United States Patent Application Publication	20250255889
Kind Code	A1
Publication Date	August 14, 2025
Inventor(s)	Letai; Anthony et al.

METHODS FOR TREATMENT SELECTION FOR CHRONIC LYMPHOCYTIC LEUKEMIA (CLL)

Abstract

As described below, the present invention features compositions, panels of biomarkers, and methods for selecting a subject with chronic lymphocytic leukemia (CLL) for treatment using an agent and/or for inclusion in a clinical trial using the agent to treat CLL.

Inventors: **Letai; Anthony (Boston, MA), Knisbacher; Binyamin (Cambridge, MA), Parvin; Salma (Boston, MA), Getz; Gad (Boston, MA), Wu; Catherine (Boston, MA)**

Applicant: **The Broad Institute, Inc.** (Cambridge, MA); **The General Hospital Corporation** (Boston, MA); **Dana-Farber Cancer Institute, Inc.** (Boston, MA); **Bar-Ilan University** (Ramat Gan, IL)

Family ID: **90455283**

Assignee: **The Broad Institute, Inc. (Cambridge, MA); The General Hospital Corporation (Boston, MA); Dana-Farber Cancer Institute, Inc. (Boston, MA); Bar-Ilan University (Ramat Gan, IL)**

Appl. No.: **19/085893**

Filed: **March 20, 2025**

Related U.S. Application Data

parent US continuation PCT/US2023/074708 20230920 PENDING child US 19085893 us-provisional-application US 63408452 20220920

Publication Classification

Int. Cl.: **A61K31/635** (20060101); **A61K31/343** (20060101); **A61K31/40** (20060101); **A61K31/4162** (20060101); **A61K31/4184** (20060101); **A61K31/496** (20060101);

A61K31/5025 (20060101); **A61K31/506** (20060101); **A61K31/519** (20060101); **A61K31/52** (20060101); **A61K45/06** (20060101); **A61P35/02** (20060101)

U.S. Cl.:

CPC **A61K31/635** (20130101); **A61K31/343** (20130101); **A61K31/40** (20130101); **A61K31/4162** (20130101); **A61K31/4184** (20130101); **A61K31/496** (20130101); **A61K31/5025** (20130101); **A61K31/506** (20130101); **A61K31/519** (20130101); **A61K31/52** (20130101); **A61K45/06** (20130101); **A61P35/02** (20180101);

Background/Summary

CROSS-REFERENCE TO RELATED APPLICATIONS [0001] This application is a continuation under 35 U.S.C. § 111(a) of PCT Application No. PCT/US2023/074708, filed Sep. 20, 2023, which claims priority to and the benefit of U.S. Provisional Patent Application No. 63/408,452, filed Sep. 20, 2022, the entire contents of each of which are incorporated by reference herein.

BACKGROUND OF THE INVENTION
[0003] Chronic lymphocytic leukemia (CLL) affected about 904,000 people globally in 2015 and resulted in 60,700 deaths. Despite recent advances in chronic lymphocytic leukemia (CLL) therapy, such as the use of targeted agents including Bruton's tyrosine kinase (BTK) inhibitor ibrutinib and the potent BCL-2 antagonist venetoclax, this disease remains incurable for most patients, who are refractory or become resistant to the agents. Thus, identifying new treatment regimens for CLL and building a precision medicine framework that can match CLL patients to the appropriate drugs are of high priority.
SUMMARY OF THE INVENTION
[0004] As described below, the present invention features compositions, panels of biomarkers, and methods for selecting a subject with chronic lymphocytic leukemia (CLL) for treatment using an agent and/or for inclusion in a clinical trial using the agent to treat CLL.
[0005] In one aspect, the invention features a method of treating a selected subject having chronic lymphocytic leukemia, the method comprising administering one of the following agents to the selected subject, wherein the subject is characterized as sensitive to the agent by having one of the following features:
TABLE-US-00001 Drug Feature Direction AZD5991 EC-i Sensitive Azacitidine CARD11 Sensitive Cerdulatinib loss_13q14.13 Sensitive GSK690693 i-CLL Sensitive Rapamycin i-CLL Sensitive Rapamycin EC-m4 Sensitive Rapamycin M-CLL Sensitive Umbralisib M-CLL Sensitive Trametinib CHD2 Sensitive Bendamustine loss_12p13.31a Sensitive Bendamustine loss_5p15.33 Sensitive Bendamustine loss_7p22.2 Sensitive Bendamustine loss_14q32.12 Sensitive Cerdulatinib loss_13q14.3 Sensitive Cerdulatinib loss_13q14.13 Sensitive JQ1 FBXW7 Sensitive Navitoclax M-CLL Sensitive Rapamycin M-CLL Sensitive Ruxolitinib loss_13q14.3 Sensitive Venetoclax loss_13q14.3 Sensitive AZD5991 loss_3p13 Sensitive Cerdulatinib loss_13q14.3 Sensitive Cerdulatinib loss_13q14.13 Sensitive Entospletinib tri_12 Sensitive GSK690693 i-CLL Sensitive JQ1 EC-i Sensitive Selinexor MYD88 Sensitive Trametinib CHD2 Sensitive Vorinostat EC-m2 Sensitive AZD5991 EC-i Sensitive Azacitidine CARD11 Sensitive Cerdulatinib loss_13q14.13 Sensitive GSK690693 i-CLL Sensitive Rapamycin i-CLL Sensitive Rapamycin EC-m4 Sensitive Rapamycin M-CLL Sensitive Umbralisib M-CLL Sensitive Trametinib CHD2 Sensitive Bendamustine loss_12p13.31a Sensitive Bendamustine loss_5p15.33 Sensitive Bendamustine loss_7p22.2 Sensitive Bendamustine loss_14q32.12 Sensitive Cerdulatinib loss_13q14.3 Sensitive Cerdulatinib loss_13q14.13 Sensitive JQ1 FBXW7 Sensitive Navitoclax M-CLL Sensitive Rapamycin M-CLL Sensitive Ruxolitinib loss_13q14.3 Sensitive Venetoclax loss_13q14.3

Sensitive AZD5991 loss_3p13 Sensitive Cerdulatinib loss_13q14.3 Sensitive Cerdulatinib loss_13q14.13 Sensitive Entospletinib tri_12 Sensitive GSK690693 i-CLL Sensitive JQ1 EC-i Sensitive Selinexor MYD88 Sensitive Trametinib CHD2 Sensitive Vorinostat EC-m2 Sensitive AZD5991 EC-i Sensitive Azacitidine CARD11 Sensitive Cerdulatinib loss_13q14.13 Sensitive GSK690693 i-CLL Sensitive Rapamycin i-CLL Sensitive Rapamycin EC-m4 Sensitive Rapamycin M-CLL Sensitive Umbralisib M-CLL Sensitive Trametinib CHD2 Sensitive Bendamustine loss_12p13.31a Sensitive Bendamustine loss_5p15.33 Sensitive Bendamustine loss_7p22.2 Sensitive Bendamustine loss_14q32.12 Sensitive Cerdulatinib loss_13q14.3 Sensitive Cerdulatinib loss_13q14.13 Sensitive JQ1 FBXW7 Sensitive Navitoclax M-CLL Sensitive Rapamycin M-CLL Sensitive Ruxolitinib loss_13q14.3 Sensitive Venetoclax loss_13q14.3 Sensitive AZD5991 loss_3p13 Sensitive Cerdulatinib loss_13q14.3 Sensitive Cerdulatinib loss_13q14.13 Sensitive Entospletinib tri_12 Sensitive GSK690693 i-CLL Sensitive JQ1 EC-i Sensitive Selinexor MYD88 Sensitive Trametinib CHD2 Sensitive Vorinostat EC-m2 Sensitive

[0006] In another aspect, the invention features a method of treating a subject having chronic lymphocytic leukemia, the method comprising administering an agent to a sensitive subject, wherein the subject's sensitivity is determined by identifying the presence of a feature from among the following expression subtypes, drives, genetic alterations, or CLL subtypes, or electing not to administer an agent to a resistant subject wherein the subject's resistance is determined by identifying the presence of a feature from among the following expression subtypes, drives, genetic alterations, or CLL subtypes, wherein the agent, feature, and sensitivity or resistance is as follows: TABLE-US-00002 Agent Feature Direction A-1331852 EC-m2 Resistant AZD5991 EC-i Sensitive Azacitidine CARD11 Sensitive Cerdulatinib loss_13q14.13 Sensitive GSK690693 i-CLL Sensitive Gandotinib EC-i Resistant Navitoclax EC-m2 Resistant Nutlin-3 priortrt_Post Resistant Nutlin-3 FBXW7 Resistant Rapamycin EC-ul Resistant Rapamycin i-CLL Sensitive Rapamycin U-CLL Resistant Rapamycin EC-m4 Sensitive Rapamycin n-CLL Resistant Rapamycin M-CLL Sensitive Umbralisib U-CLL Resistant Umbralisib n-CLL Resistant Umbralisib M-CLL Sensitive Venetoclax EC-m2 Resistant Rapamycin n-CLL Resistant Trametinib CHD2 Sensitive Bendamustine loss_12p13.31a Sensitive Bendamustine loss_5p15.33 Sensitive Bendamustine loss_7p22.2 Sensitive Bendamustine loss_14q32.12 Sensitive Cerdulatinib loss_13q14.3 Sensitive Cerdulatinib loss_13q14.13 Sensitive Gandotinib EC-i Resistant JQ1 FBXW7 Sensitive MK-2206 priortrt_Post Resistant Navitoclax U-CLL Resistant Navitoclax n-CLL Resistant Navitoclax M-CLL Sensitive Nutlin-3 priortrt_Post Resistant Rapamycin U-CLL Resistant Rapamycin M-CLL Sensitive Ruxolitinib loss_13q14.3 Sensitive Venetoclax loss_13q14.3 Sensitive AZD5991 loss_3p13 Sensitive Cerdulatinib loss_13q14.3 Sensitive Cerdulatinib loss_13q14.13 Sensitive Entospletinib tri_12 Sensitive GSK690693 i-CLL Sensitive JQ1 EC-i Sensitive Rapamycin n-CLL Resistant Selinexor MYD88 Sensitive Trametinib CHD2 Sensitive Vorinostat EC-m4 Resistant Vorinostat EC-m2 Sensitive A-1331852 EC-m2 Resistant AZD5991 EC-i Sensitive Azacitidine CARD11 Sensitive Cerdulatinib loss_13q14.13 Sensitive GSK690693 i-CLL Sensitive Gandotinib EC-i Resistant Navitoclax EC-m2 Resistant Nutlin-3 priortrt_Post Resistant Nutlin-3 FBXW7 Resistant Rapamycin EC-ul Resistant Rapamycin i-CLL Sensitive Rapamycin U-CLL Resistant Rapamycin EC-m4 Sensitive Rapamycin n-CLL Resistant Rapamycin M-CLL Sensitive Umbralisib U-CLL Resistant Umbralisib n-CLL Resistant Umbralisib M-CLL Sensitive Venetoclax EC-m2 Resistant Rapamycin n-CLL Resistant Trametinib CHD2 Sensitive Bendamustine loss_12p13.31a Sensitive Bendamustine loss_5p15.33 Sensitive Bendamustine loss_7p22.2 Sensitive Bendamustine loss_14q32.12 Sensitive Cerdulatinib loss_13q14.3 Sensitive Cerdulatinib loss_13q14.13 Sensitive Gandotinib EC-i Resistant JQ1 FBXW7 Sensitive MK-2206 priortrt_Post Resistant Navitoclax U-CLL Resistant Navitoclax n-CLL Resistant Navitoclax M-CLL Sensitive Nutlin-3 priortrt_Post Resistant Rapamycin U-CLL Resistant Rapamycin M-CLL Sensitive Ruxolitinib loss_13q14.3 Sensitive Venetoclax loss_13q14.3 Sensitive AZD5991 loss_3p13 Sensitive Cerdulatinib loss_13q14.3 Sensitive Cerdulatinib loss_13q14.13 Sensitive Entospletinib tri_12 Sensitive

GSK690693 i-CLL Sensitive JQ1 EC-i Sensitive Rapamycin n-CLL Resistant Selinexor MYD88 Sensitive Trametinib CHD2 Sensitive Vorinostat EC-m4 Resistant Vorinostat EC-m2 Sensitive A-1331852 EC-m2 Resistant AZD5991 EC-i Sensitive Azacitidine CARD11 Sensitive Cerdulatinib loss_13q14.13 Sensitive GSK690693 i-CLL Sensitive Gandotinib EC-i Resistant Navitoclax EC-m2 Resistant Nutlin-3 priortrt_Post Resistant Nutlin-3 FBXW7 Resistant Rapamycin EC-ul Resistant Rapamycin i-CLL Sensitive Rapamycin U-CLL Resistant Rapamycin EC-m4 Sensitive Rapamycin n-CLL Resistant Rapamycin M-CLL Sensitive Umbralisib U-CLL Resistant Umbralisib n-CLL Resistant Umbralisib M-CLL Sensitive Venetoclax EC-m2 Resistant Rapamycin n-CLL Resistant Trametinib CHD2 Sensitive Bendamustine loss_12p13.31a Sensitive Bendamustine loss_5p15.33 Sensitive Bendamustine loss_7p22.2 Sensitive Bendamustine loss_14q32.12 Sensitive Cerdulatinib loss_13q14.3 Sensitive Cerdulatinib loss_13q14.13 Sensitive Gandotinib EC-i Resistant JQ1 FBXW7 Sensitive MK-2206 priortrt_Post Resistant Navitoclax U-CLL Resistant Navitoclax n-CLL Resistant Navitoclax M-CLL Sensitive Nutlin-3 priortrt_Post Resistant Rapamycin U-CLL Resistant Rapamycin M-CLL Sensitive Ruxolitinib loss_13q14.3 Sensitive Venetoclax loss_13q14.3 Sensitive AZD5991 loss_3p13 Sensitive Cerdulatinib loss_13q14.3 Sensitive Cerdulatinib loss_13q14.13 Sensitive Entospletinib tri_12 Sensitive GSK690693 i-CLL Sensitive JQ1 EC-i Sensitive Rapamycin n-CLL Resistant Selinexor MYD88 Sensitive Trametinib CHD2 Sensitive Vorinostat EC-m4 Resistant Vorinostat EC-m2 Sensitive (e.g., AZD5991 is administered to a subject identified as sensitive by detecting characteristics of Ec-i in a biological sample of the subject).

[0007] In another aspect, the invention features a method of treating a chronic lymphocytic leukemia (CLL) EC-i expression subtype in a subject, the method involves administering to the subject navitoclax.

[0008] In another aspect, the invention features a method of treating a chronic lymphocytic leukemia (CLL) EC-m1 expression subtype in a subject, the method involves administering to the subject nutlin-3, navitoclax, or cerdulatinib.

[0009] In another aspect, the invention features a method of treating a chronic lymphocytic leukemia (CLL) EC-m2 expression subtype in a subject, the method involves administering to the subject abexinostat, duvelisib, idelalisib, entospletinib, or vorinostat.

[0010] In another aspect, the invention features a method of treating a chronic lymphocytic leukemia (CLL) EC-m3 expression subtype in a subject, the method involves administering to the subject venetoclax, navitoclax, or Abexinostat.

[0011] In another aspect, the invention features a method of treating a chronic lymphocytic leukemia (CLL) EC-m4 expression subtype in a subject, the method involves administering to the subject navitoclax, nutlin-3, or gandotinib.

[0012] In another aspect, the invention features a method of treating a chronic lymphocytic leukemia (CLL) EC-o expression subtype in a subject, the method involves administering to the subject gandotinib, abexinostat, or cerdulatinib.

[0013] In another aspect, the invention features a method of treating a chronic lymphocytic leukemia (CLL) EC-ul expression subtype in a subject, the method involves administering to the subject gandotinib.

[0014] In another aspect, the invention features a method of treating a chronic lymphocytic leukemia (CLL) EC-u2 expression subtype in a subject, the method involves administering to the subject ibrutinib, A-1331852, navitoclax, or rapamycin.

[0015] In another aspect, the invention features a method of treating a chronic lymphocytic leukemia (CLL) M-CLL subtype in a subject, the method involves administering to the subject navitoclax or abexinostat.

[0016] In another aspect, the invention features a method of treating a chronic lymphocytic leukemia (CLL) U-CLL subtype in a subject, the method involves administering to the subject A-1331852, atorvastatin, AZD5991, bendamustine, onalespib, trametinib, voruciclib, or zanubrutinib.

[0017] In another aspect, the invention features a method of treating a chronic lymphocytic leukemia (CLL) in a subject, the method involves administering to the subject: [0018] (a) venetoclax and an MCL1 inhibitor selected from the group consisting of AZD5991, tapotoclax, MIK665, A-1210477, ANJ810, PRT1419, AS00491, APG-3526, CT-03, and CPT-628; or [0019] (b) ibrutinib and a BCL2 inhibitor selected from the group consisting of venetoclax, ZN-d5, lisaftoclax, S55746, and AZD4320.

[0020] In another aspect, the invention features a method of treating a chronic lymphocytic leukemia (CLL) in a selected subject, the method involves administering to the subject an agent with a delta priming value listed in FIG. **14** greater than 15 associated with a driving alteration, wherein the subject is selected as having a neoplasia containing the driving alteration, wherein the driving alteration is: [0021] (a) in a gene encoding a polypeptide selected from the group consisting of ATM, CARD11, CHD2, FBXW7, ITIH2, NOTCH1, NRAS, POTI, SF3B1, TP53, and ZMYM3; or [0022] (b) in a genomic region selected from the group consisting of 7922.1, 15q24.2, 16p11.2, 19pl3.3, 1921.3, 1942.13, 2p11.2, 2q31.1, 3p21.31, 3p13, 5pl5.33, 7p22.2, 9q34.3, 1Opl2.2, 10q24.2, 10q24.32, 11q22.3, 12pl3.31a, 13q14.13, 13q14.3, 14q32.12, 16q22.1, 17pl3.3, 17p13.1, and chromosome 12, and/or 2p.

[0023] In another aspect, the invention features a method for treating a selected subject having chronic lymphocytic leukemia (CLL), the method involves: [0024] (a) characterizing the CLL as having: [0025] i. a mutated (M-CLL) or unmutated IGHV (U-CLL) subtype; [0026] ii. an expression subtype selected from EC-i, EC-m1, EC-m2, EC-m3, EC-m4, EC-o, EC-ul, or EC-u2; and/or [0027] iii. a driving alteration in a gene encoding a polypeptide selected from the group consisting of ATM, CARD11, CHD2, FBXW7, ITIH2, NOTCHI, NRAS, POTI, SF3B1, TP53, and ZMYM3; and/or [0028] iv. a driving alteration in a genomic region selected from the group consisting of 7q22.1, 15q24.2, 16p11.2, 19pl3.3, 1921.3, 1942.13, 2p11.2, 2q31.1, 3p21.31, 3p13, 5pl5.33, 7p22.2, 9q34.3, 1Opl2.2, 10q24.2, 10q24.32, 11q22.3, 12p13.31a, 13q14.13, 13q14.3, 14q32.12, 16q22.1, 17pl3.3, 17p13.1, chromosome 12, and 2p; and [0029] (b) administering an agent to the selected subject, wherein the agent has a delta priming value listed in FIG. **14** greater than 15 associated with the CLL subtype or driving alteration.

[0030] In another aspect, the invention features a method for selecting a subject having chronic lymphocytic leukemia (CLL) for inclusion in or exclusion from a clinical trial to study an agent for treatment of CLL, the method involves: [0031] (a) characterizing the CLL as having: [0032] i. a mutated (M-CLL) or unmutated IGHV (U-CLL) subtype; [0033] ii. an expression subtype selected from EC-i, EC-m1, EC-m2, EC-m3, EC-m4, EC-o, EC-ul, or EC-u2; and/or [0034] iii. a driving alteration in a gene encoding a polypeptide selected from the group consisting of ATM, CARD11, CHD2, FBXW7, ITIH2, NOTCHI, NRAS, POT1, SF3B1, TP53, and ZMYM3; and/or [0035] iv. a driving alteration in a genomic region selected from the group consisting of 7q22.1, 15q24.2, 16p11.2, 19pl3.3, 1921.3, 1942.13, 2p11.2, 2q31.1, 3p21.31, 3p13, 5p15.33, 7p22.2, 9q34.3, 1Op12.2, 10q24.2, 10q24.32, 11q22.3, 12p13.31a, 13q14.13, 13q14.3, 14q32.12, 16q22.1, 17pl3.3, 17pl3.1, chromosome 12, and 2p; and [0036] (b) selecting the subject for inclusion in the clinical trial if the agent has a positive delta priming value of greater than 15 listed in FIG. **14** for the subtype and/or driving alteration of the CLL, and otherwise excluding the subject from the clinical trial.

[0037] The method of the previous aspects, wherein the driving alteration to the genomic region is a duplication or a deletion.

[0038] In another aspect, the invention features a combination therapeutic containing venetoclax and one or more of the following: abexinostat, navitoclax, cerdulatinib, AZD5991, atorvastatin, zanubrutinib, GSK690693, trametinib, ponatinib, bendamustine, nutlin-3, and rapamycin.

[0039] In another aspect, the invention features a combination therapeutic containing venetoclax and one or more of the following: abexinostat, navitoclax, cerdulatinib, AZD5991, atorvastatin, zanubrutinib, GSK690693, trametinib, ponatinib, bendamustine, nutlin-3, and rapamycin.

