁺ T cells were labeled before conjugation with 1µM CMFDA (Molecular Probes) for 20 min at 37 °C. Immediately before FACS analysis, 7-Amino-actinomycin D (7-AAD) was added to each sample to stain dead cells.

Cytokine production

CLL patients and healthy donors PBMC were stimulated on immobilized CD3/CD28
25 antibodies (10µg/ml) for 18h. During the last 4 hours, PMA (100ng/ml), ionomycin (1µg/ml) and Golgi STOP (BD biosciences) were added to the culture medium. Cells were then washed and stained for CD3, CD19 and CD8 surface markers in FACS buffer (PBS, 1% Fetal calf serum, 1% Human serum, 0.01% Na Azide) before fixation in 2% Paraformaldehyde. Samples were then washed twice before permeabilisation in FACS buffer with 0.1% Saponin. Cell
30 suspensions were then stained with antibodies directed intracellular cytokine proteins (FITC mouse anti-human IFN γ antibody, BV421 mouse anti-human TNF α antibody, PECy7 mouse anti-human MIP1 β antibody and PE rat anti-human IL2 antibody all from BD biosciences) for 30 min at room temperature before analysis on a BD FORTESSA cytometer. The percentage of positive cells was calculated after applying a threshold corresponding to unstimulated cells.

The fold increase MFI (Median Fluorescence Intensity) was calculated by dividing the MFI of CD8⁺cytokine⁺ cells by the MFI of the total CD8 population for the considered cytokine.

Statistics, data and code availability

Detailed statistical methods used throughout the study are available as supplementary
5 information. We used R software for most statistical analyses and python software for supervised learning.

Treatment of flow cytometry raw data

The first step of data processing was to define which numerical values would be
10 extracted from fluorescence distributions obtained by each antibody used in multiparametric analysis. We compared two frequently used methods for FACS data quantification: i-) differences in the mean fluorescence intensity between the marker of interest and a normalized control (delta mean) ii-) percentage of cells within the distribution of fluorescence that is above a threshold defined by a control (isotype control).

We used the same procedure to analyze all the data extracted from cytometry. We started
15 by removing the 0.05% of larger values in each cytometry csv file in order to remove artificial high values. Then, for each marker we apply a Kolmogorov-Smirnov test to the isotype data and the marked data to test whether the distribution of the two samples are statistically different. If the p-value was greater than 5%, then the differences were decided to be not significant and we set the delta-mean and the percentage of positive cells to 0. Otherwise, we computed the
20 delta-mean as difference of the mean fluorescence in the marked sample and of the isotype sample. We then computed a threshold corresponding to the 95th percentile of the isotype data and defined the percentage of positive cells as the percentage of marked fluorescence above this threshold.

We observed that the results given by both methods were highly correlated for most
25 markers (Spearman correlation coefficient > 0.5) but for a few of them the correlation was less significant. These markers corresponded to situations where the fluorescence distribution of the staining (as detected in flow cytometry plots) were either bimodal or far from Gaussianity. In those cases, the information contained in fluorescence distribution is lost by extracting the delta mean fluorescence intensity. We thus chose to extract the percentage of positive cells above the
30 isotype control among the CD8⁺ T cell population (defined as CD3⁺ CD19⁻ CD8⁺ cells) for all markers of interest

Analysis of sample's origin

We compared the difference between percentage of positive cells on the fresh and the frozen sample for each marker using a paired two sample t-test assessing the significant

difference of the means with 0. Indeed, when the difference between percentage of positive cells of the fresh and the frozen sample for a given marker is equal to zero, the sample is not affected by PBMC/Freezing steps.

Univariate analysis of healthy vs CLL cohorts

5 For each marker we used a Mann-Whitney test (also known as Wilcoxon rank-sum test) to test whether the mean percentage of positive cells in healthy and CLL patient is significantly different. Because of the high number of statistical tests, we used the Benjamini-Hochberg correction on the previously obtained p-values.

Visualization of raw data and their correlation

10 We computed the Spearman correlation matrix of the markers fluorescence for the healthy cohort and the CLL cohort and represented it using “corrplot” package. For the heatmap, we used a specific color scale, enhancing differences near 0 and 1.