[0040] In another aspect, the invention features a combination therapeutic containing venetoclax and one or more of the following: abexinostat, navitoclax, cerdulatinib, AZD5991, atorvastatin, 30 zanubrutinib, GSK690693, trametinib, ponatinib, bendamustine, nutlin-3, and rapamycin.
[0041] In another aspect, the invention features a combination therapeutic containing venetoclax and one or more of the following: navitoclax, abexinostat, dasatinib, idelaslisib, duvelisib, cerdulatinib, bendamustine, GSK690693, nirogacestat, trametinib, and rapamycin.
[0042] In another aspect, the invention features a combination therapeutic containing venetoclax and an MCL1 inhibitor.
[0043] In another aspect, the invention features a combination therapeutic containing venetoclax and one or more of the following: abexinostat, navitoclax, cerdulatinib, AZD5991, atorvastatin, 5 zanubrutinib, GSK690693, trametinib, ponatinib, bendamustine, nutlin-3, and rapamycin. In various embodiments, the two agents are formulated separately. In various embodiments, the two agents are administered concurrently or sequentially within at least about 1, 3, 6, 9, 12, or 24 hours of one another. In various embodiments, the two agents are administered within 3, 5, 7, 10, 14, 21, or 28 days of one another.
[0044] In another aspect, the invention features a method of treating a chronic lymphocytic leukemia (CLL) in a subject, the method comprising administering to the subject A-1331852, AZD5991, azacitidine, cerdulatinib, GSK690693, umbralisib, trametinib, bendamustine, cerdulatinib, gandotinib, JQ1, MK-2206, navitoclax, nutlin-3, ruxolitinib, venetoclax, AZD5991, cerdulatinib, entospletinib, GSK690693, JQ1, rapamycin, selinexor, tgrametinib, or vorinostat.
[0045] In various embodiments of any previous aspect, Ec-i comprises an increase in GRIK3, IQGAP2, FCER1G, STK32B, GADD45A, ITGAX, KLF3, RFTN1, PTK2, DFNB31, and ZMAT1 polypeptides, or nucleic acid molecules encoding said polypeptide; [0046] wherein EC-m1 comprises an increase in one or more of TFEC, COL18A1, SLC19A1, NRIP1, KCNH2, P2RX1, ARRDC5, BEX4, and APP polypeptides, or nucleic acid molecules encoding said polypeptide; [0047] wherein EC-m2 comprises an increase in one or more of EML6, HCK, CD1C, VPS37B, CYBB, NXPH4, BTNL9, KLRK1, IQSEC1, BANKI, LEF1, SH3D21, FMOD, SEMA4A, CTLA4, ADTRP, IGSF3, IGFBP4, PDGFD, and APOD polypeptides, or nucleic acid molecules encoding said polypeptide; [0048] wherein EC-m3 comprises an increase in one or more of MS4A4E, MYL9, NT5E, MS4A6A, PITPNC1, CNTNAP2, IGF2BP3, WNT3, CLDN7, TCF7, BASP1, FLJ20373, MAP4K4, LRRK2, SAMSN1, CEACAMI, TNFRSF13B, PHF16, MID1IPI, and ABCA9 polypeptides or nucleic acid molecules encoding said polypeptide; [0049] wherein EC-m4 comprises an increase in one or more of MYBL1, NUGGC, GNG8, AEBP1, HIP1R, LATS2, RIMKLB, EML6, FADS3, MBOAT1, LCN10, DCLK2, and GLUL polypeptides or nucleic acid molecules encoding said polypeptide; [0050] wherein EC-o comprises an increase in one or more of ACSM3, TOX2, PHF16, SESN3, TBC1D9, PIP5K1B, SIK1, DUSP5, GNG7, HIVEP3, MARCKSL1, GPR183, HRK, and PITPNC1 polypeptides, or nucleic acid molecules encoding said polypeptide; [0051] wherein EC-ul comprises an increase in one or more of SEPT10, LDOC1, LPL, KANK2, SOWAHC, DUSP26, OSBPL5, WNT9A, FGFR1, GTSF1L, ADD3, AKT3, COBLL1, MNDA, FCRL3, FAM49A, FCRL2, SLC2A3, and MARCKS polypeptides, or nucleic acid molecules encoding said polypeptide; or [0052] wherein EC-u2 comprises an increase in one or more of ITGB5, BCL7A, PPP1R9A, TSPAN13, SLC12A7, SSBP3, VASH1, SPG20, IL13RA1, NR3C2, TUBG2, ZNF804A, and IL2RA polypeptides, or nucleic acid molecules encoding said polypeptide.
[0053] In various embodiments of any previous aspect, levels of the polypeptide or polypeptide are increased.
[0054] In various embodiments of any previous aspect, M-CLL is treated with navitoclax, nutlin-3, duvelisib, ibrutinib, or venetoclax.
[0055] In various embodiments of any previous aspect, U-CLL is treated with navitoclax, nutlin-3, duvelisib, ibrutinib, dasatinib, venetoclax, or idelasib.

[0056] In various embodiments of any previous aspect, venetoclax is administered in combination with an MCL1 inhibitor.
[0057] In various embodiments of any previous aspect, a subject having CLL characterized as EC-m3, or having a gain of function in 16p11.2, or a loss of function in 13a14.3 is administered venetoclax in combination with an MCL1 inhibitor.
[0058] In various embodiments of any previous aspect, a subject having a CLL characterized as 20 having a trisomy-12 driver is administered zanubrutinib or acalabrutinib.
[0059] In various embodiments of any previous aspect, a subject having CLL characterized as EC-m2, M-CLL, and/or having a trisomy-12 driver is administered zanubrutinib.
[0060] In various embodiments of any previous aspect, a subject having a CLL characterized as EC-i is administered abexinostat.
[0061] In various embodiments of any previous aspect, a subject receiving venetoclax is administered one or more of the following: abexinostat, navitoclax, cerdulatinib, AZD5991, atorvastatin, zanubrutinib, GSK690693, trametinib, ponatinib, bendamustine, nutlin-3, and rapamycin.
[0062] In various embodiments of any previous aspect, a subject receiving venetoclax is administered one or more of the following: abexinostat, navitoclax, cerdulatinib, AZD5991, atorvastatin, zanubrutinib, GSK690693, trametinib, ponatinib, bendamustine, nutlin-3, and rapamycin.
[0063] In various embodiments of any previous aspect, a subject receiving venetoclax is administered one or more of the following: abexinostat, navitoclax, cerdulatinib, AZD5991, atorvastatin, zanubrutinib, GSK690693, trametinib, ponatinib, bendamustine, nutlin-3, and rapamycin.
[0064] In various embodiments of any previous aspect, a subject receiving venetoclax is administered one or more of the following: navitoclax, abexinostat, dasatinib, idelaslisib, 5 duvelisib, cerdulatinib, bendamustine, GSK690693, nirogacestat, trametinib, and rapamycin.
[0065] In various embodiments of any previous aspect, venetoclax is administered in combination with an MCL1 inhibitor.
[0066] In various embodiments of any previous aspect, a subject receiving venetoclax is administered one or more of the following: abexinostat, navitoclax, cerdulatinib, AZD5991, atorvastatin, zanubrutinib, GSK690693, trametinib, ponatinib, bendamustine, nutlin-3, and rapamycin.
[0067] Compositions and articles defined by the invention were isolated or otherwise manufactured in connection with the examples provided below. Other features and advantages of the invention will be apparent from the detailed description, and from the claims.

Definitions

[0068] Unless defined otherwise, all technical and scientific terms used herein have the meaning commonly understood by a person skilled in the art to which this invention belongs. The following references provide one of skill with a general definition of many of the terms used in this invention: Singleton et al., Dictionary of Microbiology and Molecular Biology (2nd ed. 1994); The Cambridge Dictionary of Science and Technology (Walker ed., 1988); The Glossary of Genetics, 5th Ed., R. Rieger et al. (eds.), Springer Verlag (1991); and Hale & Marham, The Harper Collins Dictionary of Biology (1991). As used herein, the following terms have the meanings ascribed to them below, unless specified otherwise.
[0069] The terms “biomarker” and “marker” are used interchangeably herein to refer to a protein, nucleic acid molecule, clinical indicator, or other analyte that is associated with a disease. In one embodiment, a marker of chronic lymphocytic leukemia (CLL) is differentially present in a biological sample obtained from a subject having or at risk of developing chronic lymphocytic leukemia (CLL) relative to a reference. A marker is differentially present if the mean or median level of the biomarker present in the sample is statistically different from the level present in a

reference. A reference level may be, for example, the level present in a sample obtained from a healthy control subject or the level obtained from the subject at an earlier timepoint, i.e., prior to treatment. Common tests for statistical significance include, among others, t-test, ANOVA, Kruskal-Wallis, Wilcoxon, Mann-Whitney and odds ratio. Biomarkers, alone or in combination, provide measures of relative likelihood that a subject belongs to a phenotypic status of interest. Biomarkers can be used to classify a chronic lymphocytic leukemia (CLL). The differential presence of a marker of the invention in a subject sample can be useful in characterizing the subject as having or at risk of developing chronic lymphocytic leukemia (CLL), for determining the prognosis of the subject, for evaluating therapeutic efficacy, or for selecting a treatment regimen (e.g., selecting that the subject be evaluated and/or treated by a surgeon who specializes in chronic lymphocytic leukemia (CLL)). The invention includes markers that share at least about 85%, 90%, 95% or even 99% to a polypeptide sequence corresponding to a biomarker listed in Table 3A or Table 4. The invention includes markers that share at least about 85%, 90%, 95% or even 99% to a polynucleotide sequence corresponding to a gene listed in Table 3A or Table 4.

[0070] By “AT13387” is meant a chemical corresponding to CAS No. 912999-49-6, having the chemical structure

##STR00001##

and pharmaceutically acceptable salts thereof.

[0071] By “AZD7762” is meant a chemical corresponding to CAS No. 860352-01-8, having the chemical structure

##STR00002##

and pharmaceutically acceptable salts thereof.

[0072] By “dasatinib” is meant a chemical corresponding to CAS No. 302962-49-8, having the chemical structure

##STR00003##

and pharmaceutically acceptable salts thereof.

[0073] By “duvelisib” is meant a chemical corresponding to CAS No. 1201438-56-3, having the chemical structure

##STR00004##

and pharmaceutically acceptable salts thereof.

[0074] By “fludarabine” is meant a chemical corresponding to CAS No. 21679-14-1, having the chemical structure

##STR00005##

and pharmaceutically acceptable salts thereof.

[0075] By “ibrutinib” is meant a chemical corresponding to CAS No. 936563-96-1, having the chemical structure

##STR00006##

and pharmaceutically acceptable salts thereof.

[0076] By “idelalisib” is meant a chemical corresponding to CAS No. 870281-82-6, having the chemical structure

##STR00007##

and pharmaceutically acceptable salts thereof.

[0077] By “navitoclax” is meant a chemical corresponding to CAS No. 923564-51-6, having the chemical structure

##STR00008##

and pharmaceutically acceptable salts thereof.

[0078] By “PRT062607 HCL” is meant a chemical corresponding to CAS No. 1370261-97-4, having the chemical structure

##STR00009##

and pharmaceutically acceptable salts thereof.

[0079] By “selumetinib” is meant a chemical corresponding to CAS No. 606143-52-6, having the chemical structure
##STR00010##
and pharmaceutically acceptable salts thereof.
[0080] By “SNS-032” is meant a chemical corresponding to CAS No. 345627-80-7, having the chemical structure
##STR00011##
and pharmaceutically acceptable salts thereof.
[0081] By “venetoclax” is meant a chemical corresponding to CAS No. 1257044-40-8, having the chemical structure
##STR00012##
and pharmaceutically acceptable salts thereof.
[0082] By “agent” is meant any small molecule chemical compound, antibody, nucleic acid molecule, or polypeptide, or fragments thereof.
[0083] By “ameliorate” is meant to decrease, suppress, attenuate, diminish, arrest, or stabilize the development or progression of a disease.
[0084] By “alteration” or “change” is meant an increase or decrease. An alteration may be by as little as 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, or by 40%, 50%, 60%, or even by as much as 70%, 75%, 80%, 90%, or 100%.
[0085] By “analog” is meant a molecule that is not identical but has analogous functional or structural features. For example, a polypeptide analog retains the biological activity of a corresponding naturally-occurring polypeptide, while having certain biochemical modifications that enhance the analog's function relative to a naturally occurring polypeptide. Such biochemical modifications could increase the analog's protease resistance, membrane permeability, or half-life, without altering, for example, ligand binding. An analog may include an unnatural amino acid.
[0086] By “biological sample” is meant any tissue, cell, fluid, or other material derived from an organism. Non-limiting examples of biological samples include a bodily fluid (such as blood, blood serum, plasma, saliva, urine, ascites, cyst fluid, and the like); a homogenized tissue sample (e.g., a tissue sample obtained by biopsy); and a cell isolated from a patient sample.
[0087] By “capture molecule” or “capture reagent” is meant a reagent that specifically binds a nucleic acid molecule or polypeptide to label, select, or isolate the nucleic acid molecule or polypeptide. Non-limiting examples of capture molecules include polynucleotide probes, antibodies, and fragments thereof.
[0088] By “Chronic Lymphocytic Leukemia (CLL)” is meant a B cell neoplasm. In embodiments, CLL is diagnosed using:
[0089] Blood tests: These tests show the extent of cancer and any signs of infection. Blood tests measure levels of white and red blood cells, the amount of inflammation in the body, and liver and kidney function. A blood test can also look for genetic changes.
[0090] Bone marrow biopsy and aspiration: Doctors use these tests to look for leukemia cells in the bone marrow. They use thin, hollow needles to remove small samples of bone marrow and bone tissue for analysis.
[0091] Lymph node biopsy: A doctor may remove part or all of a lymph node (gland that helps your body fight infection) to examine it for signs of cancer.
[0092] Genetic testing: Doctors may use bone marrow samples to look for genetic changes that can lead to CLL. Genetic information can help guide treatment as described herein below.
[0093] Imaging: Doctors may use these tests, which produce detailed images of the body, to check for signs of cancer in other parts of the body. Imaging tests may include CT scan or ultrasound. In embodiments, CLL is characterized using features described herein.
[0094] As used herein, the terms “determining”, “assessing”, “assaying”, “measuring” and “detecting” refer to both quantitative and qualitative determinations, and as such, the term

“determining” is used interchangeably herein with “assaying,” “measuring,” and the like.
[0095] In this disclosure, “comprises,” “comprising,” “containing” and “having” and the like can have the meaning ascribed to them in U.S. Patent law and can mean “includes,” “including,” and the like; “consisting essentially of” or “consists essentially” likewise has the meaning ascribed in U.S. Patent law and the term is open-ended, allowing for the presence of more than that which is recited so long as basic or novel characteristics of that which is recited is not changed by the presence of more than that which is recited, but excludes prior art embodiments. Any embodiments specified as “comprising” a particular component(s) or element(s) are also contemplated as “consisting of” or “consisting essentially of” the particular component(s) or element(s) in some embodiments.
[0096] By “delta priming” is meant the difference in priming of a cell for apoptosis measured in the presence of an agent for treating chronic lymphocytic leukemia relative to priming of the cell in the presence of an inert carrier. In embodiments, the inert carrier is DMSO.
[0097] “Detect” refers to identifying the presence, absence or amount of the analyte to be detected.
[0098] By “driver alteration” is meant a genomic alteration that is associated with an increase in cell proliferation relative to an unaltered cell. Non-limiting examples of genes that can comprise driver alterations include ATM, CARD11, CHD2, FBXW7, ITIH2, NOTCH1, NRAS, POT1, SF3B1, TP53, and ZMYM3. Non-limiting examples of genomic region alterations that can be driver alterations include a duplication of 7q22.1, duplication of 15q24.2, duplication of 16p11.2, duplication of 19p13.3, deletion of 1q21.3, deletion of 1q42.13, deletion of 2p11.2, deletion of 2q31.1, deletion of 3p21.31, deletion of 3p13, deletion of 5p15.33, deletion of 7p22.2, deletion of 9q34.3, deletion of 10p12.2, deletion of 10q24.2, deletion of 10q24.32, deletion of 11q22.3, deletion of 12p13.31a, deletion of 13q14.13, deletion of 13q14.3, deletion of 14q32.12, deletion of 16q22.1, deletion of 17pl3.3, deletion of 17pl3.1, tri_12, and/or duplication of 2p.
[0099] By “molecular identifier” is meant an agent that when linked to a molecule of interest renders the latter detectable, via spectroscopic, photochemical, biochemical, immunochemical, or chemical means. For example, useful labels include radioactive isotopes, magnetic beads, metallic beads, colloidal particles, fluorescent dyes, electron-dense reagents, enzymes (for example, as commonly used in an ELISA), biotin, digoxigenin, or haptens.
[0100] By “disease” is meant any condition or disorder that damages or interferes with the normal function of a cell, tissue, or organ. Examples of diseases include chronic lymphocytic leukemia (CLL) and the like.
[0101] By “effective amount” is meant the amount of an agent required to ameliorate the symptoms of a disease relative to an untreated patient. The effective amount of active compound(s) used to practice the present invention for therapeutic treatment of a disease varies depending upon the manner of administration, the age, body weight, and general health of the subject. Ultimately, the attending physician or veterinarian will decide the appropriate amount and dosage regimen. Such amount is referred to as an “effective” amount.
[0102] The term “expression cluster (EC)” describes a set of genes that are co-expressed and exhibit coordinated behavior. See, for example, Abu-Jamous, B., Kelly, S. Clust: automatic extraction of optimal co-expressed gene clusters from gene expression data. Genome Biol 19, 172 (2018). https://doi.org/10.1186/s13059-018-1536-8. Expression clusters can be used to characterize disease subtypes. Expression clusters used to characterize chronic lymphocytic leukemia include the following: Ec-i, EC-m1, EC-m2, EC-m3, EC-m4, EC-o, EC-ul, and EC-u2.
[0103] In embodiments, markers useful in the panels of the invention include markers for expression cluster Ec-i, namely, GRIK3, IQGAP2, FCER1G, STK32B, GADD45A, ITGAX, KLF3, RFTN1, PTK2, DFNB31, and ZMAT1, or a sub-set thereof, as well as the nucleic acid molecules encoding such proteins. In embodiments, levels of one or more of these markers are increased.
[0104] In embodiments, markers useful in the panels of the invention include markers for

expression cluster EC-m1, namely, TFEC, COL18A1, SLC19A1, NRIP1, KCNH2, P2RX1, ARRDC5, BEX4, and APP, or a sub-set thereof, as well as the nucleic acid molecules encoding such proteins. In embodiments, levels of one or more of these markers are increased.
[0105] In embodiments, markers useful in the panels of the invention include markers for expression cluster EC-m2, namely, EML6, HCK, CD1C, VPS37B, CYBB, NXPH4, BTNL9, KLRK1, IQSECT, BANK1, LEF1, SH3D21, FMOD, SEMA4A, CTLA4, ADTRP, IGSF3, IGFBP4, PDGFD, and APOD, or a sub-set thereof, as well as the nucleic acid molecules encoding such proteins. In embodiments, levels of one or more of these markers are increased.
[0106] In embodiments, markers useful in the panels of the invention include markers for expression cluster EC-m3, namely, MS4A4E, MYL9, NT5E, MS4A6A, PITPNC1, CNTNAP2, 10 IGF2BP3, WNT3, CLDN7, TCF7, BASP1, FLJ20373, MAP4K4, LRRK2, SAMSN1, CEACAM1, TNFRSF13B, PHF16, MID1IP1, and ABCA9, or a sub-set thereof, as well as the nucleic acid molecules encoding such proteins. In embodiments, levels of one or more of these markers are increased.
[0107] In embodiments, markers useful in the panels of the invention include markers for expression cluster EC-m4, namely, MYBL1, NUGGC, GNG8, AEBP1, HIP1R, LATS2, RIMKLB, EML6, FADS3, MBOAT1, LCN10, DCLK2, and GLUL, or a sub-set thereof, as well as the nucleic acid molecules encoding such proteins. In embodiments, levels of one or more of these markers are increased.
[0108] In embodiments, markers useful in the panels of the invention include markers for expression cluster EC-o, namely, ACSM3, TOX2, PHF16, SESN3, TBC1D9, PIP5K1B, SIK1, DUSP5, GNG7, HIVEP3, MARCKSL1, GPR183, HRK, and PITPNC1, or a sub-set thereof, as well as the nucleic acid molecules encoding such proteins. In embodiments, levels of one or more of these markers are increased.
[0109] In embodiments, markers useful in the panels of the invention include markers for expression cluster EC-ul, namely, SEPT10, LDOC1, LPL, KANK2, SOWAHC, DUSP26, OSBPL5, WNT9A, FGFR1, GTSF1L, ADD3, AKT3, COBLL1, MNDA, FCRL3, FAM49A, FCRL2, SLC2A3, and MARCKS, or a sub-set thereof, as well as the nucleic acid molecules encoding such proteins. In embodiments, levels of one or more of these markers are increased.
[0110] In embodiments, markers useful in the panels of the invention include markers for expression cluster EC-u2, namely, ITGB5, BCL7A, PPP1R9A, TSPAN13, SLC12A7, SSBP3, VASH1, SPG20, IL13RA1, NR3C2, TUBG2, ZNF804A, and IL2RA, or a sub-set thereof, as well as the nucleic acid molecules encoding such proteins. The panels can comprise biomarkers for expression cluster Ec-i, EC-m1, EC-m2, EC-m3, EC-m4, EC-o, EC-ul, or EC-u2, or various combinations thereof. In embodiments, levels of one or more of these markers are increased. By “fragment” is meant a portion of a polypeptide or nucleic acid molecule. This portion contains, preferably, at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% of the entire length of the reference nucleic acid molecule or polypeptide. A fragment may contain 10, 20, 30, 40, 50, 60, 70, 80, 90, or 100, 200, 300, 400, 500, 600, 700, 800, 900, or 1000 nucleotides or amino acids.
[0111] By “increase” is meant to alter positively An increase may be by about or at least about 0.5%, 1%, 5%, 10%, 25%, 30%, 50%, 75%, or even by 100%.
[0112] The terms “isolated,” “purified,” or “biologically pure” refer to material that is free to varying degrees from components which normally accompany it as found in its native state.
[0113] “Isolate” denotes a degree of separation from an original source or surroundings. “Purify” denotes a degree of separation that is higher than isolation. A “purified” or “biologically pure” protein is sufficiently free of other materials such that any impurities do not materially affect the biological properties of the protein or cause other adverse consequences. That is, a nucleic acid or peptide of this invention is purified if it is substantially free of cellular material, viral material, or culture medium when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized. Purity and homogeneity are typically determined using

analytical chemistry techniques, for example, polyacrylamide gel electrophoresis or high performance liquid chromatography. The term “purified” can denote that a nucleic acid or protein gives rise to essentially one band in an electrophoretic gel. For a protein that can be subjected to modifications, for example, phosphorylation or glycosylation, different modifications may give rise to different isolated proteins, which can be separately purified.

[0114] By “isolated polynucleotide” is meant a nucleic acid that is free of the genes which, in the naturally-occurring genome of the organism from which the nucleic acid molecule of the invention is derived, flank the gene. The term therefore includes, for example, a recombinant DNA that is incorporated into a vector; into an autonomously replicating plasmid or virus; or into the genomic DNA of a prokaryote or eukaryote; or that exists as a separate molecule (for example, a cDNA or a genomic or cDNA fragment produced by PCR or restriction endonuclease digestion) independent of other sequences. In addition, the term includes an RNA molecule that is transcribed from a DNA molecule, as well as a recombinant DNA that is part of a hybrid gene encoding additional polypeptide sequence.

[0115] By an “isolated polypeptide” is meant a polypeptide of the invention that has been separated from components that naturally accompany it. Typically, the polypeptide is isolated when it is at least 60%, by weight, free from the proteins and naturally-occurring organic molecules with which it is naturally associated. Preferably, the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight, a polypeptide of the invention. An isolated polypeptide of the invention may be obtained, for example, by extraction from a natural source, by expression of a recombinant nucleic acid encoding such a polypeptide; or by chemically synthesizing the protein. Purity can be measured by any appropriate method, for example, column chromatography, polyacrylamide gel electrophoresis, or by HPLC analysis.

[0116] By “marker profile” is meant a characterization of the expression or expression level of two or more polypeptides or polynucleotides in a sample.

[0117] As used herein, “obtaining” as in “obtaining an agent” includes synthesizing, purchasing, or otherwise acquiring the agent.

[0118] By “polypeptide” or “amino acid sequence” is meant any chain of amino acids, regardless of length or post-translational modification. In various embodiments, the post-translational modification is glycosylation or phosphorylation. In various embodiments, conservative amino acid substitutions may be made to a polypeptide to provide functionally equivalent variants, or homologs of the polypeptide. In some aspects, the invention embraces sequence alterations that result in conservative amino acid substitutions. In some embodiments, a “conservative amino acid substitution” refers to an amino acid substitution that does not alter the relative charge or size characteristics of the protein in which the conservative amino acid substitution is made.

[0119] Variants can be prepared according to methods for altering polypeptide sequence known to one of ordinary skill in the art such as are found in references that compile such methods, e.g.

[0120] Molecular Cloning: A Laboratory Manual, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989, or Current Protocols in Molecular Biology, F. M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. Non-limiting examples of conservative substitutions of amino acids include substitutions made among amino acids within the following groups: (a) M, I, L, V; (b) F, Y, W; (c) K, R, H; (d) A, G; (e) S, T; (f) Q, N; and (g) E, D. In various embodiments, conservative amino acid substitutions can be made to the amino acid sequence of the proteins and polypeptides disclosed herein.

[0121] “Primer set” means a set of oligonucleotides. A primer set may comprise at least about 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 30, 40, 50, 60, 80, 100, 200, 250, 300, 400, 500, 600, or more primers. In embodiments, the primers are used for detection of a biomarker(s) in a sample (e.g., by PCR, targeted sequencing, biochip, or any of various other methods described herein or combinations thereof).

[0122] By “reduce” is meant to alter negatively. A reduction may be by about or at least about

0.5%, 1%, 5%, 10%, 25%, 30%, 50%, 75%, or even by 100%.
[0123] By “reference” is meant a standard or control condition. In embodiments, the reference is the level of an analyte present in a sample obtained from a subject prior to being administered a treatment, obtained from a healthy subject (e.g., a subject not having a chronic lymphocytic leukemia (CLL)), or a sample obtained from a subject at an earlier time point than a particular sample time point.
[0124] A “reference sequence” is a defined sequence used as a basis for sequence comparison. A reference sequence may be a subset of or the entirety of a specified sequence; for example, a segment of a full-length cDNA or gene sequence, or the complete cDNA or gene sequence. For polypeptides, the length of the reference polypeptide sequence will generally be at least about 16 amino acids, preferably at least about 20 amino acids, more preferably at least about 25 amino acids, and even more preferably about 35 amino acids, about 50 amino acids, or about 100 amino acids. For nucleic acids, the length of the reference nucleic acid sequence will generally be at least about 50 nucleotides, preferably at least about 60 nucleotides, more preferably at least about 75 nucleotides, and even more preferably about 100 nucleotides or about 300 nucleotides or any integer thereabout or therebetween.
[0125] By “specifically binds” is meant an agent that recognizes and binds a polypeptide or polynucleotide of the invention, but which does not substantially recognize and bind other molecules in a sample, for example, a biological sample, which naturally includes a polypeptide or polynucleotide described herein.
[0126] Nucleic acid molecules useful in the methods of the invention include any nucleic acid molecule that encodes a polypeptide of the invention or a fragment thereof. Such nucleic acid molecules need not be 100% identical with an endogenous nucleic acid sequence, but will typically exhibit substantial identity. Polynucleotides having “substantial identity” to an endogenous sequence are typically capable of hybridizing with at least one strand of a double-stranded nucleic acid molecule. By “hybridize” is meant to pair to form a double-stranded molecule between complementary polynucleotide sequences (e.g., a gene described herein), or portions thereof, under various conditions of stringency. (See, e.g., Wahl, G. M. and S. L. Berger (1987) Methods Enzymol. 152:399; Kimmel, A. R. (1987) Methods Enzymol. 152:507).
[0127] For example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and more preferably less than about 250 mM NaCl and 25 mM trisodium citrate. Low stringency hybridization can be obtained in the absence of organic solvent, e.g., formamide, while high stringency hybridization can be obtained in the presence of at least about 35% formamide, and more preferably at least about 50% formamide. Stringent temperature conditions will ordinarily include temperatures of at least about 30° C., more preferably of at least about 37° C., and most preferably of at least about 42° C. Varying additional parameters, such as hybridization time, the concentration of detergent, e.g., sodium dodecyl sulfate (SDS), and the inclusion or exclusion of carrier DNA, are well known to those skilled in the art. Various levels of stringency are accomplished by combining these various conditions as needed. In a preferred embodiment, hybridization will occur at 30° C. in 750 mM NaCl, 75 mM trisodium citrate, and 1% SDS. In a more preferred embodiment, hybridization will occur at 37° C. in 500 mM NaCl, 50 mM trisodium citrate, 1% SDS, 35% formamide, and 100 μg/ml denatured salmon sperm DNA (ssDNA). In a most preferred embodiment, hybridization will occur at 42° C. in 250 mM NaCl, 25 mM trisodium citrate, 1% SDS, 50% formamide, and 200 μg/ml ssDNA. Useful variations on these conditions will be readily apparent to those skilled in the art.
[0128] For most applications, washing steps that follow hybridization will also vary in stringency. Wash stringency conditions can be defined by salt concentration and by temperature.
[0129] As above, wash stringency can be increased by decreasing salt concentration or by increasing temperature. For example, stringent salt concentration for the wash steps will preferably

be less than about 30 mM NaCl and 3 mM trisodium citrate, and most preferably less than about 15 mM NaCl and 1.5 mM trisodium citrate. Stringent temperature conditions for the wash steps will ordinarily include a temperature of at least about 25° C., more preferably of at least about 42° C., and even more preferably of at least about 68° C. In a preferred embodiment, wash steps will occur at 25° C. in 30 mM NaCl, 3 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 42° C. in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 68° C. in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. Additional variations on these conditions will be readily apparent to those skilled in the art. Hybridization techniques are well known to those skilled in the art and are described, for example, in Benton and Davis (Science 196:180, 1977); Grunstein and Hogness (Proc. Natl. Acad. Sci., USA 72:3961, 1975); Ausubel et al. (Current Protocols in Molecular Biology, Wiley Interscience, New York, 2001); Berger and Kimmel (Guide to Molecular Cloning Techniques, 1987, Academic Press, New York); and Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York.

[0130] By “substantially identical” is meant a polypeptide or nucleic acid molecule exhibiting at least 50% identity to a reference amino acid sequence (for example, any one of the amino acid sequences described herein) or nucleic acid sequence (for example, any one of the nucleic acid sequences described herein). In embodiments, such a sequence is at least 60%, 80%, 85%, 90%, 95% or even 99% identical at the amino acid level or nucleic acid to the sequence used for comparison.

[0131] Sequence identity is typically measured using sequence analysis software (for example, Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, Wis. 53705, BLAST, BESTFIT, GAP, or PILEUP/PRETTYBOX programs). Such software matches identical or similar sequences by assigning degrees of homology to various substitutions, deletions, and/or other modifications. Conservative substitutions typically include substitutions within the following groups: glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid, asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine. In an exemplary approach to determining the degree of identity, a BLAST program may be used, with a probability score between e.sup.−3 and e.sup.−100 indicating a closely related sequence.

[0132] By “subject” is meant an animal. The animal can be a mammal. The mammal can be a human or non-human mammal, such as a bovine, equine, canine, ovine, rodent, or feline.

[0133] Ranges provided herein are understood to be shorthand for all of the values within the range. For example, a range of 1 to 50 is understood to include any number, combination of numbers, or sub-range from the group consisting of 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, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50.

[0134] As used herein, the terms “treat,” treating,” “treatment,” and the like refer to reducing or ameliorating a disorder and/or symptoms associated therewith. It will be appreciated that, although not precluded, treating a disorder or condition does not require that the disorder, condition or symptoms associated therewith be completely eliminated.

[0135] Unless specifically stated or obvious from context, as used herein, the term “or” is understood to be inclusive. Unless specifically stated or obvious from context, as used herein, the terms “a”, “an”, and “the” are understood to be singular or plural.

[0136] Unless specifically stated or obvious from context, as used herein, the term “about” is understood as within a range of normal tolerance in the art.

[0137] The recitation of a listing of chemical groups in any definition of a variable herein includes definitions of that variable as any single group or combination of listed groups. The recitation of an embodiment for a variable or aspect herein includes that embodiment as any single embodiment or in combination with any other embodiments or portions thereof.

[0138] Any compositions or methods provided herein can be combined with one or more of any of the other compositions and methods provided herein.

Description

BRIEF DESCRIPTION OF THE DRAWINGS

[0139] FIG. **1** provides a schematic overview of a design for a functional precision study integrating high-throughput dynamic BH3 profiling (HT-DBP) (top portion of FIG. **1**) and molecular profiling (lower portion of FIG. **1**). BH3 profiling is a functional tool that measures mitochondrial apoptotic priming. It uses BH3 peptides derived from the BH3 domain of pro-apoptotic BH3-only proteins to provoke a response from viable mitochondria.

[0140] FIGS. **2**A and **2**B provide a genomic landscape diagram and a bar graph. FIG. **2**A provides a genomic landscape diagram providing an overview of a chronic lymphocytic leukemia (CLL) cohort (68 samples; 65 patients (+3 pre/post treatment); 64 in CLL-map). For FIG. **2**A, whole-exome sequencing/whole genome sequencing (WES/WGS) (n=61), RNA (n=57, 10 new), methylation (n=47), CW22 (U-CLL) treatment was FCR, CW32 (U-CLL) treatment was FR_REV (later F_REV, FCR, BR+PCI), and CW48 (M-CLL) treatment was R (later R, FR). The cohort represented in FIGS. **2**A and **2**B was enriched in M-CLLs and all epitypes and expression clusters (ECs) were represented. FIG. **2**B provides a bar graph providing a breakdown of the numbers of different expression clusters observed within the CLL cohort.

[0141] FIG. **3** provides a schematic providing an overview of the high-throughput dynamic BH3 profiling (HT-DBP) screen.

[0142] FIG. **4** provides a schematic illustration of the anti-apoptotic proteins targeted by the different BH3 peptides used in the HT-DBP screen. Peptides that were promiscuous in the anti-apoptotic proteins that they targeted were considered “activators” and peptides that were selective in the anti-apoptotic proteins that they targeted were considered “sensitizers.” Different peptides, therefore, provided different information with regard to a drug's impact on a CLL cell.

[0143] FIGS. **5**A and **5**B provide plots demonstrating that dynamic BH3 profiling screens gave high quality and reproducible results. In FIG. **5**A, each point is the mean of 3 replicate comparisons per plate (A-vs-B, A-vs-C, B-vs-C). Replicates correlated across patients.

[0144] FIGS. **6**A and **6**B provide a heat map and a dendrogram. FIG. **6**A provides a heatmap that shows BH3 peptide effect similarity based on Pearson's correlations across different drugs across patients. PUMA, BIM—non-specific; BAD-BCL2 inhibitor; MS1-MCL1 inhibitor. FIG. **6**B provides a dendrogram showing peptide effect similarity.

[0145] FIGS. **7**A, **7**B, and **7**C provide a schematic and heatmaps showing the landscape of response in CLL samples. FIG. **7**A provides a schematic comparing the positive delta-priming (i.e., priming of an apoptosis in the presence of an agent for treating CLL as compared to priming in the presence of DMSO) values observed for cells contacted with the indicated BH3 peptides at the indicated concentrations in combination with the 42 drugs evaluated. FIGS. **7**B and **7**C provide heat maps showing that delta priming values were consistent for each indicated drug (left of heat maps) across all of the evaluated BH3 peptides (top of heat maps). Current first-line treatments for CLL include Venetoclax and/or Ibrutinib, which fell within the top-ten drugs with the highest positive delta priming values.

[0146] FIGS. **8**A and **8**B provide histograms showing that dynamic BH3 profiling screens provided combination therapy leads. FIG. **8**A provides a histogram showing that a combined treatment involving administration of Venetoclax (BCL2 inhibitor) and MS1 (MCL1 inhibitor) had an increased delta priming value relative to alternative treatments. FIG. **8**B provides a histogram showing that a combined treatment involving Iburtinib (Bruton's tyrosine kinase (BTK) inhibitor) and BAD (BCL2 inhibitor) had an increased delta priming value relative to alternative treatments.

DBP can be used to evaluate efficacy of such combination therapies in a clinical setting (e.g., Venetoclax combined with an MCL1 inhibitor, such as AZD5991, or Iburtinib combined with a BCL2 inhibitor, such as Venetoclax).

[0147] FIG. **9** provides a plot showing drugs that had high median delta priming in the U-CLL IGHV subtype (drugs with points above diagonal line) and drugs that had high median delta priming in the M-CLL IGHV subtype (drugs with points below diagonal line).

[0148] FIGS. **10**A, **10**B, 1° C. and 1OD provide a consensus matrix, a heat map, stacked bar plots, a scatter plot, and a line plot showing that gene expression clusters (ECs) revealed 8 distinct chronic lymphocytic leukemia (CLL) subtypes. 2 IGHV-unmutated/n-CLL clusters; 4 IGHV-mutated/m-CLL clusters; EC-i associated with i-CLL and IGLV3-21 R110 mutations; EC-m2 & EC-u2: tri(12)-enriched. FIG. **10**A provides a consensus matrix for RNA expression profiles of 610 treatment-naive CLLs by repeated hierarchical clustering with 80% resampling and varying cutoffs for number of clusters. This matrix served as input to a Bayesian non-negative matrix factorization (BayesNMF) method for inferring the total number of clusters and sample assignment to clusters. FIG. **10**B provides a heat map and a stacked bar graph together showing eight gene expression clusters (ECs, columns) identified by a Bayesian non-negative matrix factorization (BNMF) method in 610 treatment-naive samples. FIG. **10**C provides a plot showing a Kaplan Meier analysis of the impact of expression clusters on overall survival (OS) probabilities in 609 treatment-naive samples (log-rank test). FIG. **10**D provides a plot showing uniform manifold approximation and projection (UMAP) showing clustering of expression clusters (ECs).

[0149] FIG. **11** provides a dendrogram that shows, together with FIG. **10**B, that the expression clusters were distinguished by molecular features and drivers. FIG. **11** provides a dendrogram of expression clusters (ECs) with associated upregulated and downregulated biologic pathways determined by gene set enrichment analysis. The clusters varied in size and are segregated by biological features and defined subtypes of the IGHV subtypes: 2 clusters, represent U-CLLs, which give them the EC-u prefix. 4 clusters given the EC-m prefix, were strongly associated with mutated IGHV. The last cluster, named EC-i, was associated with the intermediate methylation subtype. The clusters differed by their genetic driver landscapes: a) EC-m2 and EC-u2 were strongly associated with tri(12) events, jointly containing>85% of tri(12) events; b) EC-i was defined by a specific variant in the Ig light chain, which led to constitutive B-cell receptor signaling and was shown to be associated with adverse outcome. There were 4 EC-ms, 2 EC-us, and 2 ECs that associated with tri(12). Some ECs were more defined by unique pathways, such as enhanced Oxphos in what was named EC-o and Inflammatory signaling in EC-m4. Some pathways helped distinguish between clusters of the same IGHV subtype. For example, both U-CLL clusters shared downregulation of translation, but differed in TNF-alpha signaling. Also, among M-CLLs, EC-m2 had increased antigen processing and presentation via HLA class 1b, whereas in EC-m3 this was lower. These nonclassical HLAs are thought to play a role in immune escape and are associated with poor prognosis.

[0150] FIGS. **12**A, **12**B and **12**C provide plots showing dynamic BH3 profiling responses per CLL expression subtype.

[0151] FIG. **13** provides a schematic showing how drug sensitivity experiment data can be used to inform differential effects among expression clusters. Experimental data available includes that from 246 blood cancers, 184 CLLs, and 136 CLLs with RNA-seq. See, Dietrich, et al. “Drug-perturbation-based stratification of blood cancer,” JCI, 128:427-445 (2018), the disclosure of which is incorporated herein by reference in its entirety for all purposes.

[0152] FIG. **14** provides a heatmap showing median delta-priming for the indicated molecular features. Molecular features shown in FIG. **14** include IGHV subtypes, epitypes, expression subtypes (i.e., expression clusters), mutations in driver genes, and recurrent copy-number events. A feature was included in the heatmap of FIG. **14** only if at least 2 patients in a DBP screen had the feature. Median delta-priming was computed across all BH3 peptides and across all patients within

the feature.
[0153] FIG. **15** provides a plot comparing DBP z-scores for U-CLL and viability z-scores for U-CLL.
[0154] FIG. **16** provides a plot comparing DBP z-scores for M-CLL and viability z-scores for M-CLL.
[0155] FIG. **17** provides a heatmap showing median delta priming across healthy donors for the indicated normal cell types (CD14, CD19, and CD3).
[0156] FIG. **18** provides a heatmap showing values calculated by subtracting the median delta priming values for normal cell types ("Normals") from median delta-priming for the indicated molecular features. The "Normals" column in FIG. **18** is the median delta-priming value across all peptides, donors and sample types (CD19, CD3, CD14). All other columns contain median delta-priming values for tumors associated with the denoted molecular feature after subtracting the value in the Normals column.
[0157] FIG. **19** provides a comut plot showing molecular features for a group of n=65 patients.
[0158] FIG. **20** provides a comut plot showing molecular features for a group of n=81 patients.
[0159] FIGS. **21**A and **21**B provide a heatmap and dendrogram showing peptide effect similarity for multiple peptide concentrations. PUMA 1 μM, BIM 0.01 μM, BAD 0.3 μM and MSI 2.5 μM groups were used in several analyses as part of the examples disclosed herein.
[0160] FIGS. **22**A and **22**B provide a heatmap and plot comparing median delta priming for several drugs of interest.
[0161] FIG. **23** provides a plot showing differential drug sensitivity of several expression clusters of interest. Here, a published dataset of 136 CLL patients with RNA-seq whose samples were screened with 63 drugs was used. The expression cluster classifier was applied to the RNA-seqs and the data was used to show differential sensitivity of the ECs to these different drugs, such that patients with relevant mutations could be efficaciously treated by a drug for which sensitivity is high (for example, Venetoclax).
[0162] FIGS. **24**A, **24**B, **24**C and **24**D provide plots comparing drug sensitivity results in M-CLL and U-CLL groups by comparing, at the same concentration, the mean of the two closest (higher and lower) z-scored medians or z-scored means.
[0163] FIG. **25** provides a table identifying Venetoclax sensitivities for different driver alterations at different peptide concentrations. The table identifies if the priming response exhibited by each combination indicates that a patient with the associated driver alterations would likely respond favorably to treatment, or if they would be resistant to the drug being used.
[0164] FIG. **26** provides a table identifying kinase inhibitor drug sensitivities for different peptide concentrations and driver alterations.
[0165] FIG. **27** provides a table showing the relative efficacy of Abxinostat under conditions where patients are likely to be resistant to drugs such as Nutlin-3, MK-2206, and Zanubrutinib.
[0166] FIG. **28** provides a table showing that certain BCL2 inhibitors, such as Venetoclax, can exhibit similar priming responses when combined with peptides have a BCL2 inhibiting effect, which can assist in the identification of new CLL therapies.
[0167] FIG. **29** provides a schematic illustration of the anti-apoptotic proteins targeted by the different BH3 peptides used in DBP screening; because the different peptides could be relatively more or less selective for the anti-apoptotic proteins they targeted, use of each of the peptides provided different information with regard to a drug's impact on a CLL cell.
[0168] FIG. **30** provides a table showing that MCL1 inhibitors, approximated here with the presence of MS1 peptide, likely exhibit strong effects as part of combination therapies when used together.
[0169] FIG. **31** provides a clustered heatmap showing the level of association (Pearson correlation) of each molecular feature (IGHV, or epitype, or EC subtypes and drivers) with delta-priming across the CLL samples with 0.3 μM BAD when using delta priming as a categorical variable (high>10,

vs. low<5).
[0170] FIG. **32** provides a clustered heatmap showing the level of association (Pearson correlation) of each molecular feature (IGHV, or epitype, or EC subtypes and drivers) with delta-priming across the CLL samples with 0.01p M BIM when using delta priming as a categorical variable (high>10, vs. low<5).
[0171] FIG. **33** provides a clustered heatmap showing the level of association (Pearson correlation) of each molecular feature (IGHV, or epitype, or EC subtypes and drivers) with delta-priming across the CLL samples with 2.53 μM MS1 when using delta priming as a categorical variable (high>10, vs. low<5).
[0172] FIG. **34** provides a clustered heatmap showing the level of association (Pearson correlation) of each molecular feature (IGHV, or epitype, or EC subtypes and drivers) with delta-priming across the CLL samples with 1 μM PUMA when using delta priming as a categorical variable (high>10, vs. low<5).
[0173] FIG. **35** provides a clustered heatmap showing the level of association (Pearson correlation) of each molecular feature (IGHV, or epitype, or EC subtypes and drivers) with delta-priming across the CLL samples with 0.3 μM BAD as a continuous variable (using all values).
[0174] FIG. **36** provides a clustered heatmap showing the level of association (Pearson correlation) of each molecular feature (IGHV, or epitype, or EC subtypes and drivers) with delta-priming across the CLL samples with 0.01 μM BIM as a continuous variable (using all values).
[0175] FIG. **37** provides a clustered heatmap showing the level of association (Pearson correlation) of each molecular feature (IGHV, or epitype, or EC subtypes and drivers) with delta-priming across the CLL samples with 2.5p M MS1 as a continuous variable (using all values).
[0176] FIG. **38** provides a clustered heatmap showing the level of association (Pearson correlation) of each molecular feature (IGHV, or epitype, or EC subtypes and drivers) with delta-priming across the CLL samples with 1 μM PUMA as a continuous variable (using all values).
[0177] FIG. **39** provides a heatmap showing median delta-priming for the indicated molecular features.
[0178] FIG. **40** provides a heatmap showing median delta-priming across healthy donors per normal cell type.
[0179] FIG. **41** provides a heatmap showing median delta-priming for the indicated molecular features.

DETAILED DESCRIPTION OF THE INVENTION

[0180] The invention features, among other things, compositions, panels of biomarkers, and methods for selecting a subject with chronic lymphocytic leukemia (CLL) for treatment using an agent and/or for inclusion in a clinical trial using the agent to treat CLL. Also provided herein are methods and compositions for treatment prioritization, treatment sequencing, pharmacotyping, and/or drug repurposing for CLL.
[0181] The invention is based, at least in part, on the findings presented in the Examples provided herein based on a dynamic BH3 profiling (DBP) drug screen used to assess the relative sensitivity of many CLL patient samples to an array of drugs. These sensitivities of CLL samples were compared to normal B-cell samples to evaluate the extent to which the effect of each drug was specific to diseased cells. Of note, B-cells are non-essential to the survival of the subject and, therefore, drugs that effectively lead to apoptosis of both normal and leukemic B-cells should not be ruled out as potentially valid treatment options. Applying DBP to a large set of CLL samples, assisted by High-throughput DBP (HT-DBP) enabled pharmacotyping (i.e., identifying groups of samples that were responsive or unresponsive to one or more drug treatments). Pharmacotyping can be utilized for prognosis and diagnosis, in addition to treatment assignment.
[0182] In embodiments, CLL is characterized using eight chronic lymphocytic leukemia (CLL) gene expression subtypes and their efficacy in guiding prognosis and selection of subjects for a treatment. Not being bound by theory, the gene expression subtypes correspond to gene expression

clusters enriched with unique genetic and epigenetic features, distinguished by cellular pathways, and useful as an independent prognostic factor. A machine classifier was developed to classify a chronic lymphocytic leukemia (CLL) as belonging to a particular gene expression subtype associated with a corresponding gene expression cluster. The gene expression clusters and their corresponding expression subtypes are termed Ec-i, EC-m1, EC-m2, EC-m3, EC-m4, EC-o, EC-ul, and EC-u2. Said expression subtypes are known in the art and described, for example, by Knisbacher et al., Nat Genet. 2022 November; 54(11): 1664-1674., and in PCT/US2021/045144 (BI-10756), filed Aug. 9, 2021, each of which is incorporated herein by reference in its entirety. In embodiments, the gene expression subtype is used in combination with genetic drivers and epigenetic states in a model to assist in predicting sensitivity of a CLL to a drug. In embodiments, subjects with a CLL predicted to be sensitive to a particular drug are administered the drug as part of a treatment for the CLL.