Unsupervised analysis

The unsupervised analysis of the multivariate data was conducted on the scaled data in
15 order to avoid parasite effects due to the presence of percentage values (between 0 and 1) and larger values as age. The clustering was done using hierarchical clustering on Euclidean distance of scaled data using Ward method. For visualization, we used PCA (Principal component Analysis) to reduce the parameters dimensions to a 2-D space. We used “FactoMineR” package for PCA and graphics on the scaled data. We computed the accuracy
20 of clustering using confusion matrices when the clustering output was known (CLL/Healthy, need for therapy). We also assessed the probability that the calculated accuracy was superior to no-information rate. The no information rate corresponds to the larger class (cluster) in the data set. The p-value obtained thus reflects the probability to be wrong when accepting the hypothesis that the tested classification do better than a classification that would always predict
25 the most common class.

Supervised analysis

We treat supervised analysis as a binary classification task for which algorithms are trained, for a given CLL patient described by its vector of marker expression levels, to predict whether the time to treatment will be greater than 6 months (label 1) or not (label 0). We trained
30 three different classifiers based on decision trees to perform the prediction: the classical Decision Tree classifier (CART algorithm), as well as two decision tree ensemble learning methods: Adaboost and Random Forest. Each algorithm has been evaluated in a cross-validation scheme by performing 1000 iterations of the method. For each iteration we selected a random subset of 75% of the CLL patients for supervised training of the classifier, and used

the 25% remaining CLL patients to compute its accuracy. Gini indices were extracted for each classifier to compute a normalized marker expression hierarchy in the decision rule. For analysis, we took average values of accuracy and hierarchies over the 1000 iterations.

Scoring strategy

5 To provide a score (defined as S1 score), we selected the most important markers from the Random forest algorithm. We computed the Spearman correlation matrix of the selected marker to verify the levels of correlation since absence of correlation is a prerequisite for logistic regression (see R code). We first aim at deciding how many of these markers are really needed to obtain a good prediction of the need of treatment within 6 months. To this aim we
10 conducted a 3-fold cross validation scheme to compare the accuracies of different models (using the 2, 3, 4, 5, 6 most important markers). We created three groups of patient ensuring that each group contains treated and non-treated patient, and used sequentially two groups as learning cohort and a group as testing cohort. On each learning cohort, we computed the S1 score using a logistic regression (R glm function) and chose an optimal threshold with ROCR package. We
15 then applied the learned S1 score to the testing cohort and computed accuracy and F-measure. We performed these step on 10 different groups of patients and calculated the mean accuracy and mean F-measure.

Results:

20 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) (data not shown). We focused our study on whole blood to preserve tumor microenvironment and to ensure that CD8⁺ T cells
25 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 (data not shown, (10-11)).

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
30 significantly different expression in CLL patients and healthy donors (data not shown). 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 (data not shown). 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 (data not shown): 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 (data not shown). 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%) (data not shown) 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) (data not shown). While adhesion molecule and lytic molecule expression correlated positively with dimension 1, chemokine receptor and activation/differentiation molecule expression negatively correlated with dimension 1 (data not shown). We also observed that, 4 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 4 markers in a multi-step gating strategy (Table 1) 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 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 (data not shown). Moreover, the cytotoxicity of CD8⁺ T cells towards conventional targets or autologous tumor B cells was reduced (data not shown) despite high levels of lytic molecules expression (data not shown). 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 years for CLL patients and 50-66 years 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 (data not shown).

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⁺ (data not shown).

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 towards 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 phenotypical changes among CD8⁺ T cells. 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 (data not shown). Significant clustering of patients according to Binet stage or IGVH mutational status were also not observed (data not shown). 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 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 criteria. We generated a 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 1A). 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 towards therapy from the one who don't (data not shown). Interestingly, the expression profile of the differentiation markers associated with evolution towards therapy (increased representation of EM cells and decreased representation of CM cells among CD8⁺ T cells) was also observed, to a lesser extent, in patients with more advanced Binet stage (B and C) and patients with unmutated IGVH genes (data not shown).

To confirm the existence of this specific phenotypic signature of CD8⁺ T cells in patients evolving towards therapy, we tested 2 additional learning algorithms from which we could get a feature hierarchy after learning. With Adaboost algorithm (Figure 1B), we found again differentiation status subsets (EM, CM) among the 5 most important markers to predict the need for therapy. Those 2 subsets were also the most important markers in the Decision tree learning (Figure 1C).