[0183] Dynamic BH3 Profiling Dynamic BH3 profiling (DBP) is a drug screening assay that measures the relative priming of cells in a biological sample for cell death by apoptosis in the presence of a specific compound and/or a pro-apoptotic peptide. By “Dynamic BH3 profiling” is meant measuring drug-induced changes in mitochondrial apoptotic priming. Mitochondrial apoptotic priming is a measure of how close to the apoptotic threshold a cell is. A highly primed cell has relatively less anti-apoptotic binding site availability and is closer to the apoptotic threshold than a poorly primed cell, which has more anti-apoptotic availability to buffer an apoptotic assault and is further from the apoptotic threshold. BH3 peptides derived from the BH3 domain of pro-apoptotic BH3-only proteins provoke a response from viable mitochondria. Cytochrome c released from the mitochondria after a short incubation with BH3 peptide is used as a surrogate for priming. In general, the more sensitive a mitochondrion is to a BH3 peptide, the more primed it is. A drug treatment that enhances priming will cause mitochondria to undergo MOMP more easily when incubated with a fixed concentration of a promiscuously binding BH3 peptide, such as BIM BH3 peptide, compared to control-treated cells. See, for example, Potter, D. S. & Letai, A. To prime, or not to prime: that is the question. Cold Spring Harb. Symp. Quant. Biol. 81, 131-140 (2016); Montero, J. et al. Drug-induced death signaling strategy rapidly predicts cancer response to chemotherapy. Cell 160, 977-989 (2015); Daniels, V. W. et al. Metabolic perturbations sensitize triple-negative breast cancers to apoptosis induced by BH3 mimetics. Sci.

[0184] Signal 14, eabc7405 (2021); Bhola, P. D. et al. High-throughput dynamic BH3 profiling may quickly and accurately predict effective therapies in solid tumors. Sci. Signal 13, eaayl451 (2020); and Potter, D. S., Du, R., Bhola, P., Bueno, R. & Letai, A. Dynamic BH3 profiling identifies active BH3 mimetic combinations in non-small cell lung cancer. Cell Death Dis. 12, 741 (2021).

[0185] Dynamic BH3 profiling and peptides for use therein are described, for example, in Certo, et al., “Mitochondria primed by death signals determine cellular addiction to antiapoptotic BCL-2 family members,” Cancer Cell, 9:351-365 (2006); and in Foight, et al. “Designed BH3 Peptides with High Affinity and Specificity for Targeting Mcl-1 in Cells,” ACS Chemical Biology, 9:1962-1968 (2014), the disclosures of which are incorporated herein by reference in their entireties for all purposes. The DBP assay enables one to compare the response of a patient tumor sample to a specific drug relative to other drugs and relative to an inert control (e.g., DMSO). Similarly, DBP allows one to evaluate sensitivity of a leukemia sample to various pro-apoptotic peptides, which can be promiscuous activators of apoptosis (BIM and PUMA which both target e.g., BCL-2, BCL-XL and MCL-1) or selective activators of apoptosis (BAD which targets BCL-2 and BCL-XL; and MS1 which targets MCL-1). Drug and peptide combinations that are effective only or especially when administered at the same time can be detected as well.

Chronic Lymphocytic Leukemia (CLL)

[0186] Chronic lymphocytic leukemia (CLL) is a type of cancer in which the bone marrow makes too many lymphocytes. Early on there are typically no symptoms. Later, non-painful lymph node

swelling, feeling tired, fever, night sweats, or weight loss for no clear reason may occur. Enlargement of the spleen and low red blood cells (anemia) may also occur. It typically worsens gradually (i.e., “chronic”) over years.

[0187] Chronic lymphocytic leukemia (CLL) is a B cell neoplasm with variable natural history that is conventionally categorized into two major subtypes distinguished by the extent of somatic mutations in the heavy chain variable region of immunoglobulin genes (IGHV).

Selection of Subjects for Treatment

[0188] Panels comprising biomarkers of the invention are used to characterize chronic lymphocytic leukemia (CLL) in a subject to select the subject for treatment with an agent, for prognosis, and/or to characterize the CLL as belonging to an expression subtype (e.g., Ec-i, EC-m1, EC-m2, EC-m3, EC-m4, EC-o, EC-ul, and/or EC-u2). The panels of the invention are used in combination with a classification model, as described in the Examples provided herein, to categorize a chronic lymphocytic leukemia as belonging to an expression subtype selected from Ec-i, EC-m1, EC-m2, EC-m3, EC-m4, EC-o, EC-ul, and EC-u2. In certain embodiments, panels of the invention are used to select a treatment for the subject. In some embodiments, panels of the invention are used to select a subject for inclusion in a clinical study; for example, a subject is selected for treatment if the subject has a CLL of an expression subtype associated with a positive response to a drug being evaluated in the clinical study. In embodiments, the expression subtype is used as an input to an integrated model for predicting a clinical outcome for a subject having CLL. The integrated model can include as inputs, expression subtype, genetic drivers, and epigenetic states.

[0189] Non-limiting examples of genetic drivers include DNA mutations, copy number alterations, and/or structural variants in one or more of the following genes and/or genomic 10 regions: Genes: ADAMTS4, ANK], ARIDIA, ARID5B, ARPC4, ASXL1, ATM, BAI2, BAZ2A, BCOR, BIRC3, BRAF, BRCC3, CARD]], CCND2, CDC25B, CDCA7, CDKN]B, CENPB, CHD2, CHKB, CNOT3, CREB1, CREBBP, CUL9, DDX3X, DICER1, DIS3, DYRKIA, EEFIA1, EGR2, EWSR1, FAM50A, FAM65C, FBXW7, FUBPJ, GNBI, GPS2, GSR, IKBKB, IKZF3, INO80, IRF4, ITIH2, ITPKB, KLHL6, KMT2D, KRAS, MAP2K1, MAP2K2, MAP4K5, MAPK4, MBDI, MED], MED12, MGA, MSL3, MUM], MYD88, MYLK4, NCAPG, NEK8, NFKBI, NFKBIB, NFKBIE, NOTCH], NRAS, NSD], NXF], POLR3B, POT], PTPN]], RAF], RELA, RFX7, RPS]5, RPS16, RPS23, RRMI, RSC]A], RUFY], SAMHDI, SCN8A, SENP7, SETD2, SF3B1, SP140, SPEN, TFCP2, TP53, TRAF3, TRMT1, USP8, XPOI, ZC3H18, ZMYM3, and ZNF292; Genomic regions: I0p12.2, 10q21.3, 10q24.2, 10q24.32, 11q, 1113.4, 11q22.3, 12pl3.31, 12pl3.31, 13q14.13, 13q14.2, 13q14.3, 14q32.12, 14q32.33, 15q15.1, 15q24.2, 15q25.2, 15q26.1, 16p11.2, 16p11.2, 16pl3.3, 16q22.1, 17p, 17p11.2, 17p13.1, 17p13.3, 17q, 17q11.2, 17q21.32, 17q22, 17q23.1, 17q23.3, 17q25.1, 18p, 18q11.2, 18q21.2, 19p, 19p13.11, 19pl3.12, 19pl3.3, 19q, 19q13.33, lp31.3, lp35.2, 1p36.11, p360.21, 1921.3, 1q22, 1923.2, 1932.2, 1942.12, 1942.13, 20p, 20p11.22, 22q12.1, 22q13.2, 2p, 2p11.2, 2p13.3, 2p15, 2p23.3, 2q12.2, 2q13, 2q31.1, 3p, 3p13, 3p21.31, 3p22.2, 3p22.3, 4p, 4q35.1, 5pl5.33, 5q32, 5q35.3, 6p21.32, 6p22.1, 6q, 6q21, 6q25.3, 7p22.1, 7p22.2, 7q11.23, 7q22.1, 7q36.1, 8p, 8p11.23, 8q, 8q12.1, 8q22.1, 9p21.3, and 9q34.3, 12 (including trisomy of chromosome 12).

[0190] In some embodiments, results from a dynamic BH3 profiling (DBP) and/or high-throughput DBP (HT-DBP) screen, as described in the Examples provided herein, can be compared to existing or future DBP and/or HT-DBP screens to assign a subject's CLL to a specific pharmacotype and to prioritize treatment.

[0191] In some embodiments, a specific treatment plan is advised or disadvised for subjects with a specific subtype of CLL. Subtypes include but are not limited to IGHV-mutated CLL (M-CLL), IGHV-unmutated CLL (U-CLL), methylation subtypes of CLL [CLLs that resemble naive B-cells (n-CLL), intermediate methylation state CLLs (i-CLL) and/or CLLs that resemble memory B-cells (m-CLL)], RNA expression subtypes (EC-m1, EC-m2, EC-m3, EC-m4, EC-ul, EC-u2, EC-o, EC-i, and/or the more general EC-m and/or EC-u).

[0192] In some embodiments, treatments are recommended for subjects whose CLL has a specific genetic mutation, genetic copy-number alteration and/or a genetic structural variation that is associated with a response or lack of response to one or more specific drugs or drug classes. These alterations can arise in the germline or somatically in the leukemic cells or the leukemia's precursor cells.
[0193] In some embodiments, the method comprises determining whether a subject sample (e.g., a CLL sample) will or will not respond to a specific drug or drug class. In certain embodiments, the drugs comprise Abexinostat, Acalabrutinib, Azacitidine, AZD8055, Carfilzomib, Cerdulatinib, Crizotinib, Dasatinib, Duvelisib, Entospletinib, Erastin, Fludarabine, Gandotinib, GSK690693, Idelalisib, JQ1, Lenalidomide, MK2206, Navitoclax, Nirogacestat, Nutlin-3, Osimertinib, Ponatinib, Rapamycin, Ricolinostat, Ruxolitinib, Selinexor, Sorafenib, Sunitinib, Umbralisib, Vecabrutinib, Vorinostat, A-1331852, atorvastatin, AZD5991, Bendamustine, Onalespib, Trametinib, Voruciclib, Zanubrutinib, Ibrutinib, and/or Venetoclax.
[0194] In some embodiments, response or resistance to these drugs extends to drug classes they represent and/or to other drugs that target the same molecules, processes and/or biological pathways, including for example those listed in Table 1A or Table 1B.
[0195] In some embodiments, the methods further comprise obtaining the sample (e.g., the cancer sample) from a subject. In certain embodiments, the method further comprises treating CLL and/or CLL subtypes U-CLL, M-CLL, n-CLL, i-CLL, m-CLL, EC-m1, EC-m2, EC-m3, EC-m4, EC-ul, EC-u2, EC-o, EC-i) in the subject by administering cancer therapy to the subject (e.g., a chemotherapy, a radiation therapy, an immunotherapy).
[0196] In some embodiments the method comprises administering a combination of drugs concurrently or defining a set of drugs to administer sequentially. In embodiments, the combination of drugs comprises two or more drugs listed in Table 1A or 1B. In some instances, the method comprises administering venetoclax to a subject in combination with an MCL1 inhibitor (e.g., AZD5991). In some cases, the method comprises administering iburtinib to a subject in combination with a BCL2 inhibitor (e.g., venetoclax). In embodiments, the BCL2 inhibitor comprises venetoclax, ZN-d5 (Zentalis), lisaftoclax (APG-2575, Ascentage), S55746 (Servier/Novartis), and/or AZD4320 (Astra-Zeneca). In some cases, the MCL1 inhibitor comprises AZD5991, tapotoclax (AMG-176, AMGEN), MIK665 (Servier/Novartis), A-1210477 (AbbVie), ANJ810 (Anji Oncology), PRT1419 (Prelude Therapeutics), AS00491, APG-3526 (Ascentage Pharma), CT-03, and//or CPT-6281 (Captor Therapeutics).
[0197] In some embodiments, DBP and/or HT-DBP is applied to determine the pharmacotype of a subject, which can be used for designing a treatment plan, prognosis and/or diagnosis as CLL or a molecular subtype of CLL.
[0198] In some embodiments, DNA sequencing, RNA sequencing, DNA methylation assays (e.g., reduced-representation bisulfite sequencing, methylation arrays, whole-genome bisulfite sequencing, targeted bisulfite sequencing) and/or proteomics are applied to a sample from a subject to recommend or unrecommend treatment with one or more of the aforementioned drugs.
[0199] The invention provides methods for using the expression subtype of a chronic lymphocytic leukemia (CLL) to predict the sensitivity or resistance of a CLL to a drug. The invention further provides methods for selecting a subject with chronic lymphocytic leukemia (CLL) for treatment with a drug to which the CLL is predicted to be sensitive. The invention also provides methods for selecting subjects having chronic lymphocytic leukemia for inclusion in a clinical trial or other drug study where subjects with CLL predicted to be sensitive to a drug being studied in the trial or study are included in the trial or study and/or subjects with CLL predicted to be resistant to the drug are excluded from the trial or study.
[0200] Based on their expression subtype, subjects are selected for treatment with one or more of the agents listed in Table 1A or 1B.
[0201] In some embodiments, a subject having a CLL with a particular expression subtype is

selected for treatment with an agent targeting a gene or polypeptide associated with the expression subtype. In various embodiments, the association of a gene or polypeptide with an expression subtype is determined according to the associations (e.g., increase or decrease in expression levels) indicated in Table 3A.

[0202] In some embodiments, a subject having a CLL determined to have a driver mutation, is administered an agent targeting the gene and/or a product of the gene (e.g., an agent reducing expression or activity of the gene and/or polypeptide). In embodiments, the drug sensitivity and drug resistance information provided in FIGS. **12**A-**12**C relating to particular drugs and expression subtypes can be extrapolated to apply to those drugs having a similar or the same main target, and/or the same target category (A) or (B) as a drug listed in.

[0203] The correlation of test results with an expression subtype involves applying a classification algorithm (e.g., a machine learning classifier) of some kind to the results to determine the expression subtype. The classification algorithm may be as simple as determining whether or not the amounts of the markers are above or below a particular cut-off number. When multiple biomarkers are used, the classification algorithm may be a linear regression formula.

[0204] Alternatively, the classification algorithm may be the product of any of a number of learning algorithms described herein.

[0205] In the case of complex classification algorithms, it may be necessary to perform the algorithm on the data, thereby determining the expression subtype using a computer, e.g., a programmable digital computer. In either case, one can then record the status on tangible medium, for example, in computer-readable format such as a memory drive or disk or simply printed on paper. The result also could be reported on a computer screen.

[0206] Panels The present disclosure provides panels of biomarkers and the use of such panels for characterizing chronic lymphocytic leukemia (CLL). As would be understood, references herein to a biomarker, a panel of biomarkers, or other similar phrase indicates one or more of the biomarkers listed below, in Tables 3A and 4, or otherwise described herein.

[0207] In one embodiment, markers useful in the panels of the invention include, for example, ABCA9, ACAP3, ACSM3, ADAP2, AF127936.7, ARHGAP33, ARMC7, ARRDC5, ARSD, ARSI, ASB2, ATP1A3, ATP2B1, ATPIF1, BASP1, BCL2A1, BCL7A, BCS1L, CAMK2A, CLDN23, CMTM7, COBLL1, CRELD2, CRY1, CTAGE9, CTLA4, DDR1, DKFZP761J1410, DPF3, EML6, ERRFIl, ESPNL, EZH2, FAHD2B, FAM109A, FBXO27, FGL2, FLJ20373, FMOD, GADD45A, GNAO1, GPR160, GPR34, GUCD1, HCK, HDAC4, HIP1R, HMCES, IGSF3, IQSEC1, ITGAX, KCNH3, KCNN3, KCTD3, KDM1B, KLK1, KSR1, LCN10, LINC00865, LPL, LRRK2, LUZP1, MAP4K4, MAPK4, MAST4, MPRIP, MRO, MSI2, MVB12B, MYBL1, MYC, MYL5, MYL9, MYO3A, NEDD9, NFKBIZ, NR2F6, NRIP1, NRSN2, NUGGC, P2RX1, *PELI*3, PIGB, PIP5K1B, PITPNC1, PLD1, PTPN7, QDPR, REPS2, RHBDF2, RIMKLB, RP11-134N1.2, RP11-265P11.1, RP11-453F18__B.1, RP11-456H18.2, RP1-90J20.12, SAMSN1, SCPEP1, SH3D21, SLC44A1, SLC4A7, SLC4A8, SMIM10, SPN, SSBP3, STAM, STX5, SYNGR3, TAS1R3, TBC1D2B, TBC1D9, TFEC, TIMELESS, TNFRSF13B, TNR, TOX2, TRIM7, TUBG2, VSIG10, WNT5A, ZMYND8, and ZNF804A, or a sub-set thereof, as well as the nucleic acid molecules encoding such proteins. In another embodiment, markers useful in the panels of the invention include, for example, ACAP3, ACSM3, AEBP1, AKT3, ARHGAP33, ARHGAP42, ARMC7, ARRDC5, ATPIF1, BACH2, BASP1, BCL7A, C17orfl00, CBLB, CD72, CD86, CEACAMI, CHPT1, CLDN7, CMTM7, CNTNAP1, COBLL1, COL18A1, CRY1, CTLA4, EGR3, EML6, EZH2, FADS3, FCER1G, FCRL2, FGL2, FLJ20373, FMOD, GADD45A, GLIPRI, GNB4, GPR160, GPR34, GRIK3, GUCD1, HCK, HIP1R, HIVEP3, HMCES, IGF2BP3, IGSF3, IL21R, INPP5F, IQGAP2, IQSEC1, ITGAX, ITGB5, JDP2, KANK2, KCNH2, KDM1B, KLF3, LATS2, LCN10, LEF1, LPL, LRRK2, LUZP1, MAP4K4, MID1IPI, MMP14, MPRIP, MSI2, MYBL1, MYL9, MYLIP, MZB1, NBPF3, NRIP1, NRSN2, NUGGC, NXPH4, P2RX1, P2RX5, P2RY14, PDGFD, PIP5K1B, PITPNC1, PON2, PRICKLE1, PTPN7, RCN3, RDX, RHBDF2, RIMKLB, RNF135,

RP11-145M9.4, RP11-268J15.5, RP11-463012.3, RP5-1028K7.2, SAMSN1, SCCPDH, SCD, SCPEP1, SDC3, SECTM1, SESN3, SH3BP2, SH3D21, SLC16A5, SLC19A1, SLC4A7, SPN, SSBP3, STX5, SUSD1, TBC1D2B, TBC1D9, TBKBP1, TCF7, TFEC, TGFBR3, TIGIT, TIMELESS, TMEM133, TNFRSF13B, TOX2, TRAK2, TTC39C, TUBG2, VPS37B, VSIG10, WNT9A, ZAP70, ZNF667-AS1, ZNF804A, and ZSWIM6, or a subset thereof, as well as the nucleic acid molecules encoding such proteins. Fragments of the aforementioned polypeptides useful in the methods of the invention are sufficient to bind an antibody that specifically recognizes the protein from which the fragment is derived.

[0208] In some instances, markers useful for the panels of the invention include markers for U-CLL, namely XPO1, BCOR, KRAS, RPS23, RRM1, RAF1, MAP2K2, LRP1B, or a subset thereof, as well as the nucleic acid molecules encoding such proteins. In some cases, markers useful for the panels of the invention include markers for M-CLL, namely MYD88, KLHL6, ITPKB, TCL1A, DICER1, or a subset thereof, as well as the nucleic acid molecules encoding such proteins.

[0209] In embodiments, markers useful in the panels of the invention include markers for expression cluster Ec-i, namely, GRIK3, IQGAP2, FCER1G, STK32B, GADD45A, ITGAX, KLF3, RFTN1, PTK2, DFNB31, and ZMAT1, or a sub-set thereof, as well as the nucleic acid molecules encoding such proteins.

[0210] In embodiments, markers useful in the panels of the invention include markers for expression cluster EC-m1, namely, TFEC, COL18A1, SLC19A1, NRIP1, KCNH2, P2RX1, ARRDC5, BEX4, and APP, or a sub-set thereof, as well as the nucleic acid molecules encoding such proteins.

[0211] In embodiments, markers useful in the panels of the invention include markers for expression cluster EC-m2, namely, EML6, HCK, CD1C, VPS37B, CYBB, NXPH4, BTNL9, KLRK1, IQSEC1, BANKI, LEF1, SH3D21, FMOD, SEMA4A, CTLA4, ADTRP, IGSF3, IGFBP4, PDGFD, and APOD, or a sub-set thereof, as well as the nucleic acid molecules encoding such proteins.

[0212] In embodiments, markers useful in the panels of the invention include markers for expression cluster EC-m3, namely, MS4A4E, MYL9, NT5E, MS4A6A, PITPNC1, CNTNAP2, 5 IGF2BP3, WNT3, CLDN7, TCF7, BASP1, FLJ20373, MAP4K4, LRRK2, SAMSN1, CEACAM1, TNFRSF13B, PHF16, MID1IP1, and ABCA9, or a sub-set thereof, as well as the nucleic acid molecules encoding such proteins.

[0213] In embodiments, markers useful in the panels of the invention include markers for expression cluster EC-m4, namely, MYBL1, NUGGC, GNG8, AEBP1, HIP1R, LATS2, RIMKLB, EML6, FADS3, MBOAT1, LCN10, DCLK2, and GLUL, or a sub-set thereof, as well as the nucleic acid molecules encoding such proteins.

[0214] In embodiments, markers useful in the panels of the invention include markers for expression cluster EC-o, namely, ACSM3, TOX2, PHF16, SESN3, TBC1D9, PIP5KIB, SIK1, DUSP5, GNG7, HIVEP3, MARCKSL1, GPR183, HRK, and PITPNC1, or a sub-set thereof, as well as the nucleic acid molecules encoding such proteins.

[0215] In embodiments, markers useful in the panels of the invention include markers for expression cluster EC-ul, namely, SEPT10, LDOC1, LPL, KANK2, SOWAHC, DUSP26, OSBPL5, WNT9A, FGFR1, GTSF1L, ADD3, AKT3, COBLL1, MNDA, FCRL3, FAM49A, FCRL2, SLC2A3, and MARCKS, or a sub-set thereof, as well as the nucleic acid molecules encoding such proteins.

[0216] In embodiments, markers useful in the panels of the invention include markers for expression cluster EC-u2, namely, ITGB5, BCL7A, PPP1R9A, TSPAN13, SLC12A7, SSBP3, VASH1, SPG20, ILT3RA1, NR3C2, TUBG2, ZNF804A, and IL2RA, or a sub-set thereof, as well as the nucleic acid molecules encoding such proteins. The panels can comprise biomarkers for expression cluster Ec-i, EC-m1, EC-m2, EC-m3, EC-m4, EC-o, EC-ul, or EC-u2, or various combinations thereof.