In conclusion, our supervised analysis highlights a phenotypic signature correlating with evolution towards 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 six 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 (defined as S1 score) 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. We used a cross validation scheme where we split the cohort in 3 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 S1 score on remaining patients and assess the validity of the model. Then, based on the learning data, we adjusted an optimal threshold with ROCR package to optimize the number of “false positive” and “true negative”. The individuals whose S1 score was above the cutoff, were predicted as “treated” and

“untreated” otherwise. 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 2B). We observed that increasing 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 towards therapy.

As a control, we calculated the S1 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) (data not shown).

To test whether the observed signature of CD8⁺ T cells from CLL patients that will evolve towards 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 don’t (data not shown).

In conclusion, these results identify a phenotypic signature of CD8⁺ T cells in CLL patients that evolve towards 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 phenotype 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 (data not shown).

As mentioned before, we observed that the expression of several phenotypic markers is altered by cell isolation/freezing procedures (data not shown). Nevertheless, the use of frozen samples is convenient for several reasons: a) frozen samples are much more easily available and sharable than fresh samples; b) they can be processed in a more automatized 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. 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 2C). We observed that increasing the number of markers taken into account in the logistic regression to 3 markers (CM, EM, CXCR4) (but not above) did improve the accuracy or the F-measure.

These observations confirm that a signature of 3 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 4 groups of patients based on the time elapsed between diagnosis and phenotyping (0-2 years, 2-5 years, 5-10 years and >10 years) 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 (data not shown). 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 is 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.

REFERENCES:

Throughout this application, various references describe the state of the art to which this invention pertains. The disclosures of these references are hereby incorporated by reference into the present disclosure.

1. Fridman, W. H., Pages, F., Sautes-Fridman, C. & Galon, J. The immune contexture in human tumours: impact on clinical outcome. Nat Rev Cancer 12, 298-306, doi:10.1038/nrc3245 (2012).

2. Kalos, M. & June, C. H. Adoptive T cell transfer for cancer immunotherapy in the era of synthetic biology. *Immunity* 39, 49-60, doi:10.1016/j.immuni.2013.07.002 (2013).

3. Chen, D. S. & Mellman, I. Oncology meets immunology: the cancer-immunity cycle. *Immunity* 39, 1-10, doi:10.1016/j.immuni.2013.07.012 (2013).

5 4. Dunn, G. P., Bruce, A. T., Ikeda, H., Old, L. J. & Schreiber, R. D. Cancer immunoediting: from immunosurveillance to tumor escape. *Nat Immunol* 3, 991-998, doi:10.1038/ni1102-991 (2002).

5. Scarfo, L., Ferreri, A. J. & Ghia, P. Chronic lymphocytic leukaemia. *Crit Rev Oncol Hematol* 104, 169-182, doi:10.1016/j.critrevonc.2016.06.003 (2016).

10 6. Mellstedt, H. & Choudhury, A. T and B cells in B-chronic lymphocytic leukaemia: Faust, Mephistopheles and the pact with the Devil. *Cancer Immunol Immunother* 55, 210-220, doi:10.1007/s00262-005-0675-4 (2006).

7. Ramsay A. G., Johnson A. J., Lee A. M., Gorgun G., Le Dieu R, Blum W., et al. Chronic lymphocytic leukemia T cells show impaired immunological synapse formation that
15 can be reversed with an immunomodulating drug. *J Clin Invest* 118:2427-37, doi:10.1172/JCI35017 (2008).

8. Kabanova A., Sanseviero F., Candi V., Gamberucci A., Gozzetti A., Campoccia G., et al. Human Cytotoxic T Lymphocytes Form Dysfunctional Immune Synapses with B Cells Characterized by Non-Polarized Lytic Granule Release. *Cell Rep*; 15:9-18, doi:
20 10.1016/j.celrep.2016.02.084 (2016).

9. Riches J. C., Ramsay A. G., Gribben J. G. T-cell function in chronic lymphocytic leukaemia. *Semin Cancer Biol*; 20:431-8, doi: 10.1016/j.semcancer.2010.09.006 (2010)

10. Romeu M. A., Mestre M., Gonzalez L., Valls A., Verdaguer J., Corominas M., et al. Lymphocyte immunophenotyping by flow cytometry in normal adults. Comparison of fresh
25 whole blood lysis technique, Ficoll-Paque separation and cryopreservation. *J Immunol Methods*.154:7-10 (1992).