[0217] The invention further features the use of such panels for characterizing chronic lymphocytic leukemia (CLL). In embodiments, the panels are used in combination with a classifier (e.g., a machine learning classifier) to identify a CLL as belonging to a particular expression subtype. The panels are advantageously used for guiding selection of a subject for a CLL treatment.

Biomarkers

[0218] Measurements of expression levels of biomarkers (e.g., polypeptide and/or polynucleotides encoding polypeptides present in expression clusters described herein) are used in combination with a model (e.g., a machine learning classifier) to identify a chronic lymphocytic leukemia as belonging to a particular expression subtype. In particular embodiments, a biomarker is an organic biomolecule that is differentially present in a sample taken from a subject of one phenotypic status (e.g., having a disease, such as chronic lymphocytic leukemia (CLL)) as compared with another phenotypic status (e.g., not having the disease). A biomarker is differentially present between different phenotypic statuses if the mean or median expression level of the biomarker in the different groups is calculated to be statistically significant. Common tests for statistical significance include, among others, t-test, ANOVA, Kruskal-Wallis, Wilcoxon, Mann-Whitney and odds ratio. Biomarkers, alone or in combination, provide measures of relative risk that a subject belongs to one phenotypic status or another. Therefore, they are useful as markers for characterizing a disease (e.g., chronic lymphocytic leukemia (CLL)).

[0219] A biomarker of the invention may be detected in a biological sample of the subject (e.g., tissue, fluid), including, but not limited to blood, blood serum, plasma, saliva, urine, ascites, cyst fluid, a homogenized tissue sample (e.g., a tissue sample obtained by biopsy), a cell isolated from a patient sample, and the like.

[0220] The disclosure provides panels comprising isolated biomarkers. The biomarkers can be isolated from biological fluids. They can be isolated by any method known in the art. In certain embodiments, this isolation is accomplished using the mass and/or binding characteristics of the markers. For example, a sample comprising the biomolecules can be subject to chromatographic fractionation and subject to further separation by, e.g., acrylamide gel electrophoresis.

[0221] Knowledge of the identity of the biomarker also allows their isolation by immunoaffinity chromatography. In some embodiments, biomarkers described herein are fixed to a substrate (e.g., chips, beads, microfluidic platforms, membranes).

Detection of Biomarkers

[0222] The biomarkers of this disclosure can be detected by any suitable method. The methods described herein can be used individually or in combination for a more accurate detection of the biomarkers (e.g., biochip in combination with mass spectrometry, immunoassay in combination with mass spectrometry, and the like).

[0223] Detection paradigms that can be employed in the disclosure include, but are not limited to, optical methods, electrochemical methods (voltammetry and amperometry techniques), atomic force microscopy, and radio frequency methods, e.g., multipolar resonance spectroscopy.

[0224] Illustrative of optical methods, in addition to microscopy, both confocal and non-confocal, are detection of fluorescence, luminescence, chemiluminescence, absorbance, reflectance, transmittance, and birefringence or refractive index (e.g., surface plasmon resonance, ellipsometry, a resonant mirror method, a grating coupler waveguide method or interferometry).

[0225] These and additional methods are described below.

Detection by sequencing and/or probes

[0226] In particular embodiments, the biomarkers of the invention are measured by a sequencing- and/or probe-based technique (e.g., RNA-seq).

[0227] RNA sequencing (RNA-Seq) is a powerful tool for transcriptome profiling. In embodiments, to mitigate sequence-dependent bias resulting from amplification complications to allow truly digital RNA-Seq, a set of barcode sequences can be used to ensure that every cDNA molecule prepared from an mRNA sample is uniquely labeled by random attachment of barcode

sequences to both ends (see, e.g., Shiroguchi K, et al. Proc Natl Acad Sci USA. 2012 Jan. 24;109(4):1347-52). After PCR, paired-end deep sequencing can be applied to read the two barcodes and cDNA sequences. Rather than counting the number of reads, RNA abundance can be measured based on the number of unique barcode sequences observed for a given cDNA sequence. The barcodes may be optimized to be unambiguously identifiable. This method is a representative example of how to quantify a whole transcriptome from a sample.

[0228] Detecting a target polynucleotide sequence or fragment thereof associated with a biomarker that hybridizes to a probe sequence may involve sequencing, FACS, qPCR, RT-PCR, a genotyping array, and/or a NanoString assay (see, e.g., Malkov, et al. "Multiplexed measurements of gene signatures in different analytes using the Nanostring nCounter™ Assay System", BMC Research Notes, 2: Article No: 80 (2009)), or any of various other techniques known to one of skill in the art. Various detection methods may be used and are described as follows.

[0229] Preparation of a library for sequencing may involve an amplification step. Amplification may involve thermocycling or isothermal amplification (such as through the methods RPA or LAMP). Cross-linking may involve overlap-extension PCR or use of ligase to associate multiple amplification products with each other. Amplification can refer to any method employing a primer and a polymerase capable of replicating a target sequence with reasonable fidelity. Amplification may be carried out by natural or recombinant DNA polymerases such as TaqGold™, T7 DNA polymerase, Klenow fragment of *E. coli* DNA polymerase, and reverse transcriptase. A preferred amplification method is PCR. In particular, the isolated RNA can be subjected to a reverse transcription assay that is coupled with a quantitative polymerase chain reaction (RT-PCR) in order to quantify the expression level of a biomarker.

[0230] Detection of the expression level of a biomarker can be conducted in real time in an amplification assay (e.g., qPCR). In one aspect, the amplified products can be directly visualized with fluorescent DNA-binding agents including but not limited to DNA intercalators and DNA groove binders. Because the amount of the intercalators incorporated into the double-stranded DNA molecules is typically proportional to the amount of the amplified DNA products, one can conveniently determine the amount of the amplified products by quantifying the fluorescence of the intercalated dye using conventional optical systems in the art. DNA-binding dyes suitable for this application include, as non-limiting examples, SYBR green, SYBR blue, DAPI, propidium iodine, Hoeste, SYBR gold, ethidium bromide, acridines, proflavine, acridine orange, acriflavine, fluorcoumanin, ellipticine, daunomycin, chloroquine, distamycin D, chromomycin, homidium, mithramycin, ruthenium polypyridyls, anthramycin, and the like.

[0231] Other fluorescent labels such as sequence specific probes can be employed in the amplification reaction to facilitate the detection and quantification of the amplified products. Probe-based quantitative amplification relies on the sequence-specific detection of a desired amplified product. It utilizes fluorescent, target-specific probes (e.g., TaqMan® probes) resulting in increased specificity and sensitivity. Methods for performing probe-based quantitative amplification are taught, for example, in U.S. Pat. No. 5,210,015.

[0232] Sequencing may be performed on any high-throughput platform. Methods of sequencing oligonucleotides and nucleic acids are well known in the art (see, e.g., WO93/23564, WO98/28440 and WO98/13523; U.S. Pat. App. Pub. No. 2019/0078232; U.S. Pat. Nos. 5,525,464; 5,202,231; 5,695,940; 4,971,903; 5,902,723; 5,795,782; 5,547,839 and 5,403,708; Sanger et al., Proc. Natl. Acad. Sci. USA 74:5463 (1977); Drmanac et al., Genomics 4:114 (1989); Koster et al., Nature Biotechnology 14:1123 (1996); Hyman, Anal. Biochem. 174:423 (1988); Rosenthal, International Patent Application Publication 761107 (1989); Metzker et al., Nucl. Acids Res. 22:4259 (1994); Jones, Biotechniques 22:938 (1997); Ronaghi et al., Anal. Biochem. 242:84 (1996); Ronaghi et al., Science 281:363 (1998); Nyren et al., Anal. Biochem. 151:504 (1985); Canard and Arzumanov, Gene 11:1 (1994); Dyatkina and Arzumanov, Nucleic Acids Symp Ser 18:117 (1987); Johnson et al., Anal. Biochem. 136:192 (1984); and Elgen and Rigler, Proc. Natl. Acad. Sci. USA 91(13):5740

(1994), all of which are expressly incorporated by reference herein in their entirety).
[0233] The sequencing of a polynucleotide can be carried out using any suitable commercially available sequencing technology. In embodiments, the sequencing of a polynucleotide is carried out using a chain termination method of DNA sequencing (e.g., Sanger sequencing). In some embodiments, commercially available sequencing technology is a next-generation sequencing technology, including as non-limiting examples combinatorial probe anchor synthesis (cPAS), DNA nanoball sequencing, droplet-based or digital microfluidics, heliscope single molecule sequencing, nanopore sequencing (e.g., Oxford Nanopore technologies), GeneGap sequencing, massively parallel signature sequencing (MPSS), microfluidic Sanger sequencing, microscopy-based techniques (e.g., transmission electronic microscopy DNA sequencing), RNA polymerase (RNAP) sequencing, single-molecule real-time (SMRT) sequencing, SOLiD sequencing, ion semiconductor sequencing, polony sequencing, Pyrosequencing (454), sequencing by hybridization, sequencing by synthesis (e.g., Illumina™ sequencing), sequencing with mass spectrometry, and tunneling currents DNA sequencing.
[0234] In embodiments, levels of biomarkers in a sample are quantified using targeted sequencing. Methods for targeted sequencing are well known in the art (see, e.g., Rehm, “Disease-targeted sequencing: a cornerstone in the clinic”, Nature Reviews Genetics, 14:295-300 (2013)).
[0235] In embodiments, a probe comprises a molecular identifier, such as a fluorescent or chemiluminescent label, a radioactive isotope label, an enzymatic ligand, or the like. The molecular identifier can be a fluorescent label or an enzyme tag, such as digoxigenin, (3-galactosidase, urease, alkaline phosphatase or peroxidase, avidin/biotin complex.
[0236] Methods used to detect or quantify binding of a probe to a target biomarker will typically depend upon the molecular identifier. For example, radiolabels may be detected using photographic film or a phosphoimager. Fluorescent markers may be detected and quantified using a photodetector to detect emitted light. Enzymatic labels can be detected by providing the enzyme with a substrate and measuring the reaction product produced by the action of the enzyme on the substrate; and colorimetric labels can be detected by visualizing a colored label.
[0237] Specific non-limiting examples of molecular identifiers include radioisotopes, such as 32P, 14C, 1251, 3H, and 1311, fluorescein, rhodamine, dansyl chloride, umbelliferone, luciferase, peroxidase, alkaline phosphatase, P-galactosidase, P-glucosidase, horseradish peroxidase, glucoamylase, lysozyme, saccharide oxidase, microperoxidase, biotin, and ruthenium. In the case where biotin is employed as a molecular identifier, streptavidin bound to an enzyme (e.g., peroxidase) may further be added to facilitate detection of the biotin.
[0238] Examples of fluorescent molecular identifiers include, but are not limited to, Atto dyes, 4-acetamido-4′-isothiocyanatostilbene-2,2′ disulfonic acid; acridine and derivatives: acridine, acridine isothiocyanate; 5-(2′-aminoethyl)aminonaphthalene-1-sulfonic acid (EDANS); 4-amino-N-[3-vinyl sulfonyl)phenyl]naphthalimide-3,5 disulfonate; N-(4-anilino-1-naphthyl)maleimide; anthranilamide; BODIPY; Brilliant Yellow; coumarin and derivatives; coumarin, 7-amino-4-methylcoumarin (AMC, Coumarin 120), 7-amino-4-trifluoromethylcouluarin (Coumaran 151); cyanine dyes; cyanosine; 4′,6-diaminidino-2-phenylindole (DAPI); 5′5”-dibromopyrogallol-sulfonaphthalein (Bromopyrogallol Red); 7-diethylamino-3-(4′-isothiocyanatophenyl)-4-methylcoumarin; diethylenetriamine pentaacetate; 4,4′-diisothiocyanatodihydro-stilbene-2,2′-disulfonic acid; 4,4′-diisothiocyanatostilbene-2,2′-disulfonic acid; 5-[dimethylamino]naphthalene-1-sulfonyl chloride (DNS, dansylchloride); 4-dimethylaminophenylazophenyl-4′-isothiocyanate (DABITC); eosin and derivatives; eosin, eosin isothiocyanate, erythrosin and derivatives; erythrosin B, erythrosin, isothiocyanate; ethidium; fluorescein and derivatives; 5-carboxyfluorescein (FAM), 5-(4,6-dichlorotriazin-2-yl)aminofluorescein (DTAF), 2′,7′-dimethoxy-4′5′-dichloro-6-carboxyfluorescein, fluorescein, fluorescein isothiocyanate, QFITC, (XRITC); fluorescamine; IR144; IR1446; Malachite Green isothiocyanate; 4-methylumbelliferoneortho cresolphthalein; nitrotyrosine; pararosaniline; Phenol Red; B-phycoerythrin; o-phthaldialdehyde;

pyrene and derivatives: pyrene, pyrene butyrate, succinimidyl 1-pyrene; butyrate quantum dots; Reactive Red 4 (Cibacron™ Brilliant Red 3B-A) rhodamine and derivatives: 6-carboxy-X-rhodamine (ROX), 6-carboxyrhodamine (R6G), lissamine rhodamine B sulfonyl chloride rhodamine (Rhod), rhodamine B, rhodamine 123, rhodamine X isothiocyanate, sulforhodamine B, sulforhodamine 101, sulfonyl chloride derivative of sulforhodamine 101 (Texas Red); N,N,N′,N′ tetramethyl-6-carboxyrhodamine (TAMRA); tetramethyl rhodamine; tetramethyl rhodamine isothiocyanate (TRITC); riboflavin; rosolic acid; terbium chelate derivatives; Cy3; Cy5; Cy5.5; Cy7; IRD 700; IRD 800; La Jolta Blue; phthalo cyanine; and naphthalo cyanine.

[0239] A fluorescent molecular identifier may be a fluorescent protein, such as blue fluorescent protein, cyan fluorescent protein, green fluorescent protein, red fluorescent protein, yellow fluorescent protein or any photoconvertible protein. Colorimetric molecular identifiers, bioluminescent molecular identifiers and/or chemiluminescent molecular identifiers may be used in embodiments of the disclosure.

[0240] Detection of a molecular identifier may involve detecting energy transfer between molecules in a hybridization complex by perturbation analysis, quenching, or electron transport between donor and acceptor molecules, the latter of which may be facilitated by double stranded match hybridization complexes. The fluorescent molecular identifier may be a perylene or a terrylen. In the alternative, the fluorescent molecular identifier may be a fluorescent bar code.

[0241] The molecular identifier may be light sensitive, wherein the label is light-activated and/or light cleaves the one or more linkers to release the molecular cargo. The light-activated molecular cargo may be a major light-harvesting complex (LHCII). In another embodiment, the fluorescent molecular label may induce free radical formation.

[0242] In an advantageous embodiment, agents may be uniquely labeled in a dynamic manner (see, e.g., international patent application serial no. PCT/US2013/61182 filed Sep. 23, 2012).

[0243] The unique labels are, at least in part, nucleic acid in nature, and may be generated by sequentially attaching two or more detectable oligonucleotide tags to each other and each unique label may be associated with a separate agent. A detectable oligonucleotide tag may be an oligonucleotide that may be detected by sequencing of its nucleotide sequence and/or by detecting non-nucleic acid detectable moieties to which it may be attached.

[0244] In embodiments, the molecular identifier is a microparticle, including, as non-limiting examples, quantum dots (Empodocles, et al., Nature 399:126-130, 1999), or gold nanoparticles (Reichert et al., Anal. Chem. 72:6025-6029, 2000).

Detection by Immunoassay

[0245] In particular embodiments, the biomarkers of the invention are measured by immunoassay. An immunoassay typically utilizes an antibody (or other agent that specifically binds the marker) to detect the presence or level of a biomarker in a sample. Antibodies can be produced by methods well known in the art, e.g., by immunizing animals with the biomarkers. Biomarkers can be isolated from samples based on their binding characteristics. Alternatively, if the amino acid sequence of a polypeptide biomarker is known, the polypeptide can be synthesized and used to generate antibodies by methods well known in the art.

[0246] This disclosure contemplates traditional immunoassays including, for example, Western blot, sandwich immunoassays including ELISA and other enzyme immunoassays, fluorescence-based immunoassays, and chemiluminescence. Nephelometry is an assay done in liquid phase, in which antibodies are in solution. Binding of the antigen to the antibody results in changes in absorbance, which is measured. Other forms of immunoassay include magnetic immunoassay, radioimmunoassay, and real-time immunoquantitative PCR (iqPCR).

[0247] Immunoassays can be carried out on solid substrates (e.g., chips, beads, microfluidic platforms, membranes) or on any other forms that supports binding of the antibody to the marker and subsequent detection. A single marker may be detected at a time or a multiplex format may be used. Multiplex immunoanalysis may involve planar microarrays (protein chips) and bead-based

microarrays (suspension arrays).
[0248] In a SELDI-based immunoassay, a biospecific capture reagent for the biomarker is attached to the surface of an MS probe, such as a pre-activated ProteinChip array. The biomarker is then specifically captured on the biochip through this reagent, and the captured biomarker is detected by mass spectrometry.
Detection by Biochip
[0249] In embodiments, a sample is analyzed by means of a biochip (also known as a microarray). The polypeptides and nucleic acid molecules of the disclosure are useful as hybridizable array elements in a biochip. Biochips generally comprise solid substrates and have a generally planar surface, to which a capture reagent (also called an adsorbent or affinity reagent) is attached. Frequently, the surface of a biochip comprises a plurality of addressable locations, each of which has the capture reagent bound there.
[0250] The array elements are organized in an ordered fashion such that each element is present at a specified location on the substrate. Useful substrate materials include membranes, composed of paper, nylon or other materials, filters, chips, glass slides, and other solid supports. The ordered arrangement of the array elements allows hybridization patterns and intensities to be interpreted as expression levels of particular genes or proteins. Methods for making nucleic acid microarrays are known to the skilled artisan and are described, for example, in U.S. Pat. No. 5,837,832, Lockhart, et al. (Nat. Biotech. 14:1675-1680, 1996), and Schena, et al. (Proc. Natl. Acad. Sci. 93:10614-10619, 1996), herein incorporated by reference in its entrety. Methods for making polypeptide microarrays are described, for example, by Ge (Nucleic Acids Res. 28: e3. i-e3. vii, 2000), MacBeath et al., (Science 289:1760-1763, 2000), Zhu et al. (Nature Genet. 26:283-289), and in U.S. Pat. No. 6,436,665, hereby incorporated by reference in its entirety.
Detection by Protein Biochip
[0251] In embodiments, a sample is analyzed by means of a protein biochip (also known as a protein microarray). Such biochips are useful in high-throughput low-cost screens to identify alterations in the expression or post-translation modification of a biomarker, or a fragment thereof. In embodiments, a protein biochip of the disclosure binds a biomarker present in a sample and detects an alteration in the level of the biomarker. Typically, a protein biochip features a protein, or fragment thereof, bound to a solid support. Suitable solid supports include membranes (e.g., membranes composed of nitrocellulose, paper, or other material), polymer-based films (e.g., polystyrene), beads, or glass slides. For some applications, proteins (e.g., antibodies that bind a marker of the disclosure) are spotted on a substrate using any convenient method known to the skilled artisan (e.g., by hand or by inkjet printer).
[0252] In embodiments, the protein biochip is hybridized with a detectable probe. Such probes can be polypeptides, nucleic acid molecules, antibodies, or small molecules. For some applications, polypeptide and nucleic acid molecule probes are derived from a biological sample taken from a patient, such as a bodily fluid (such as blood, blood serum, plasma, saliva, urine, ascites, cyst fluid, and the like); a homogenized tissue sample (e.g., a tissue sample obtained by biopsy); or a cell isolated from a patient sample. Probes can also include antibodies, candidate peptides, nucleic acids, or small molecule compounds derived from a peptide, nucleic acid, or chemical library. Hybridization conditions (e.g., temperature, pH, protein concentration, and ionic strength) are optimized to promote specific interactions. Such conditions are known to the skilled artisan and are described, for example, in Harlow, E. and Lane, D., Using Antibodies: A Laboratory Manual. 1998, New York: Cold Spring Harbor Laboratories. After removal of non-specific probes, specifically bound probes are detected, for example, by fluorescence, enzyme activity (e.g., an enzyme-linked calorimetric assay), direct immunoassay, radiometric assay, or any other suitable detectable method known to the skilled artisan.
[0253] Many protein biochips are described in the art. These include, for example, protein biochips produced by Ciphergen Biosystems, Inc. (Fremont, CA), Zyomyx (Hayward, CA), Packard

BioScience Company (Meriden, CT), Phylos (Lexington, MA), Invitrogen (Carlsbad, CA), Biacore (Uppsala, Sweden) and Procognia (Berkshire, UK). Examples of such protein biochips are described in the following patents or published patent applications: U.S. Pat. Nos. 6,225,047; 6,537,749; 6,329,209; and 5,242,828; PCT International Publication Nos. WO 00/56934; WO 03/048768; and WO 99/51773.

Detection by Nucleic Acid Biochip

[0254] In aspects of the invention, a sample is analyzed by means of a nucleic acid biochip (also known as a nucleic acid microarray). To produce a nucleic acid biochip, oligonucleotides may be synthesized or bound to the surface of a substrate using a chemical coupling procedure and an ink jet application apparatus, as described in PCT application WO95/251116 (Baldeschweiler et al.). Alternatively, a gridded array may be used to arrange and link cDNA fragments or oligonucleotides to the surface of a substrate using a vacuum system, thermal, UV, mechanical or chemical bonding procedure.

[0255] A nucleic acid molecule (e.g. RNA or DNA) derived from a biological sample may be used to produce a hybridization probe as described herein. The biological samples are generally derived from a patient, e.g., as a bodily fluid (such as blood, blood serum, plasma, saliva, urine, ascites, cyst fluid, and the like); a homogenized tissue sample (e.g., a tissue sample obtained by biopsy); or a cell isolated from a patient sample. For some applications, cultured cells or other tissue preparations may be used. The mRNA is isolated according to standard methods, and cDNA is produced and used as a template to make complementary RNA suitable for hybridization. Such methods are well known in the art. The RNA is amplified in the presence of fluorescent nucleotides, and the labeled probes are then incubated with the microarray to allow the probe sequence to hybridize to complementary oligonucleotides bound to the biochip.

[0256] Incubation conditions are adjusted such that hybridization occurs with precise complementary matches or with various degrees of less complementarity depending on the degree of stringency employed, as defined above. The removal of nonhybridized probes may be accomplished, for example, by washing. The washing steps that follow hybridization can also vary in stringency, as defined above.

[0257] Detection systems for measuring the absence, presence, and amount of hybridization for all of the distinct nucleic acid sequences are well known in the art. For example, simultaneous detection is described in Heller et al., Proc. Natl. Acad. Sci. 94:2150-2155, 1997. In embodiments, a scanner is used to determine the levels and patterns of fluorescence.

Detection by Mass Spectrometry

[0258] In embodiments, the biomarkers of this disclosure are detected by mass spectrometry (MS). Mass spectrometry is a well-known tool for analyzing chemical compounds that employs a mass spectrometer to detect gas phase ions. Mass spectrometers are well known in the art and include, but are not limited to, time-of-flight, magnetic sector, quadrupole filter, ion trap, ion cyclotron resonance, electrostatic sector analyzer and hybrids of these. The method may be performed in an automated (Villanueva, et al., Nature Protocols (2006) 1(2):880-891) or semi-automated format. This can be accomplished, for example, with the mass spectrometer operably linked to a liquid chromatography device (LC-MS/MS or LC-MS) or gas chromatography device (GC-MS or GC-MS/MS). Methods for performing mass spectrometry are well known and have been disclosed, for example, in US Patent Application Publication Nos: 2005/0023454; 2005/0035286; U.S. Pat. No. 5,800,979 and the references disclosed therein.

Laser Desorption/Ionization

[0259] In embodiments, the mass spectrometer is a laser desorption/ionization mass spectrometer. In laser desorption/ionization mass spectrometry, the analytes are placed on the surface of a mass spectrometry probe, a device adapted to engage a probe interface of the mass spectrometer and to present an analyte to ionizing energy for ionization and introduction into a mass spectrometer. A laser desorption mass spectrometer employs laser energy, typically from an ultraviolet laser, but

also from an infrared laser, to desorb analytes from a surface, to volatilize and ionize them and make them available to the ion optics of the mass spectrometer. The analysis of proteins by LDI can take the form of MALDI or of SELDI.

[0260] Laser desorption/ionization in a single time of flight instrument typically is performed in linear extraction mode. Tandem mass spectrometers can employ orthogonal extraction modes.

Matrix-assisted Laser Desorption/Ionization (MALDI) and Electrospray Ionization (ESI)

[0261] In embodiments, the mass spectrometric technique for use as disclosed herein is matrix-assisted laser desorption/ionization (MALDI) or electrospray ionization (ESI). In related embodiments, the procedure is MALDI with time of flight (TOF) analysis, known as MALDI-TOF MS. This involves forming a matrix on a membrane with an agent that absorbs the incident light strongly at the particular wavelength employed. The sample is excited by UV or IR laser light into the vapor phase in the MALDI mass spectrometer. Ions are generated by the vaporization and form an ion plume. The ions are accelerated in an electric field and separated according to their time of travel along a given distance, giving a mass/charge (m/z) reading which is very accurate and sensitive. MALDI spectrometers are well known in the art and are commercially available from, for example, PerSeptive Biosystems, Inc. (Framingham, Mass., USA).

[0262] Magnetic-based serum processing can be combined with traditional MALDI-TOF. Through this approach, improved peptide capture is achieved prior to matrix mixture and deposition of the sample on MALDI target plates. Accordingly, in embodiments, methods of peptide capture are enhanced through the use of derivatized magnetic bead based sample processing.

[0263] MALDI-TOF MS allows scanning of the fragments of many proteins at once. Thus, many proteins can be run simultaneously on a polyacrylamide gel, subjected to a method of the disclosure to produce an array of spots on a collecting membrane, and the array may be analyzed. Subsequently, automated output of the results is provided by using a server (e.g., ExPASy) to generate the data in a form suitable for computers.

[0264] Other techniques for improving the mass accuracy and sensitivity of the MALDI-TOF MS can be used to analyze the fragments of protein obtained on a collection membrane. These include, but are not limited to, the use of delayed ion extraction, energy reflectors, ion-trap modules, and the like. In addition, post source decay and MS-MS analysis are useful to provide further structural analysis. With ESI, the sample is in the liquid phase and the analysis can be by ion-trap, TOF, single quadrupole, multi-quadrupole mass spectrometers, and the like. The use of such devices (other than a single quadrupole) allows MS-MS or MS” analysis to be performed. Tandem mass spectrometry allows multiple reactions to be monitored at the same time.

[0265] Capillary infusion may be employed to introduce the biomarker to a desired mass spectrometer implementation, for instance, because it can efficiently introduce small quantities of a sample into a mass spectrometer without destroying the vacuum. Capillary columns are routinely used to interface the ionization source of a mass spectrometer with other separation techniques including, but not limited to, gas chromatography (GC) and liquid chromatography (LC). GC and LC can serve to separate a solution into its different components prior to mass analysis. Such techniques are readily combined with mass spectrometry. One variation of the technique is the coupling of high-performance liquid chromatography (HPLC) to a mass spectrometer for integrated sample separation/and mass spectrometer analysis.

[0266] Quadrupole mass analyzers may also be employed as needed to practice the disclosure. Fourier-transform ion cyclotron resonance (FTMS) can also be used for some embodiments. It offers high resolution and the ability of tandem mass spectrometry experiments. FTMS is based on the principle of a charged particle orbiting in the presence of a magnetic field. Coupled to ESI and MALDI, FTMS offers high accuracy with errors as low as 0.001%.

Surface-Enhanced Laser Desorption/Ionization (SELDI)

[0267] In embodiments, the mass spectrometric technique for use herein is “Surface Enhanced Laser Desorption and Ionization” or “SELDI,” as described, for example, in U.S. Pat. Nos.

5,719,060; 6,225,047, both to Hutchens and Yip. This refers to a method of desorption/ionization gas phase ion spectrometry (e.g., mass spectrometry) in which an analyte (here, one or more of the biomarkers) is captured on the surface of a SELDI mass spectrometry probe.
[0268] SELDI has also been called “affinity capture mass spectrometry.” It also is called “Surface-Enhanced Affinity Capture” or “SEAC”. This version involves the use of probes that have a material on the probe surface that captures analytes through a non-covalent affinity interaction (adsorption) between the material and the analyte. The material is variously called an “adsorbent,” a “capture reagent,” an “affinity reagent” or a “binding moiety.” Such probes can be referred to as “affinity capture probes” and as having an “adsorbent surface.” The capture reagent can be any material capable of binding an analyte. The capture reagent is attached to the probe surface by physisorption or chemisorption. In certain embodiments, the probes have the capture reagent already attached to the surface. In other embodiments, the probes are pre-activated and include a reactive moiety that is capable of binding the capture reagent, e.g., through a reaction forming a covalent or coordinate covalent bond. Epoxide and acyl-imidizole are useful reactive moieties to covalently bind polypeptide capture reagents such as antibodies or cellular receptors. Nitrilotriacetic acid and iminodiacetic acid are useful reactive moieties that function as chelating agents to bind metal ions that interact non-covalently with histidine containing peptides. Adsorbents are generally classified as chromatographic adsorbents and biospecific adsorbents.
[0269] “Chromatographic adsorbent” refers to an adsorbent material typically used in chromatography. Chromatographic adsorbents include, for example, ion exchange materials, metal chelators (e.g., nitrilotriacetic acid or iminodiacetic acid), immobilized metal chelates, hydrophobic interaction adsorbents, hydrophilic interaction adsorbents, dyes, simple biomolecules (e.g., nucleotides, amino acids, simple sugars and fatty acids) and mixed mode adsorbents (e.g., hydrophobic attraction/electrostatic repulsion adsorbents).
[0270] A biospecific adsorbent is an adsorbent comprising a biomolecule, e.g., a nucleic acid molecule (e.g., an aptamer), a polypeptide, a polysaccharide, a lipid, a steroid or a conjugate of these (e.g., a glycoprotein, a lipoprotein, a glycolipid, a nucleic acid (e.g., DNA)-protein conjugate). In certain instances, the biospecific adsorbent can be a macromolecular structure such as a multiprotein complex, a biological membrane or a virus. Examples of biospecific adsorbents are antibodies, receptor proteins and nucleic acids. Biospecific adsorbents typically have higher specificity for a target analyte than chromatographic adsorbents. Further examples of adsorbents for use in SELDI can be found in U.S. Pat. No. 6,225,047. A “bioselective adsorbent” refers to an adsorbent that binds to an analyte with an affinity of at least 10.sup.−8 M.
[0271] Protein biochips produced by Ciphergen comprise surfaces having chromatographic or biospecific adsorbents attached thereto at addressable locations. Ciphergen's ProteinChip® arrays include NP20 (hydrophilic); H4 and H50 (hydrophobic); SAX-2, Q-10 and (anion exchange); WCX-2 and CM-10 (cation exchange); IMAC-3, IMAC-30 and IMAC-50 (metal chelate); and PS-10, PS-20 (reactive surface with acyl-imidazole, epoxide) and PG-20 (protein G coupled through acyl-imidazole). Hydrophobic ProteinChip arrays have isopropyl or nonylphenoxy-poly(ethylene glycol)methacrylate functionalities. Anion exchange ProteinChip arrays have quaternary ammonium functionalities. Cation exchange ProteinChip arrays have carboxylate functionalities. Immobilized metal chelate ProteinChip arrays have nitrilotriacetic acid functionalities (IMAC 3 and IMAC 30) or O-methacryloyl-N,N-bis-carboxymethyl tyrosine functionalities (IMAC 50) that adsorb transition metal ions, such as copper, nickel, zinc, and gallium, by chelation. Preactivated ProteinChip arrays have acyl-imidazole or epoxide functional groups that can react with groups on proteins for covalent binding.
[0272] Such biochips are further described in: U.S. Pat. No. 6,579,719 (Hutchens and Yip, “Retentate Chromatography,” Jun. 17, 2003); U.S. Pat. No. 6,897,072 (Rich et al., “Probes for a Gas Phase Ion Spectrometer,” May 24, 2005); U.S. Pat. No. 6,555,813 (Beecher et al., “Sample Holder with Hydrophobic Coating for Gas Phase Mass Spectrometer,” Apr. 29, 2003); U.S. Patent

Application Publication No. U.S. 2003/0032043 A1 (Pohl and Papanu, "Latex Based Adsorbent Chip," Jul. 16, 2002); and PCT International Publication No. WO 03/040700 (Um et al., "Hydrophobic Surface Chip," May 15, 2003); U.S. Patent Application Publication No. US 2003/0218130 A1 (Boschetti et al., "Biochips With Surfaces Coated With Polysaccharide-Based Hydrogels," Apr. 14, 2003) and U.S. Pat. No. 7,045,366 (Huang et al., "Photocrosslinked Hydrogel Blend Surface Coatings" May 16, 2006).

[0273] In general, a probe with an adsorbent surface is contacted with the sample for a period of time sufficient to allow the biomarker or biomarkers that may be present in the sample to bind to the adsorbent. After an incubation period, the substrate is washed to remove unbound material. Any suitable washing solutions can be used; preferably, aqueous solutions are employed. The extent to which molecules remain bound can be manipulated by adjusting the stringency of the wash. The elution characteristics of a wash solution can depend, for example, on pH, ionic strength, hydrophobicity, degree of chaotropism, detergent strength, and temperature. Unless the probe has both SEAC and SEND properties (as described herein), an energy absorbing molecule then is applied to the substrate with the bound biomarkers.

[0274] In yet another method, one can capture the biomarkers with a solid-phase bound immuno-adsorbent that has antibodies that bind the biomarkers. After washing the adsorbent to remove unbound material, the biomarkers are eluted from the solid phase and detected by applying to a SELDI biochip that binds the biomarkers and analyzing by SELDI.

[0275] The biomarkers bound to the substrates are detected in a gas phase ion spectrometer such as a time-of-flight mass spectrometer. The biomarkers are ionized by an ionization source such as a laser, the generated ions are collected by an ion optic assembly, and then a mass analyzer disperses and analyzes the passing ions. The detector then translates information of the detected ions into mass-to-charge ratios. Detection of a biomarker typically will involve detection of signal intensity. Thus, both the quantity and mass of the biomarker can be determined.

Classification Algorithms

[0276] The present disclosure provides methods for characterizing a chronic lymphocytic leukemia (CLL) as belonging to an expression subtype (e.g., Ec-i, EC-m1, EC-m2, EC-m3, EC-m4, EC-o, EC-ul, and EC-u2). The expression subtype is useful in predicting clinical outcome for a CLL patient and/or for guiding therapy.

[0277] In some embodiments, data derived from the assays for detection of biomarkers (e.g., RNA-seq) that are generated using samples such as "known samples" can then be used to "train" a classification model. Exemplary methods for developing a model for classifying a chronic lymphocytic leukemia as belonging to an expression subtype are described in the Examples provided herein. A "known sample" is a sample that has been pre-classified. The data used to form the classification model can be referred to as a "training data set." Once trained, the classification model (e.g., a machine learning classifier) can be used to classify the expression subtype of a chronic lymphocytic leukemia (CLL) based upon levels of biomarkers detected in a sample. The sample can be taken from a subject having CLL. This can be useful, for example, in guiding selection of a treatment for a subject or for prognostic purposes.

[0278] The training data set that is used to form the classification model may comprise raw data or pre-processed data. In embodiments, a classifier can be trained using a random forest classifier, as described in the Examples provided herein.

[0279] Classification models can be formed using any suitable statistical classification (or "learning") method that attempts to segregate bodies of data into classes based on objective parameters present in the data. Classification methods may be either supervised or unsupervised. Examples of supervised and unsupervised classification processes are described in Jain, "Statistical Pattern Recognition: A Review", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 1, January 2000, the teachings of which are incorporated by reference herein in their entirety.

[0280] In supervised classification, training data containing examples of known categories are presented to a learning mechanism, which learns one or more sets of relationships that define each of the known classes. New data may then be applied to the learning mechanism, which then classifies the new data using the learned relationships. Examples of supervised classification processes include linear regression processes (e.g., multiple linear regression (MLR), partial least squares (PLS) regression and principal components regression (PCR)), binary decision trees (e.g., recursive partitioning processes such as CART—classification and regression trees), artificial neural networks such as back propagation networks, discriminant analyses (e.g., Bayesian classifier or Fischer analysis), logistic classifiers, and support vector classifiers (support vector machines).
[0281] In embodiments, a supervised classification method is a recursive partitioning process. Recursive partitioning processes use recursive partitioning trees to classify data derived from unknown samples. Further details about recursive partitioning processes are provided in U.S. Patent Application Publication No. 2002/0138208 A1 to Paulse et al., “Method for analyzing mass spectra.”
[0282] In embodiments, the classification models that are created can be formed using unsupervised learning methods. Unsupervised classification attempts to learn classifications based on similarities in the training data set, without pre-classifying the spectra from which the training data set was derived. Unsupervised learning methods include cluster analyses. A cluster analysis attempts to divide the data into “clusters” or groups that ideally should have members that are very similar to each other, and very dissimilar to members of other clusters. Similarity is then measured using some distance metric, which measures the distance between data items, and clusters together data items that are closer to each other. Clustering techniques include the MacQueen's K-means algorithm and the Kohonen's Self-Organizing Map algorithm.
[0283] Learning algorithms asserted for use in classifying biological information are described, for example, in PCT International Publication No. WO 01/31580 (Barnhill et al., “Methods and devices for identifying patterns in biological systems and methods of use thereof”), U.S. Patent Application Publication No. 2002/0193950 A1 (Gavin et al., “Method or analyzing mass spectra”), U.S. Patent Application Publication No. 2003 0004402 A1 (Hitt et al., “Process for discriminating between biological states based on hidden patterns from biological data”), and U.S. Patent Application Publication No. 2003/0055615 A1 (Zhang and Zhang, “Systems and methods for processing biological expression data”).
[0284] The classification models can be formed on and used on any suitable digital computer. Suitable digital computers include micro, mini, or large computers using any standard or specialized operating system, such as a Unix, Windows® or Linux® based operating system. The digital computer that is used may be physically separate from a device that is used to detect biomarkers, or it may be coupled to the device.
[0285] The training data set and the classification models according to embodiments of the invention can be embodied by computer code that is executed or used by a digital computer. The computer code can be stored on any suitable computer readable media including optical or magnetic disks, sticks, tapes, etc., and can be written in any suitable computer programming language including C, C++, visual basic, etc.
Hardware and Software
[0286] The present invention also provides a computer system useful in analyzing data associated with biomarker expression, patient selection, and related computations (e.g., calculations associated with a machine learning classifier).
[0287] A computer system (or digital device) may be used to receive, transmit, display and/or store results, analyze the results, and/or produce a report of the results and analysis. A computer system may be understood as a logical apparatus that can read instructions from media (e.g. software) and/or network port (e.g. from the internet), which can optionally be connected to a server having fixed media. A computer system may comprise one or more of a CPU, disk drives, input devices

such as keyboard and/or mouse, and a display (e.g. a monitor). Data communication, such as transmission of instructions or reports, can be achieved through a communication medium to a server at a local or a remote location. The communication medium can include any means of transmitting and/or receiving data. For example, the communication medium can be a network connection, a wireless connection, or an internet connection. Such a connection can provide for communication over the World Wide Web. It is envisioned that data relating to the present invention can be transmitted over such networks or connections (or any other suitable means for transmitting information, including but not limited to mailing a physical report, such as a print-out) for reception and/or for review by a receiver. One can record results of calculations (e.g., sequence analysis or a listing of hybrid capture probe sequences) made by a computer on tangible medium, for example, in computer-readable format such as a memory drive or disk, as an output displayed on a computer monitor or other monitor, or simply printed on paper. The results can be reported on a computer screen. The receiver can be but is not limited to an individual, or electronic system (e.g. one or more computers, and/or one or more servers).

[0288] In some embodiments, the computer system may comprise one or more processors. Processors may be associated with one or more controllers, calculation units, and/or other units of a computer system, or implanted in firmware as desired. If implemented in software, the routines may be stored in any computer readable memory such as in RAM, ROM, flash memory, a magnetic disk, a laser disk, or other suitable storage medium. Likewise, this software may be delivered to a computing device via any known delivery method including, for example, over a communication channel such as a telephone line, the internet, a wireless connection, etc., or via a transportable medium, such as a computer readable disk, flash drive, etc. The various steps may be implemented as various blocks, operations, tools, modules and techniques which, in turn, may be implemented in hardware, firmware, software, or any combination of hardware, firmware, and/or software. When implemented in hardware, some or all of the blocks, operations, techniques, etc. may be implemented in, for example, a custom integrated circuit (IC), an application specific integrated circuit (ASIC), a field programmable logic array (FPGA), a programmable logic array (PLA), etc.

[0289] A client-server, relational database architecture can be used in embodiments of the invention. A client-server architecture is a network architecture in which each computer or processor on the network is either a client or a server. Server computers are typically powerful computers dedicated to managing disk drives (file servers), printers (print servers), or network traffic (network servers). Client computers include PCs (personal computers) or workstations on which users run applications, as well as example output devices as disclosed herein. Client computers rely on server computers for resources, such as files, devices, and even processing power. In some embodiments of the invention, the server computer handles all of the database functionality. The client computer can have software that handles all the front-end data management and can also receive data input from users.

[0290] A machine readable medium that may comprise computer-executable code may take many forms, including but not limited to, a tangible storage medium, a carrier wave medium or physical transmission medium. Non-volatile storage media include, for example, optical or magnetic disks, such as any of the storage devices in any computer(s) or the like, such as may be used to implement the databases, etc. shown in the drawings. Volatile storage media include dynamic memory, such as main memory of such a computer platform. Tangible transmission media include coaxial cables; copper wire and fiber optics, including the wires that comprise a bus within a computer system. Carrier-wave transmission media may take the form of electric or electromagnetic signals, or acoustic or light waves such as those generated during radio frequency (RF) and infrared (IR) data communications. Common forms of computer-readable media therefore include for example: a floppy disk, a flexible disk, hard disk, magnetic tape, any other magnetic medium, a CD-ROM, DVD or DVD-ROM, any other optical medium, punch cards paper tape, any other physical storage medium with patterns of holes, a RAM, a ROM, a PROM and EPROM, a FLASH-EPROM, any

other memory chip or cartridge, a carrier wave transporting data or instructions, cables or links transporting such a carrier wave, or any other medium from which a computer may read programming code and/or data. Many of these forms of computer readable media may be involved in carrying one or more sequences of one or more instructions to a processor for execution.
[0291] The subject computer-executable code can be executed on any suitable device which may comprise a processor, including a server, a PC, or a mobile device such as a smartphone or tablet. Any controller or computer optionally includes a monitor, which can be a cathode ray tube ("CRT") display, a flat panel display (e.g., active matrix liquid crystal display, liquid crystal display, etc.), or others. Computer circuitry is often placed in a box, which includes numerous integrated circuit chips, such as a microprocessor, memory, interface circuits, and others. The box also optionally includes a hard disk drive, a floppy disk drive, a high capacity removable drive such as a writeable CD-ROM, and other common peripheral elements. Inputting devices such as a keyboard, mouse, or touch-sensitive screen, optionally provide for input from a user.
[0292] The computer can include appropriate software for receiving user instructions, either in the form of user input into a set of parameter fields, e.g., in a GUI, or in the form of preprogrammed instructions, e.g., preprogrammed for a variety of different specific operations.
Pharmaceutical Compositions
[0293] As reported herein, the panels of biomarkers presented herein can be used in a method to select a subject for treatment with an agent. In embodiments, the treatment is administered as part of a clinical trial. Accordingly, the disclosure provides chemotherapeutic compositions for treatment of chronic lymphocytic leukemia (CLL). Non-limiting examples of agents suitable for use in the methods provided herein include those listed in Tables 1A and 1B or otherwise listed herein. The compositions should be sterile and contain a therapeutically effective amount of the polypeptides or nucleic acid molecules in a unit of weight or volume suitable for administration to a subject.
[0294] In embodiments, the composition contains a drug selected from one of those listed in Tables 1A and 1B below, and the like (e.g., alternative drugs effective in the treatment of chronic lymphocytic leukemia (CLL)). In embodiments, the drug has the same main target, or the same target category (A) or (B) as a drug listed in Tables 1A and 1B.
TABLE-US-00003 TABLE 1A Agents for treatment of CLL. drug_name drug_name_alias drug_category target_category main_targets A-1331852 BH3 mimetic Apoptosis (BH3) BCL-XL Abexinostat HDAC inhibitor Epigenetics pan-HDAC (mostly HDAC1) Acalabrutinib ACP-196 BTK inhibitor B cell receptor BTK signaling Atorvastatin Statin Antilipemic HMG-COA Azacitidine DNA methylation Epigenetics inhibitor AZD5991 BH3 mimetic Apoptosis (BH3) MCL1 AZD8055 mTOR inhibitor mTOR signaling mTOR Bendamustine DNA alkylator DNA damage DNA response Carfilzomib PR-171 Proteasome Proteasome inhibitor Cerdulatinib PRT062070 SYK and BCR and JAK1, JAK2, JAK/STAT JAK/STAT JAK3, SYK, inhibitor signaling TYK2 Crizotinib PF-02341066 ALK inhibitor ALK/MET/ROS ALK, c- Met/HGF, ROS Dasatinib BMS-354825 BCR/ABL inhibitor BCR-ABL BCR-ABL, SRC Duvelisib IPI-145 PI3K inhibitor PI3K/AKT PI3K signaling Entospletinib GS-9973 SYK inhibitor B cell receptor SYK signaling Erastin Ferroptosis inducer Ferroptosis VDAC Fludarabine NSC 118218 Antimetabolite DNA synthesis Gandotinib LY2784544 JAK/STAT JAK/STAT JAK2, FLT3, inhibitor signaling FLT4, FGFR2, TYK2, TRKB GSK690693 PI3K/AKT/mTOR PI3K/AKT AKT1, AKT2, inhibitor signaling AKT3 Ibrutinib PCI-32765 BTK inhibitor B cell receptor BTK signaling Idelalisib CAL-101 PI3K inhibitor PI3K/AKT PI3K signaling JQ1 Bromodomain Epigenetics BET inhibitor bromodomain inhibitor Lenalidomide CC-5013 Immunomodulatory Immunomodulati IKZF1, IKZF3 drug (IMiD) on MK-2206 MK2206 AKT inhibitor PI3K/AKT AKT1, AKT2, signaling AKT3 Navitoclax ABT-263 BH3 mimetic Apoptosis (BH3) BCL-XL, BCL-2 and BCL-W Nirogacestat PF-3084014 γ-secretase NOTCH γ-secretase inhibitor signaling Nutlin-3 MDM2 inhibitor DNA damage MDM2 response Onalespib AT13387 HSP90 inhibitor Heat shock HSP90

protein Osimertinib AZD9291 EGFR inhibitor EGFR signaling EGFR (mutated) Ponatinib AP24534 BCR-ABL inhibitor BCR-ABL BCR-ABL, PDGFR, VEGFR, FGFR, SRC Rapamycin AY-22989 mTOR inhibitor mTOR signaling mTORC1, FKBP12 Ricolinostat ACY-1215 HDAC inhibitor Epigenetics HDAC6 (less: HDAC1, HDAC2, HDAC3) Ruxolitinib INCB18424 JAK/STAT JAK/STAT JAK1, JAK2 inhibitor signaling Selinexor KPT-330 XPO1 inhibitor Nuclear export XPO1 Sorafenib BAY-43-9006 MAPK inhibitor MAPK signaling RAF, PDGF, VEGFR2, VEGFR3, KIT Sunitinib SU-11248 VEGF inhibitor VEGF signaling Trametinib GSK1120212 MAPK inhibitor MAPK signaling MEK1, MEK2 Umbralisib TGR1202 PI3K inhibitor PI3K/AKT PI3Kδ signaling Vecabrutinib SNS-062 BTK inhibitor B cell receptor signaling Venetoclax ABT-199 BH3 mimetic Apoptosis (BH3) BCL2 Vorinostat MK0683 HDAC inhibitor Epigenetics Voruciclib NSC-3590 CDK inhibitor Cell cycle CDK9, MCL1, control CDK4/6, CDK1 Zanubrutinib BGB-3111 BTK inhibitor B cell receptor BTK signaling