11. Ng A. A., Lee B. T., Teo T. S., Poidinger M. & Connolly J. E. Optimal cellular preservation for high dimensional flow cytometric analysis of multicentre trials. *J Immunol Methods*; 385:79-89, doi: 10.1016/j.jim.2012.08.010 (2012).

30 12. 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*; 18:678-87, doi: 10.1158/1078-0432.CCR-11-2630 (2012).

13. Nikolich-Žugich J. The twilight of immunity: emerging concepts in aging of the immune system. *Nature Immunology*; 19:10-9, doi: 10.1038/s41590-017-0006-x (2018).

14. Mackus W. J., Frakking F.N., Grummels A., Gamadia L. E., De Bree G. J., Hamann D., et al. Expansion of CMV-specific CD8+CD45RA+CD27- T cells in B-cell chronic
5 lymphocytic leukemia. *Blood*; 102:1057-63 (2003).

10

15

20

CLAIMS:

1. A method for predicting the need for therapy of a patient suffering from a cancer comprising: i) determining, in a sample obtained from the patient, the expression level of markers expressed by the CD8⁺ T lymphocytes; ii) calculating the CD8⁺ T cell compartment S1 score of the patient, using the expression levels from step i), according to the logistic regression defined from a cancer learning cohort; iii) comparing the S1 score obtained for a given patient, to the optimal threshold value defined from the cancer learning cohort ; iv) providing a prognosis of evolution toward therapy within 4 to 8 months when the S1 score determined at step ii) is higher than the optimal threshold value, or providing a prognosis of not needing therapy within 4 to 8 months when the S1 score is lower than the optimal threshold.
2. A method according to claim 1, wherein the cancer is the chronic lymphocytic leukemia (CLL).
3. A method according to claims 1 or 2 wherein the markers expressed by CD8⁺ T cells whose expression level is determined in step i) are Central Memory (CM) and Effector Memory (EM).
4. A method according to claims 1 or 2 wherein the markers expressed by CD8⁺ T cells whose expression level is determined in step i) are CM and EM and at least one marker selected among CXCR4, CXCR3, CD25, CCR5, LAG-3, CXCR5, CD5, Perforin, CD58, B7-H3, ICAM-1, CD38, CD11a, CD57, CD137, CD69, GzB, GzA, Naïve, CTLA-4, BTLA, HLA2, PD1, EMRA, CD127, CCR4, GAL-3.
5. A method according to claims 1 or 2 wherein the markers expressed by CD8⁺ T cells whose expression level is determined in step i) are CM, EM and CXCR4.
6. A method according to claims 1 or 2 wherein the markers expressed by CD8⁺ T cells whose expression level is determined in step i) are CM, EM, CXCR4 and at least one marker selected among CXCR3, CD25, CCR5, LAG-3, CXCR5, CD5, Perforin, CD58, B7-H3, ICAM-1, CD38, CD11a, CD57, CD137, CD69, GzB, GzA, Naïve, CTLA-4, BTLA, HLA2, PD1, EMRA, CD127, CCR4, GAL-3.

7. A method according to claims 1 or 2 wherein the markers expressed by CD8+ T cells whose expression level is determined in step i) are CM, EM, CXCR4, CXCR3, CD25 and CCR5.
8. A method according to any one claims 1 to 7 wherein the prediction of the need for therapy is done within 6 months.
9. A method for treating a CLL in a patient comprising: i) predicting the need for therapy of the patient according to any one claims 3 to 7 and ii) administering a chemotherapeutic compound, if the patient has been considered with need for therapy in step i).