TABLE-US-00004 TABLE 1B Agents for treatment of CLL. Drug Target 1 Idelalisib (CAL-101) PI3K-delta 2 Duvelisib (IPI-145) PI3K-delta, gamma 3 Umbralisib PI3K-d 4 Gandotinib JAK2 5 Trametinib MEK1/MEK2 6 Sorafenib BRAF 7 Rapamycin mTOR 8 AZD8055 mTOR 9 MK2206 AKT 10 Acalabrutinib BTK 11 SNS-062 BTK 12 Entosplentinib SYK 13 Cerdulatinib JAK/SYK 14 Crizotinib ALK 15 Dasatinib Abl 16 Ponatinib TK 17 Ruxolitinib JAK1/2 18 Voruciclib CDK9 19 Sunitinib RTK 20 Carfilzomib Proteosome 21 Lenalidomide Immunomodulator 22 Fludarabine DNA damaging agent 23 Bendamustine Alkylating agent 24 Azacitidine DNA methylating agent 25 Nutlin-3 MDM2 26 Abexinostat HDAC 27 Vorinostat HDAC 28 Ricolinostat selective HDAC 6 29 JQ1 Bromodomain 30 ABT-263 BCL-2/BCL-XL/BCLw 31 AZD5991 MCL-1 32 A-1331852 BCL-xl 33 Erastin Ferroptosis inducer 34 Atoruvastatin HMG-CoA reductase 35 Onalespib HSP90 36 PF-3084014 gamma-secretase 37 GSK690396 AKT 38 Selinexor Nuclear export 39 Venetoclax BCL2 40 Ibrutinib BTK 41 AZD9291 EGFR 42 Zanabrutinib BTK

TABLE-US-00005 TABLE 2 Gene Drivers and their Sensitivity to CLL Treatment Agents dprim_pear- peptide drug feature son_corr direction PUMA A-1331852 loss_6q21 −0.5859268 Resistant 1 uM PUMA A-1331852 CNOT3 0.4470944 Sensitive 1 uM PUMA A-1331852 EC-m2 −0.3621573 Resistant 1 uM PUMA AZD5991 EC-i 0.45349954 Sensitive 1 uM PUMA AZD5991 BIRC3 0.4088154 Sensitive 1 uM PUMA AZD5991 loss_7p22.2 0.44897072 Sensitive 1 uM PUMA AZD5991 loss_16p11.2 0.37742434 Sensitive 1 uM PUMA AZD5991 loss_17q11.2 0.4088154 Sensitive 1 uM PUMA AZD5991 loss_1q42.13 0.35362592 Sensitive 1 uM PUMA AZD5991 loss_9p21.3 0.4088154 Sensitive 1 uM PUMA Abexinostat loss_17q23.1 0.33522272 Sensitive 1 uM PUMA Abexinostat loss_6q21 0.3538112 Sensitive 1 uM PUMA Abexinostat EC-m2 0.45571182 Sensitive 1 uM PUMA Atorvastatin RFX7 −0.3377731 Resistant 1 uM PUMA Atorvastatin FBXW7 −0.3807836 Resistant 1 uM PUMA Atorvastatin loss_2q31.1 −0.3576906 Resistant 1 uM PUMA Atorvastatin ZMYM3 −0.3377731 Resistant 1 uM PUMA Atorvastatin loss_7p22.2 −0.3484779 Resistant 1 uM PUMA Atorvastatin gain_16p11.2 −0.3743338 Resistant 1 uM PUMA Atorvastatin loss_10p12.2 −0.3576906 Resistant 1 uM PUMA Azacitidine CARD11 0.50143359 Sensitive 1 uM PUMA Bendamustine loss_14q32.12 0.33413206 Sensitive 1 uM PUMA Bendamustine gain_7q22.1 0.36184927 Sensitive 1 uM PUMA Carfilzomib i-CLL 0.43213969 Sensitive 1 uM PUMA Carfilzomib loss_3p21.31 0.35409837 Sensitive 1 uM PUMA Carfilzomib loss_14q32.12 0.35409837 Sensitive 1 uM PUMA Carfilzomib gain_7q22.1 0.47287271 Sensitive 1 uM PUMA Carfilzomib MYD88 0.44097787 Sensitive 1 uM PUMA Cerdulatinib loss_13q14.3 0.33070389 Sensitive 1 uM PUMA Cerdulatinib loss_13q14.13 0.45233099 Sensitive 1 uM PUMA Crizotinib EC-i −0.3956383 Resistant 1 uM PUMA Crizotinib NOTCH1 −0.3508287 Resistant 1 uM PUMA Crizotinib IGLV321_R110 −0.3508287 Resistant 1 uM PUMA Crizotinib loss_6q −0.3508287 Resistant 1 uM PUMA Crizotinib RPS15 −0.3508287 Resistant 1 uM PUMA Dasatinib NOTCH1 −0.478335 Resistant 1 uM PUMA Dasatinib IGLV321_R110 −0.478335 Resistant 1 uM PUMA Dasatinib loss_17q23.1 0.35195754 Sensitive 1 uM PUMA Dasatinib loss_6q −0.478335 Resistant 1 uM PUMA Dasatinib RPS15 −0.478335

Resistant 1 uM PUMA Duvelisib loss_17q23.1 0.34171963 Sensitive 1 uM PUMA Duvelisib loss_17p13.1 0.34171963 Sensitive 1 uM PUMA Duvelisib EC-m2 0.3945238 Sensitive 1 uM PUMA Entospletinib loss_19p13.11 0.40294784 Sensitive 1 uM PUMA Erastin loss_8p −0.4591263 Resistant 1 uM PUMA Erastin EC-m3 −0.3301407 Resistant 1 uM PUMA Erastin EC-i −0.3373168 Resistant 1 uM PUMA Erastin loss_7q36.1 −0.4591263 Resistant 1 uM PUMA GSK690693 i-CLL 0.41029257 Sensitive 1 uM PUMA GSK690693 BIRC3 0.43691279 Sensitive 1 uM PUMA GSK690693 ITIH2 0.43691279 Sensitive 1 uM PUMA GSK690693 loss_17q11.2 0.43691279 Sensitive 1 uM PUMA GSK690693 loss_9p21.3 0.43691279 Sensitive 1 uM PUMA Gandotinib EC-i −0.3507015 Resistant 1 uM PUMA Gandotinib GNB1 0.33399627 Sensitive 1 uM PUMA Ibrutinib loss_6q21 −0.5196509 Resistant 1 uM PUMA Idelalisib DICER1 −0.3326945 Resistant 1 uM PUMA Idelalisib MED1 −0.3326945 Resistant 1 uM PUMA Idelalisib EC-m4 0.34331218 Sensitive 1 uM PUMA Idelalisib KMT2D 0.3371123 Sensitive 1 uM PUMA JQ1 loss_8p −0.6160372 Resistant 1 uM PUMA JQ1 EC-i −0.5959866 Resistant 1 uM PUMA JQ1 loss_1q21.3 −0.3770763 Resistant 1 uM PUMA JQ1 loss_12p13.31a −0.5064866 Resistant 1 uM PUMA JQ1 loss_18q21.2 0.33831656 Sensitive 1 uM PUMA JQ1 loss_5p15.33 −0.5064866 Resistant 1 uM PUMA JQ1 loss_2q31.1 −0.4197346 Resistant 1 uM PUMA JQ1 loss_3p13 0.36761499 Sensitive 1 uM PUMA JQ1 loss_16q22.1 −0.373144 Resistant 1 uM PUMA JQ1 ATM −0.4530874 Resistant 1 uM PUMA JQ1 loss_7p22.2 −0.3770763 Resistant 1 uM PUMA JQ1 gain_16p11.2 −0.390517 Resistant 1 uM PUMA JQ1 loss_10q24.32 0.33831656 Sensitive 1 uM PUMA JQ1 loss_3p21.31 −0.6160372 Resistant 1 uM PUMA JQ1 loss_14q32.12 −0.6160372 Resistant 1 uM PUMA JQ1 loss_10p12.2 −0.4197346 Resistant 1 uM PUMA JQ1 loss_1p36.11 0.33831656 Sensitive 1 uM PUMA JQ1 loss_7q36.1 −0.6160372 Resistant 1 uM PUMA JQ1 MYD88 −0.6160372 Resistant 1 uM PUMA JQ1 loss_17p13.3 −0.5064866 Resistant 1 uM PUMA MK-2206 priortrt_Post −0.3482936 Resistant 1 uM PUMA Navitoclax EC-m4 0.33245488 Sensitive 1 uM PUMA Navitoclax loss_6q21 −0.5997546 Resistant 1 uM PUMA Navitoclax loss_14q32.12 0.35735421 Sensitive 1 uM PUMA Navitoclax EC-m2 −0.4445467 Resistant 1 uM PUMA Nirogacestat BIRC3 0.4303189 Sensitive 1 uM PUMA Nirogacestat loss_1q21.3 0.33363306 Sensitive 1 uM PUMA Nirogacestat loss_7p22.2 0.36058747 Sensitive 1 uM PUMA Nirogacestat loss_17q11.2 0.4303189 Sensitive 1 uM PUMA Nirogacestat loss_9p21.3 0.4303189 Sensitive 1 uM PUMA Nutlin-3 priortrt_Post −0.4306035 Resistant 1 uM PUMA Nutlin-3 FBXW7 −0.3313826 Resistant 1 uM PUMA Nutlin-3 EC-m4 0.3496268 Sensitive 1 uM PUMA Nutlin-3 loss_17p13.1 −0.3687944 Resistant 1 uM PUMA Onalespib loss_6q21 −0.5471562 Resistant 1 uM PUMA Osimertinib BIRC3 0.38799985 Sensitive 1 uM PUMA Osimertinib loss_3p21.31 −0.3558358 Resistant 1 uM PUMA Osimertinib loss_17q11.2 0.38799985 Sensitive 1 uM PUMA Osimertinib loss_9p21.3 0.38799985 Sensitive 1 uM PUMA Osimertinib MYD88 −0.4725921 Resistant 1 uM PUMA Ponatinib RFX7 −0.4213822 Resistant 1 uM PUMA Ponatinib priortrt_Post −0.5032364 Resistant 1 uM PUMA Ponatinib loss_2q31.1 −0.3631316 Resistant 1 uM PUMA Ponatinib ZMYM3 −0.4213822 Resistant 1 uM PUMA Ponatinib loss_10p12.2 −0.3631316 Resistant 1 uM PUMA Ponatinib m-CLL 0.41758483 Sensitive 1 uM PUMA Rapamycin RFX7 −0.5651278 Resistant 1 uM PUMA Rapamycin EC-u1 −0.4444035 Resistant 1 uM PUMA Rapamycin i-CLL 0.46970843 Sensitive 1 uM PUMA Rapamycin priortrt_Post −0.3633407 Resistant 1 uM PUMA Rapamycin U-CLL −0.4657341 Resistant 1 uM PUMA Rapamycin EC-m4 0.41842197 Sensitive 1 uM PUMA Rapamycin ZMYM3 −0.5651278 Resistant 1 uM PUMA Rapamycin n-CLL −0.4739873 Resistant 1 uM PUMA Rapamycin M-CLL 0.46573411 Sensitive 1 uM PUMA Ricolinostat loss_15q15.1b −0.3847567 Resistant 1 uM PUMA Ricolinostat i-CLL −0.3930618 Resistant 1 uM PUMA Ricolinostat loss_1q21.3 −0.3766571 Resistant 1 uM PUMA Ricolinostat loss_12p13.31a −0.3467068 Resistant 1 uM PUMA Ricolinostat loss_5p15.33 −0.3467068 Resistant 1 uM PUMA Ricolinostat loss_2q31.1 −0.3442189 Resistant 1 uM PUMA Ricolinostat CARD11 0.34960477 Sensitive 1 uM PUMA Ricolinostat loss_16q22.1 −0.3467068 Resistant 1 uM PUMA Ricolinostat loss_7p22.2 −0.3766571 Resistant 1 uM PUMA Ricolinostat

gain_16p11.2 −0.4183658 Resistant 1 uM PUMA Ricolinostat loss_3p21.31 −0.4057486 Resistant 1 uM PUMA Ricolinostat loss_14q32.12 −0.4057486 Resistant 1 uM PUMA Ricolinostat loss_10p12.2 −0.3442189 Resistant 1 uM PUMA Ricolinostat loss_20p11.22 −0.3847567 Resistant 1 uM PUMA Ricolinostat m-CLL 0.38059876 Sensitive 1 uM PUMA Ricolinostat loss_17p13.3 −0.3467068 Resistant 1 uM PUMA Ruxolitinib U-CLL −0.3389186 Resistant 1 uM PUMA Ruxolitinib loss_11q22.3 −0.3372231 Resistant 1 uM PUMA Ruxolitinib n-CLL −0.3553018 Resistant 1 uM PUMA Ruxolitinib M-CLL 0.33891856 Sensitive 1 uM PUMA Selinexor loss_8p 0.35546713 Sensitive 1 uM PUMA Selinexor loss_7q36.1 0.35546713 Sensitive 1 uM PUMA Sorafenib RFX7 −0.5439862 Resistant 1 uM PUMA Sorafenib FBXW7 −0.3983393 Resistant 1 uM PUMA Sorafenib ZMYM3 −0.5439862 Resistant 1 uM PUMA Sunitinib gain_19p13.3 −0.3705342 Resistant 1 uM PUMA Sunitinib RFX7 −0.5237902 Resistant 1 uM PUMA Sunitinib FBXW7 −0.4644002 Resistant 1 uM PUMA Sunitinib loss_2q31.1 −0.3637593 Resistant 1 uM PUMA Sunitinib ZMYM3 −0.5237902 Resistant 1 uM PUMA Sunitinib loss_10p12.2 −0.3637593 Resistant 1 uM PUMA Trametinib CHD2 0.53055601 Sensitive 1 uM PUMA Trametinib TP53 0.37315291 Sensitive 1 uM PUMA Trametinib EC-m1 −0.3588799 Resistant 1 uM PUMA Trametinib CUL9 −0.353067 Resistant 1 uM PUMA Umbralisib NOTCH1 −0.3958792 Resistant 1 uM PUMA Umbralisib IGLV321_R110 −0.3960778 Resistant 1 uM PUMA Umbralisib loss_12p13.31a 0.38748565 Sensitive 1 uM PUMA Umbralisib U-CLL −0.4629576 Resistant 1 uM PUMA Umbralisib loss_5p15.33 0.38748565 Sensitive 1 uM PUMA Umbralisib n-CLL −0.3356869 Resistant 1 uM PUMA Umbralisib loss_6q −0.3960778 Resistant 1 uM PUMA Umbralisib M-CLL 0.46295764 Sensitive 1 uM PUMA Umbralisib RPS15 −0.3960778 Resistant 1 uM PUMA Venetoclax loss_6q21 −0.6708865 Resistant 1 uM PUMA Venetoclax EC-m2 −0.5106933 Resistant 1 uM PUMA Vorinostat i-CLL −0.3317217 Resistant 1 uM PUMA Vorinostat loss_6q21 0.336181 Sensitive 1 uM PUMA Vorinostat EC-m2 0.3339204 Sensitive 1 uM PUMA Voruciclib i-CLL 0.37258884 Sensitive 1 uM PUMA Voruciclib SF3B1 −0.3336378 Resistant 1 uM PUMA Zanubrutinib gain_19p13.3 −0.3860208 Resistant 1 uM PUMA Zanubrutinib BIRC3 −0.5322225 Resistant 1 uM PUMA Zanubrutinib loss_1q21.3 −0.3844555 Resistant 1 uM PUMA Zanubrutinib FBXW7 −0.3945898 Resistant 1 uM PUMA Zanubrutinib ITIH2 −0.5322225 Resistant 1 uM PUMA Zanubrutinib loss_7p22.2 −0.3844555 Resistant 1 uM PUMA Zanubrutinib gain_16p11.2 −0.4882074 Resistant 1 uM PUMA Zanubrutinib loss_17p13.1 −0.5322225 Resistant 1 uM PUMA Zanubrutinib loss_17q11.2 −0.5322225 Resistant 1 uM PUMA Zanubrutinib loss_9p21.3 −0.5322225 Resistant 1 uM MS1 AZD5991 CNOT3 0.75935526 Sensitive 2.5 uM MS1 AZD8055 NOTCH1 −0.3618136 Resistant 2.5 uM MS1 AZD8055 IGLV321_R110 −0.3618136 Resistant 2.5 uM MS1 AZD8055 loss_18q21.2 −0.3305838 Resistant 2.5 uM MS1 AZD8055 loss_3p13 −0.3305838 Resistant 2.5 uM MS1 AZD8055 loss_6q −0.3618136 Resistant 2.5 uM MS1 AZD8055 loss_1p36.11 −0.3305838 Resistant 2.5 uM MS1 AZD8055 RPS15 −0.3618136 Resistant 2.5 uM MS1 Abexinostat priortrt_Post 0.35791198 Sensitive 2.5 uM MS1 Abexinostat FBXW7 0.3832463 Sensitive 2.5 uM MS1 Abexinostat ITIH2 0.34485221 Sensitive 2.5 uM MS1 Atorvastatin RFX7 −0.8667203 Resistant 2.5 uM MS1 Atorvastatin loss_11q22.3 −0.4356328 Resistant 2.5 uM MS1 Atorvastatin loss_10q24.2 −0.5072844 Resistant 2.5 uM MS1 Atorvastatin loss_2q31.1 −0.4261955 Resistant 2.5 uM MS1 Atorvastatin ZMYM3 −0.8667203 Resistant 2.5 uM MS1 Atorvastatin n-CLL −0.3847935 Resistant 2.5 uM MS1 Atorvastatin loss_10p12.2 −0.4261955 Resistant 2.5 uM MS1 Azacitidine BIRC3 0.44950434 Sensitive 2.5 uM MS1 Azacitidine ITIH2 0.44950434 Sensitive 2.5 uM MS1 Azacitidine loss_19p13.11 0.33873888 Sensitive 2.5 uM MS1 Azacitidine loss_17q11.2 0.44950434 Sensitive 2.5 uM MS1 Azacitidine loss_9p21.3 0.44950434 Sensitive 2.5 uM MS1 Bendamustine BIRC3 0.43549527 Sensitive 2.5 uM MS1 Bendamustine CNOT3 0.70329503 Sensitive 2.5 uM MS1 Bendamustine loss_17q11.2 0.43549527 Sensitive 2.5 uM MS1 Bendamustine loss_9p21.3 0.43549527 Sensitive 2.5 uM MS1 Cerdulatinib loss_13q14.13 0.36118573 Sensitive 2.5 uM MS1 Crizotinib RFX7 −0.5099227 Resistant 2.5 uM MS1 Crizotinib