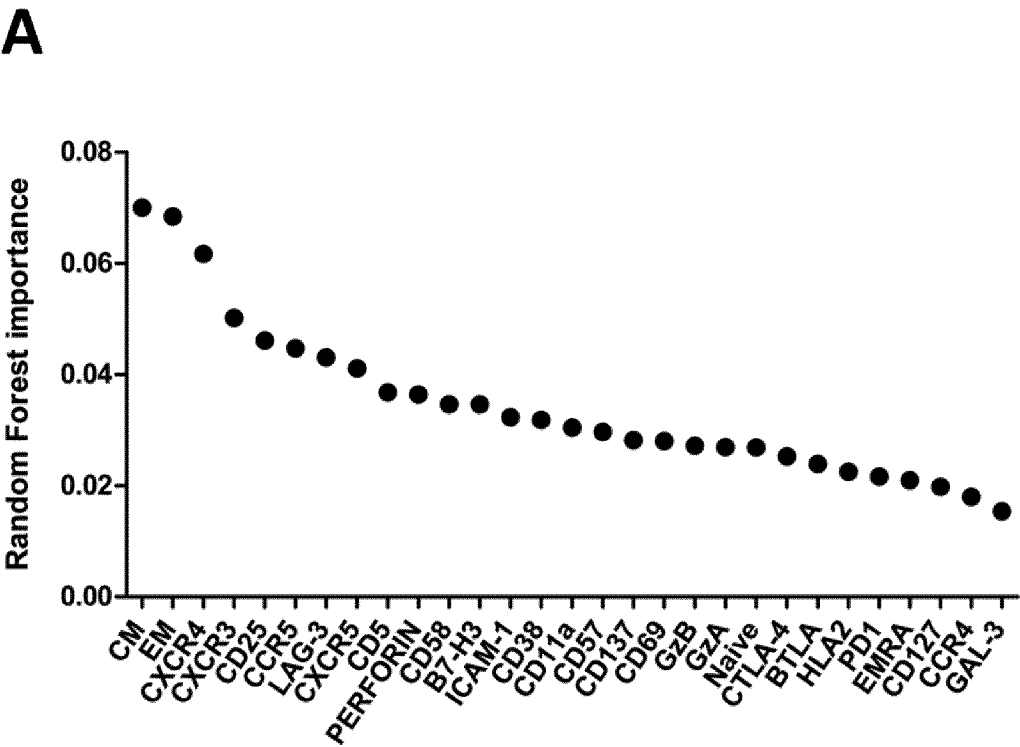


Figure 1A

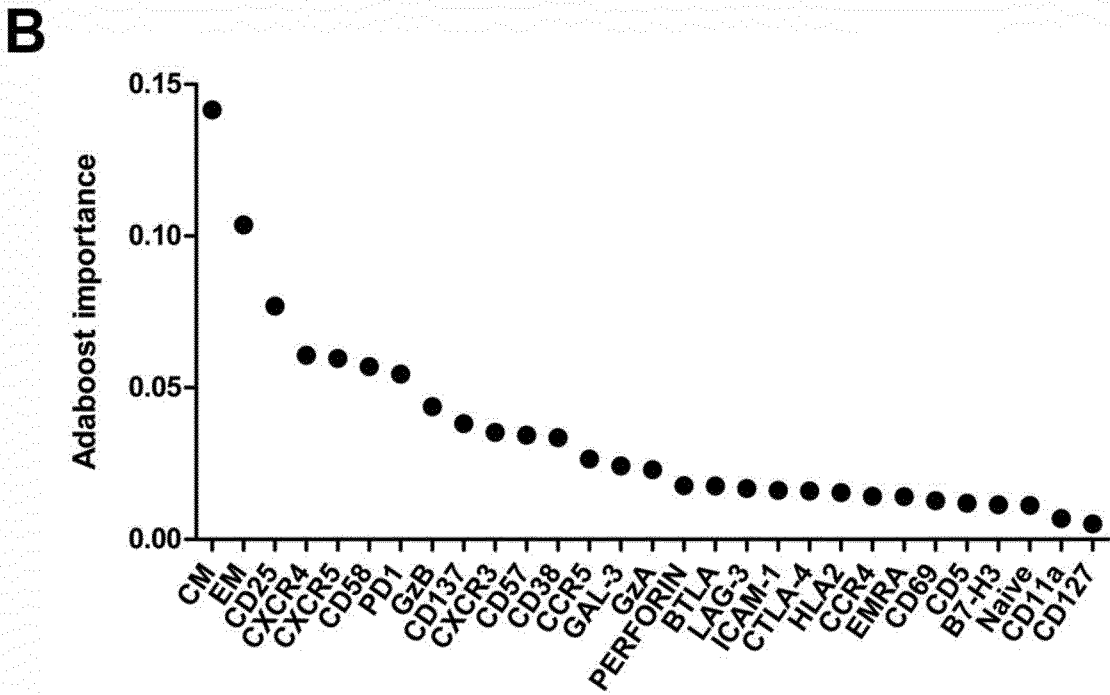


Figure 1B

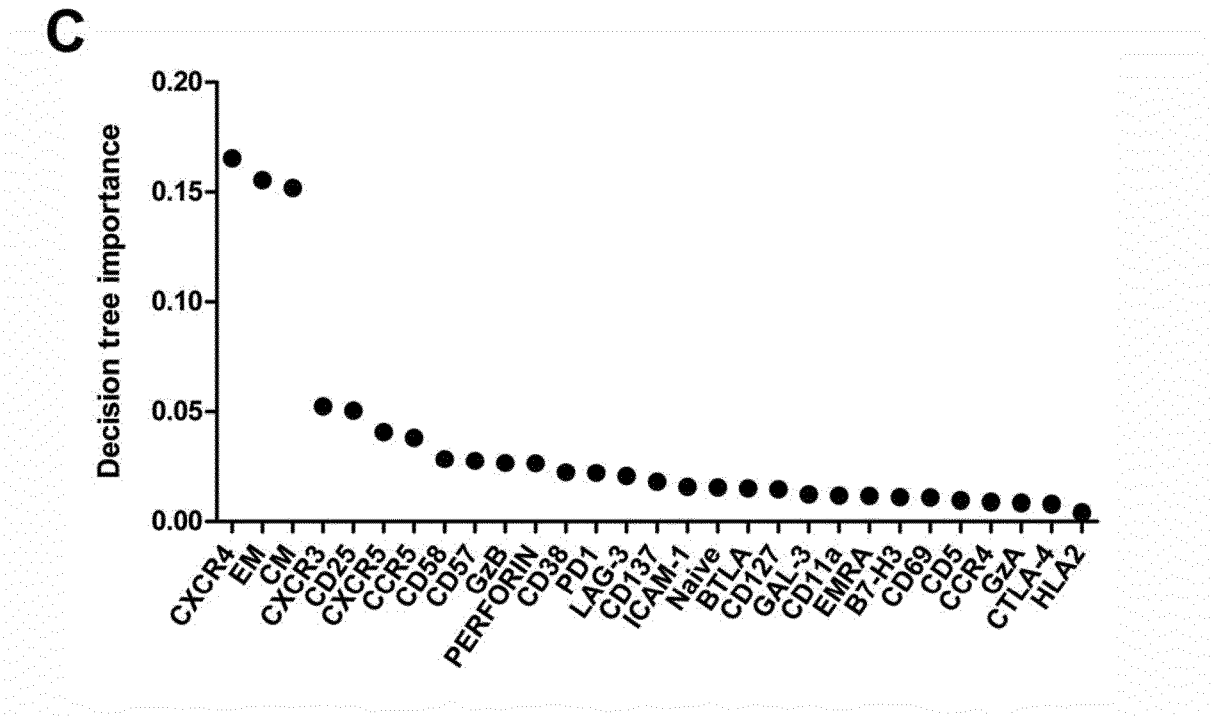


Figure 1C

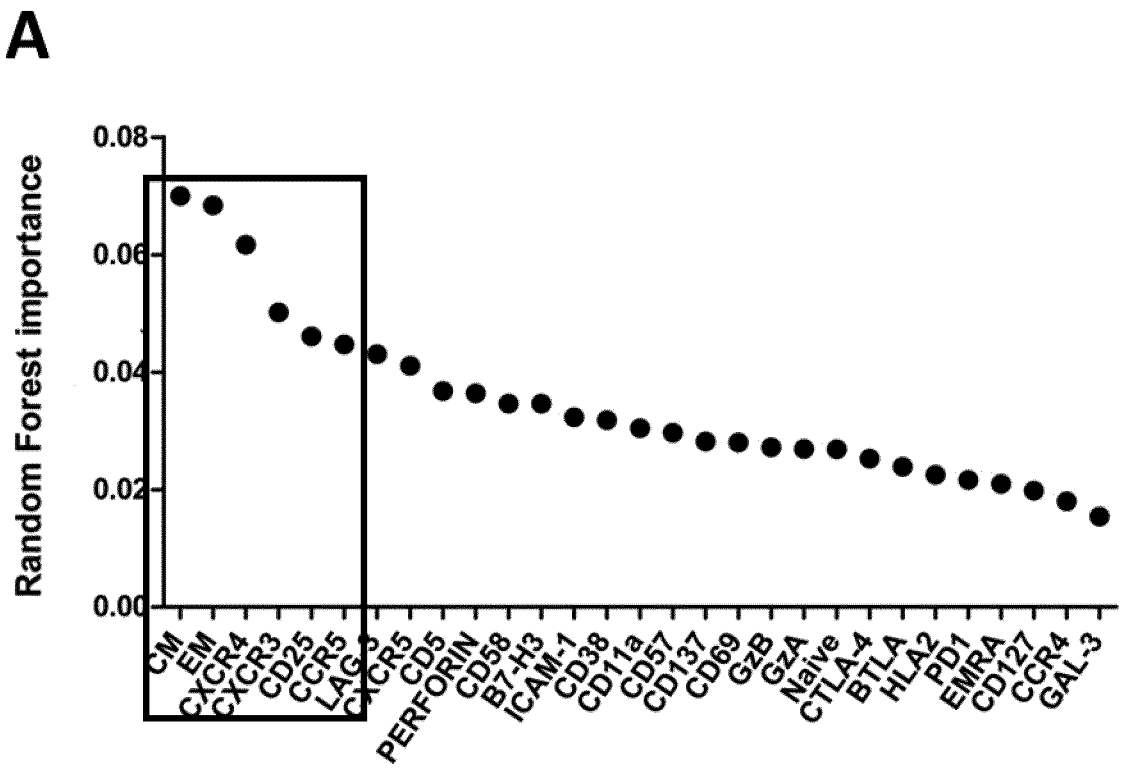


Figure 2A

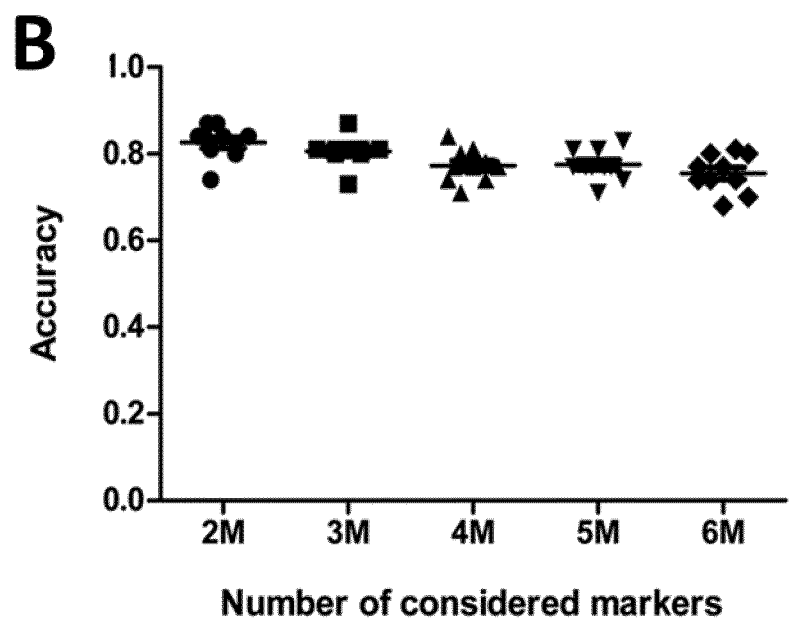


Figure 2B

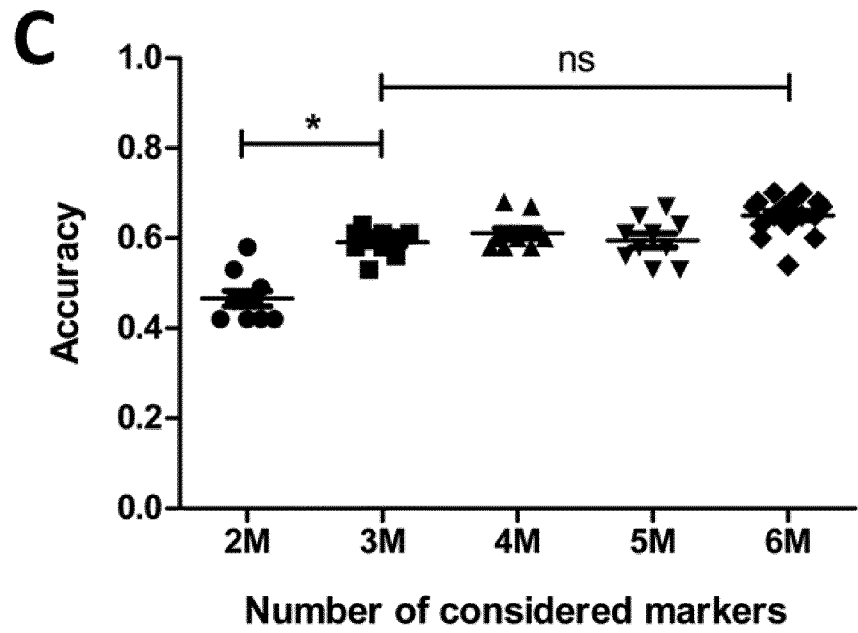


Figure 2C

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2020/050553

A. CLASSIFICATION OF SUBJECT MATTER
INV. G01N33/50 G01N33/569 G01N33/574
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, BIOSIS, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2013/192339 A1 (UNIV CALIFORNIA [US]) 27 December 2013 (2013-12-27) the whole document claims 1-94 paragraphs [0091], [0110] -----	1-9
Y	DEEPAK MITTAL ET AL: "New insights into cancer immunoediting and its three component phases-elimination, equilibrium and escape", CURRENT OPINION IN IMMUNOLOGY., vol. 27, 1 April 2014 (2014-04-01), pages 16-25, XP055451457, GB ISSN: 0952-7915, DOI: 10.1016/j.coi.2014.01.004 the whole document page 7, paragraph 2 ----- -/-	1-9



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

9 March 2020

Date of mailing of the international search report

16/03/2020

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Jenkins, Gareth

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2020/050553

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	MICHEL ENAMORADO ET AL: "Enhanced anti-tumour immunity requires the interplay between resident and circulating memory CD8+ T cells", NATURE COMMUNICATIONS, vol. 8, 17 July 2017 (2017-07-17), page 16073, XP055560573, DOI: 10.1038/ncomms16073 the whole document abstract	1-9
Y	MICHELE W.L. TENG ET AL: "From mice to humans: developments in cancer immunoediting", JOURNAL OF CLINICAL INVESTIGATION, vol. 125, no. 9, 4 August 2015 (2015-08-04), pages 3338-3346, XP055560580, GB ISSN: 0021-9738, DOI: 10.1172/JCI80004 the whole document page 3340, column 1, paragraph 2	1-9
Y	DATABASE BIOSIS [Online] BIOSCIENCES INFORMATION SERVICE, PHILADELPHIA, PA, US; November 2012 (2012-11), GASSNER FRANZ JOSEF ET AL: "T Cell Exhaustion Contributes to Immune Evasion in Chronic Lymphocytic Leukaemia", XP002789186, Database accession no. PREV201300229996 abstract & BLOOD, vol. 120, no. 21, November 2012 (2012-11), page 1773, 54TH ANNUAL MEETING AND EXPOSITION OF THE AMERICAN-SOCIETY-OF-HEMATOLOGY (ASH); ATLANTA, GA, USA; DECEMBER 08 -11, 2012 ISSN: 0006-4971(print)	1-9
Y	WOLF H. FRIDMAN ET AL: "The immune contexture in cancer prognosis and treatment", NATURE REVIEWS CLINICAL ONCOLOGY, vol. 14, no. 12, 25 July 2017 (2017-07-25), pages 717-734, XP055560588, NY, US ISSN: 1759-4774, DOI: 10.1038/nrclinonc.2017.101 the whole document abstract figure 3	1-9

-/--

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2020/050553

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	PAULINE GONNORD ET AL: "Multiparametric analysis of CD8 + T cell compartment phenotype in chronic lymphocytic leukemia reveals a signature associated with progression toward therapy", ONCOIMMUNOLOGY, 7 February 2019 (2019-02-07), pages 1-13, XP055560603, DOI: 10.1080/2162402X.2019.1570774 the whole document -----	1-9

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2020/050553

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2013192339	A1	27-12-2013	NONE