NOTCH1 −0.3921101 Resistant 2.5 uM MS1 Crizotinib IGLV321_R110 −0.3921101 Resistant 2.5 uM MS1 Crizotinib priortrt_Post −0.3729429 Resistant 2.5 uM MS1 Crizotinib ZMYM3 −0.5099227 Resistant 2.5 uM MS1 Crizotinib loss_6q −0.3921101 Resistant 2.5 uM MS1 Crizotinib RPS15 −0.3921101 Resistant 2.5 uM MS1 Dasatinib EC-i −0.4190906 Resistant 2.5 uM MS1 Dasatinib NOTCH1 −0.6405408 Resistant 2.5 uM MS1 Dasatinib IGLV321_R110 −0.6405408 Resistant 2.5 uM MS1 Dasatinib loss_6q −0.6405408 Resistant 2.5 uM MS1 Dasatinib tri_12 0.39018403 Sensitive 2.5 uM MS1 Dasatinib RPS15 −0.6405408 Resistant 2.5 uM MS1 Duvelisib loss_17q23.1 0.35409209 Sensitive 2.5 uM MS1 Duvelisib loss_17p13.1 0.35409209 Sensitive 2.5 uM MS1 Entospletinib loss_19p13.11 0.33669984 Sensitive 2.5 uM MS1 Fludarabine BIRC3 0.44305439 Sensitive 2.5 uM MS1 Fludarabine ITIH2 0.50360592 Sensitive 2.5 uM MS1 Fludarabine loss_17q11.2 0.44305439 Sensitive 2.5 uM MS1 Fludarabine loss_9p21.3 0.44305439 Sensitive 2.5 uM MS1 GSK690693 EC-m3 −0.3847545 Resistant 2.5 uM MS1 GSK690693 i-CLL 0.40885601 Sensitive 2.5 uM MS1 GSK690693 loss_13q14.13 0.37283349 Sensitive 2.5 uM MS1 GSK690693 EC-o 0.33701895 Sensitive 2.5 uM MS1 Ibrutinib loss_19p13.11 0.43395101 Sensitive 2.5 uM MS1 Idelalisib tri_12 0.38216015 Sensitive 2.5 uM MS1 JQ1 EC-m4 0.39012985 Sensitive 2.5 uM MS1 Lenalidomide CNOT3 0.33071669 Sensitive 2.5 uM MS1 MK-2206 EC-i −0.4274473 Resistant 2.5 uM MS1 MK-2206 IGLV321_R110 −0.5268628 Resistant 2.5 uM MS1 MK-2206 loss_6q −0.5268628 Resistant 2.5 uM MS1 MK-2206 RPS15 −0.5268628 Resistant 2.5 uM MS1 Navitoclax loss_8p 0.42290643 Sensitive 2.5 uM MS1 Navitoclax gain_16p11.2 0.3378385 Sensitive 2.5 uM MS1 Navitoclax loss_3p21.31 0.48327152 Sensitive 2.5 uM MS1 Navitoclax loss_14q32.12 0.36089533 Sensitive 2.5 uM MS1 Navitoclax loss_7q36.1 0.42290643 Sensitive 2.5 uM MS1 Navitoclax MYD88 0.4615326 Sensitive 2.5 uM MS1 Nutlin-3 EC-m4 0.33733821 Sensitive 2.5 uM MS1 Onalespib loss_19p13.11 0.4067422 Sensitive 2.5 uM MS1 Osimertinib EC-i −0.3441832 Resistant 2.5 uM MS1 Osimertinib NOTCH1 −0.4871931 Resistant 2.5 uM MS1 Osimertinib IGLV321_R110 −0.4871931 Resistant 2.5 uM MS1 Osimertinib loss_6q −0.4871931 Resistant 2.5 uM MS1 Osimertinib RPS15 −0.4871931 Resistant 2.5 uM MS1 Ponatinib RFX7 −0.7386591 Resistant 2.5 uM MS1 Ponatinib NOTCH1 −0.4169651 Resistant 2.5 uM MS1 Ponatinib IGLV321_R110 −0.4169651 Resistant 2.5 uM MS1 Ponatinib priortrt_Post −0.4869062 Resistant 2.5 uM MS1 Ponatinib FBXW7 −0.3833616 Resistant 2.5 uM MS1 Ponatinib loss_10q24.2 −0.5360884 Resistant 2.5 uM MS1 Ponatinib loss_2q31.1 −0.431849 Resistant 2.5 uM MS1 Ponatinib ZMYM3 −0.7386591 Resistant 2.5 uM MS1 Ponatinib loss_6q −0.4169651 Resistant 2.5 uM MS1 Ponatinib loss_10p12.2 −0.431849 Resistant 2.5 uM MS1 Ponatinib RPS15 −0.4169651 Resistant 2.5 uM MS1 Rapamycin RFX7 −0.6708833 Resistant 2.5 uM MS1 Rapamycin EC-u1 −0.4081365 Resistant 2.5 uM MS1 Rapamycin priortrt_Post −0.5247781 Resistant 2.5 uM MS1 Rapamycin U-CLL −0.4001589 Resistant 2.5 uM MS1 Rapamycin FBXW7 −0.3836273 Resistant 2.5 uM MS1 Rapamycin loss_11q22.3 −0.3838154 Resistant 2.5 uM MS1 Rapamycin loss_10q24.2 −0.4815459 Resistant 2.5 uM MS1 Rapamycin loss_2q31.1 −0.3824336 Resistant 2.5 uM MS1 Rapamycin ZMYM3 −0.6708833 Resistant 2.5 uM MS1 Rapamycin n-CLL −0.5147572 Resistant 2.5 uM MS1 Rapamycin M-CLL 0.4001589 Sensitive 2.5 uM MS1 Rapamycin loss_10p12.2 −0.3824336 Resistant 2.5 uM MS1 Ricolinostat CNOT3 0.46109945 Sensitive 2.5 uM MS1 Ricolinostat loss_10q24.32 −0.3885268 Resistant 2.5 uM MS1 Ricolinostat m-CLL 0.39146931 Sensitive 2.5 uM MS1 Ruxolitinib NOTCH1 −0.4522592 Resistant 2.5 uM MS1 Ruxolitinib IGLV321_R110 −0.4522592 Resistant 2.5 uM MS1 Ruxolitinib loss_6q −0.4522592 Resistant 2.5 uM MS1 Ruxolitinib RPS15 −0.4522592 Resistant 2.5 uM MS1 Sorafenib gain_19p13.3 −0.4424339 Resistant 2.5 uM MS1 Sorafenib RFX7 −0.8770851 Resistant 2.5 uM MS1 Sorafenib priortrt_Post −0.5222341 Resistant 2.5 uM MS1 Sorafenib loss_11q22.3 −0.434434 Resistant 2.5 uM MS1 Sorafenib loss_10q24.2 −0.5202678 Resistant 2.5 uM MS1 Sorafenib loss_2q31.1 −0.4677062 Resistant 2.5 uM MS1 Sorafenib ZMYM3 −0.8770851 Resistant 2.5 uM MS1 Sorafenib loss_10p12.2 −0.4677062 Resistant 2.5 uM MS1 Sorafenib m-CLL 0.34460982 Sensitive 2.5 uM MS1 Sunitinib RFX7 −0.6585943 Resistant

2.5 uM MS1 Sunitinib priortrt_Post −0.4090967 Resistant 2.5 uM MS1 Sunitinib U-CLL −0.3746284 Resistant 2.5 uM MS1 Sunitinib FBXW7 −0.3482133 Resistant 2.5 uM MS1 Sunitinib loss_10q24.2 −0.3670408 Resistant 2.5 uM MS1 Sunitinib loss_2q31.1 −0.3521378 Resistant 2.5 uM MS1 Sunitinib ZMYM3 −0.6585943 Resistant 2.5 uM MS1 Sunitinib n-CLL −0.3690725 Resistant 2.5 uM MS1 Sunitinib M-CLL 0.37462839 Sensitive 2.5 uM MS1 Sunitinib loss_10p12.2 −0.3521378 Resistant 2.5 uM MS1 Trametinib CHD2 0.3824262 Sensitive 2.5 uM MS1 Trametinib EC-i −0.4494912 Resistant 2.5 uM MS1 Trametinib NOTCH1 −0.4082263 Resistant 2.5 uM MS1 Trametinib IGLV321_R110 −0.5391527 Resistant 2.5 uM MS1 Trametinib loss_6q −0.5391527 Resistant 2.5 uM MS1 Trametinib RPS15 −0.5391527 Resistant 2.5 uM MS1 Umbralisib EC-i −0.346901 Resistant 2.5 uM MS1 Umbralisib NOTCH1 −0.3542668 Resistant 2.5 uM MS1 Umbralisib IGLV321_R110 −0.4531376 Resistant 2.5 uM MS1 Umbralisib U-CLL −0.3532312 Resistant 2.5 uM MS1 Umbralisib loss_6q −0.4531376 Resistant 2.5 uM MS1 Umbralisib M-CLL 0.35323115 Sensitive 2.5 uM MS1 Umbralisib gain_7q22.1 0.40480838 Sensitive 2.5 uM MS1 Umbralisib RPS15 −0.4531376 Resistant 2.5 uM MS1 Venetoclax POT1 0.34442157 Sensitive 2.5 uM MS1 Venetoclax loss_13q14.13 0.34896275 Sensitive 2.5 uM MS1 Venetoclax gain_16p11.2 0.3631573 Sensitive 2.5 uM MS1 Voruciclib SAMHD1 −0.3336984 Resistant 2.5 uM MS1 Zanubrutinib gain_19p13.3 −0.3488323 Resistant 2.5 uM MS1 Zanubrutinib RFX7 −0.4526979 Resistant 2.5 uM MS1 Zanubrutinib FBXW7 −0.37558 Resistant 2.5 uM MS1 Zanubrutinib ZMYM3 −0.4526979 Resistant 2.5 uM BAD A-1331852 CNOT3 0.38194078 Sensitive 0.3 uM BAD AZD5991 EC-i 0.37911078 Sensitive 0.3 uM BAD AZD5991 POT1 0.42166522 Sensitive 0.3 uM BAD AZD5991 loss_7p22.2 0.330507 Sensitive 0.3 uM BAD AZD5991 loss_14q32.12 0.40648567 Sensitive 0.3 uM BAD AZD5991 loss_16p11.2 0.55266804 Sensitive 0.3 uM BAD AZD5991 loss_1q42.13 0.53341994 Sensitive 0.3 uM BAD Abexinostat EC-m2 0.39731553 Sensitive 0.3 uM BAD Atorvastatin gain_19p13.3 0.35935709 Sensitive 0.3 uM BAD Atorvastatin RFX7 0.49701967 Sensitive 0.3 uM BAD Atorvastatin EC-u1 0.42618459 Sensitive 0.3 uM BAD Atorvastatin U-CLL 0.35026526 Sensitive 0.3 uM BAD Atorvastatin FBXW7 0.40543892 Sensitive 0.3 uM BAD Atorvastatin ZMYM3 0.49701967 Sensitive 0.3 uM BAD Atorvastatin n-CLL 0.39787953 Sensitive 0.3 uM BAD Atorvastatin M-CLL −0.3502653 Resistant 0.3 uM BAD Atorvastatin m-CLL −0.420056 Resistant 0.3 uM BAD Azacitidine loss_19p13.11 0.39193137 Sensitive 0.3 uM BAD Azacitidine CARD11 0.47557006 Sensitive 0.3 uM BAD Azacitidine MYD88 0.38614039 Sensitive 0.3 uM BAD Bendamustine EC-i 0.36174873 Sensitive 0.3 uM BAD Bendamustine loss_12p13.31a 0.36311374 Sensitive 0.3 uM BAD Bendamustine loss_5p15.33 0.36311374 Sensitive 0.3 uM BAD Bendamustine loss_7p22.2 0.37937529 Sensitive 0.3 uM BAD Bendamustine loss_14q32.12 0.4636513 Sensitive 0.3 uM BAD Bendamustine loss_16p11.2 0.61602839 Sensitive 0.3 uM BAD Bendamustine loss_1q42.13 0.45587516 Sensitive 0.3 uM BAD Cerdulatinib loss_13q14.3 0.48803187 Sensitive 0.3 uM BAD Cerdulatinib loss_13q14.13 0.44644712 Sensitive 0.3 uM BAD Crizotinib RFX7 0.55715299 Sensitive 0.3 uM BAD Crizotinib EC-i −0.3482038 Resistant 0.3 uM BAD Crizotinib EC-u1 0.37851483 Sensitive 0.3 uM BAD Crizotinib NOTCH1 −0.3688823 Resistant 0.3 uM BAD Crizotinib IGLV321_R110 −0.3688823 Resistant 0.3 uM BAD Crizotinib loss_11q22.3 0.36905109 Sensitive 0.3 uM BAD Crizotinib ZMYM3 0.55715299 Sensitive 0.3 uM BAD Crizotinib n-CLL 0.39436071 Sensitive 0.3 uM BAD Crizotinib loss_6q −0.3688823 Resistant 0.3 uM BAD Crizotinib RPS15 −0.3688823 Resistant 0.3 uM BAD Dasatinib NOTCH1 −0.5130779 Resistant 0.3 uM BAD Dasatinib IGLV321_R110 −0.5130779 Resistant 0.3 uM BAD Dasatinib loss_6q −0.5130779 Resistant 0.3 uM BAD Dasatinib RPS15 −0.5130779 Resistant 0.3 uM BAD Duvelisib loss_17q23.1 0.37806892 Sensitive 0.3 uM BAD Duvelisib loss_17p13.1 0.37806892 Sensitive 0.3 uM BAD Entospletinib loss_17q23.1 0.35991923 Sensitive 0.3 uM BAD Entospletinib loss_19p13.11 0.47996356 Sensitive 0.3 uM BAD Entospletinib loss_17p13.1 0.37235435 Sensitive 0.3 uM BAD Erastin EC-i −0.3447734 Resistant 0.3 uM BAD GSK690693 POT1 −0.3915694 Resistant 0.3 uM BAD GSK690693 loss_3p13 −0.3480723 Resistant 0.3 uM

BAD GSK690693 EC-o 0.36161943 Sensitive 0.3 uM BAD Gandotinib EC-i −0.4144924 Resistant 0.3 uM BAD Gandotinib GNB1 0.33154559 Sensitive 0.3 uM BAD JQ1 loss_8p −0.3437573 Resistant 0.3 uM BAD JQ1 gain_19p13.3 0.35896456 Sensitive 0.3 uM BAD JQ1 i-CLL 0.33271819 Sensitive 0.3 uM BAD JQ1 BIRC3 0.63547763 Sensitive 0.3 uM BAD JQ1 loss_12p13.31a −0.3995505 Resistant 0.3 uM BAD JQ1 FBXW7 0.5086376 Sensitive 0.3 uM BAD JQ1 ITIH2 0.63547763 Sensitive 0.3 uM BAD JQ1 loss_5p15.33 −0.3995505 Resistant 0.3 uM BAD JQ1 loss_3p21.31 −0.3437573 Resistant 0.3 uM BAD JQ1 loss_17p13.1 0.50698974 Sensitive 0.3 uM BAD JQ1 loss_14q32.12 −0.3437573 Resistant 0.3 uM BAD JQ1 loss_7q36.1 −0.3437573 Resistant 0.3 uM BAD JQ1 EC-m2 0.38769258 Sensitive 0.3 uM BAD JQ1 loss_17q11.2 0.63547763 Sensitive 0.3 uM BAD JQ loss_9p21.3 0.63547763 Sensitive 0.3 uM BAD JQ1 MYD88 −0.3437573 Resistant 0.3 uM BAD JQ1 loss_17p13.3 −0.3995505 Resistant 0.3 uM BAD Lenalidomide loss_10q24.32 −0.354437 Resistant 0.3 uM BAD MK-2206 EC-i −0.4300923 Resistant 0.3 uM BAD MK-2206 IGLV321_R110 −0.4533833 Resistant 0.3 uM BAD MK-2206 priortrt_Post −0.3325602 Resistant 0.3 uM BAD MK-2206 loss_6q −0.4533833 Resistant 0.3 uM BAD MK-2206 RPS15 −0.4533833 Resistant 0.3 uM BAD Navitoclax loss_8p 0.34214958 Sensitive 0.3 uM BAD Navitoclax EC-m3 0.33802914 Sensitive 0.3 uM BAD Navitoclax EC-u1 −0.3717303 Resistant 0.3 uM BAD Navitoclax loss_12p13.31a 0.33632285 Sensitive 0.3 uM BAD Navitoclax U-CLL −0.4616729 Resistant 0.3 uM BAD Navitoclax loss_5p15.33 0.33632285 Sensitive 0.3 uM BAD Navitoclax n-CLL −0.4049764 Resistant 0.3 uM BAD Navitoclax loss_3p21.31 0.39049159 Sensitive 0.3 uM BAD Navitoclax loss_14q32.12 0.45889323 Sensitive 0.3 uM BAD Navitoclax M-CLL 0.46167289 Sensitive 0.3 uM BAD Navitoclax loss_7q36.1 0.34214958 Sensitive 0.3 uM BAD Navitoclax loss_17p13.3 0.33632285 Sensitive 0.3 uM BAD Nirogacestat loss_18q21.2 −0.3785608 Resistant 0.3 uM BAD Nirogacestat loss_3p13 −0.356381 Resistant 0.3 uM BAD Nirogacestat loss_1p36.11 −0.3785608 Resistant 0.3 uM BAD Nutlin-3 priortrt_Post −0.416614 Resistant 0.3 uM BAD Nutlin-3 U-CLL −0.4093096 Resistant 0.3 uM BAD Nutlin-3 M-CLL 0.40930963 Sensitive 0.3 uM BAD Onalespib loss_13q14.13 0.35335742 Sensitive 0.3 uM BAD Onalespib loss_16p11.2 0.39143469 Sensitive 0.3 uM BAD Onalespib loss_1q42.13 0.35967286 Sensitive 0.3 uM BAD Osimertinib NOTCH1 −0.362555 Resistant 0.3 uM BAD Osimertinib IGLV321_R110 −0.362555 Resistant 0.3 uM BAD Osimertinib loss_6q −0.362555 Resistant 0.3 uM BAD Osimertinib RPS15 −0.362555 Resistant 0.3 uM BAD Osimertinib MYD88 −0.4311851 Resistant 0.3 uM BAD Ponatinib RFX7 0.38527486 Sensitive 0.3 uM BAD Ponatinib EC-i −0.3953926 Resistant 0.3 uM BAD Ponatinib NOTCH1 −0.3559831 Resistant 0.3 uM BAD Ponatinib IGLV321_R110 −0.3559831 Resistant 0.3 uM BAD Ponatinib i-CLL −0.447291 Resistant 0.3 uM BAD Ponatinib ZMYM3 0.38527486 Sensitive 0.3 uM BAD Ponatinib loss_3p13 −0.3534033 Resistant 0.3 uM BAD Ponatinib loss_6q −0.3559831 Resistant 0.3 uM BAD Ponatinib RPS15 −0.3559831 Resistant 0.3 uM BAD Ponatinib m-CLL 0.35855774 Sensitive 0.3 uM BAD Rapamycin RFX7 −0.5061225 Resistant 0.3 uM BAD Rapamycin EC-u1 −0.3689429 Resistant 0.3 uM BAD Rapamycin priortrt_Post −0.4255941 Resistant 0.3 uM BAD Rapamycin U-CLL −0.3794115 Resistant 0.3 uM BAD Rapamycin FBXW7 −0.4413568 Resistant 0.3 uM BAD Rapamycin loss_11q22.3 −0.3321682 Resistant 0.3 uM BAD Rapamycin ZMYM3 −0.5061225 Resistant 0.3 uM BAD Rapamycin n-CLL −0.4134887 Resistant 0.3 uM BAD Rapamycin M-CLL 0.37941148 Sensitive 0.3 uM BAD Ricolinostat CARD11 0.33195138 Sensitive 0.3 uM BAD Ruxolitinib ITIH2 −0.3811264 Resistant 0.3 uM BAD Ruxolitinib loss_13q14.3 0.40676061 Sensitive 0.3 uM BAD Ruxolitinib tri_12 −0.3860581 Resistant 0.3 uM BAD Selinexor ITIH2 0.33151808 Sensitive 0.3 uM BAD Sorafenib RFX7 0.43405693 Sensitive 0.3 uM BAD Sorafenib i-CLL −0.449044 Resistant 0.3 uM BAD Sorafenib ZMYM3 0.43405693 Sensitive 0.3 uM BAD Sorafenib GNB1 0.39982624 Sensitive 0.3 uM BAD Sunitinib RFX7 −0.4492304 Resistant 0.3 uM BAD Sunitinib priortrt_Post −0.3661761 Resistant 0.3 uM BAD Sunitinib U-CLL −0.3404206 Resistant 0.3 uM BAD Sunitinib FBXW7 −0.3307806 Resistant 0.3 uM BAD Sunitinib KLHL6 0.35719433 Sensitive 0.3 uM BAD Sunitinib

loss_11q22.3 −0.3414054 Resistant 0.3 uM BAD Sunitinib loss_2q31.1 −0.3825925 Resistant 0.3 uM BAD Sunitinib ZMYM3 −0.4492304 Resistant 0.3 uM BAD Sunitinib M-CLL 0.34042063 Sensitive 0.3 uM BAD Sunitinib loss_10p12.2 −0.3825925 Resistant 0.3 uM BAD Trametinib CHD2 0.37700492 Sensitive 0.3 uM BAD Trametinib TP53 0.45481052 Sensitive 0.3 uM BAD Umbralisib TP53 0.3904925 Sensitive 0.3 uM BAD Umbralisib NOTCH1 −0.3568164 Resistant 0.3 uM BAD Umbralisib IGLV321_R110 −0.5181869 Resistant 0.3 uM BAD Umbralisib U-CLL −0.3778123 Resistant 0.3 uM BAD Umbralisib KMT2D 0.40655247 Sensitive 0.3 uM BAD Umbralisib loss_6q −0.5181869 Resistant 0.3 uM BAD Umbralisib M-CLL 0.37781234 Sensitive 0.3 uM BAD Umbralisib RPS15 −0.5181869 Resistant 0.3 uM BAD Vecabrutinib NOTCH1 −0.4023258 Resistant 0.3 uM BAD Vecabrutinib IGLV321_R110 −0.3364496 Resistant 0.3 uM BAD Vecabrutinib loss_13q14.3 0.3499221 Sensitive 0.3 uM BAD Vecabrutinib loss_6q −0.3364496 Resistant 0.3 uM BAD Vecabrutinib RPS15 −0.3364496 Resistant 0.3 uM BAD Venetoclax EC-m3 0.36744156 Sensitive 0.3 uM BAD Venetoclax loss_13q14.3 0.44284898 Sensitive 0.3 uM BAD Voruciclib i-CLL 0.43205742 Sensitive 0.3 uM BAD Voruciclib BIRC3 0.36724922 Sensitive 0.3 uM BAD Voruciclib ITIH2 0.36724922 Sensitive 0.3 uM BAD Voruciclib loss_17q11.2 0.36724922 Sensitive 0.3 uM BAD Voruciclib loss_9p21.3 0.36724922 Sensitive 0.3 uM BAD Zanubrutinib BIRC3 0.38964535 Sensitive 0.3 uM BAD Zanubrutinib priortrt_Post −0.3621853 Resistant 0.3 uM BAD Zanubrutinib ITIH2 0.38964535 Sensitive 0.3 uM BAD Zanubrutinib tri_12 0.37718588 Sensitive 0.3 uM BAD Zanubrutinib loss_17p13.1 0.38964535 Sensitive 0.3 uM BAD Zanubrutinib loss_17q11.2 0.38964535 Sensitive 0.3 uM BAD Zanubrutinib loss_9p21.3 0.38964535 Sensitive 0.3 uM BIM A-1331852 EC-u2 0.36150504 Sensitive 0.01 uM BIM A-1331852 CARD11 −0.3550751 Resistant 0.01 uM BIM A-1331852 MYD88 −0.347394 Resistant 0.01 uM BIM AZD5991 loss_18q21.2 0.34858214 Sensitive 0.01 uM BIM AZD5991 loss_15q26.1 0.35512307 Sensitive 0.01 uM BIM AZD5991 loss_3p13 0.37182779 Sensitive 0.01 uM BIM AZD5991 ANK1 0.35512307 Sensitive 0.01 uM BIM AZD5991 loss_1p36.11 0.34858214 Sensitive 0.01 uM BIM Abexinostat loss_17q23.1 0.33522511 Sensitive 0.01 uM BIM Abexinostat priortrt_Post 0.43037204 Sensitive 0.01 uM BIM Abexinostat CNOT3 0.33160882 Sensitive 0.01 uM BIM Abexinostat EC-m2 0.41609565 Sensitive 0.01 uM BIM Acalabrutinib ITIH2 0.38155867 Sensitive 0.01 uM BIM Atorvastatin gain_19p13.3 −0.4238085 Resistant 0.01 uM BIM Atorvastatin RFX7 −0.8943479 Resistant 0.01 uM BIM Atorvastatin EC-u1 −0.3689004 Resistant 0.01 uM BIM Atorvastatin FBXW7 −0.5147083 Resistant 0.01 uM BIM Atorvastatin loss_10q24.2 −0.3883399 Resistant 0.01 uM BIM Atorvastatin loss_2q31.1 −0.3447146 Resistant 0.01 uM BIM Atorvastatin ZMYM3 −0.8943479 Resistant 0.01 uM BIM Atorvastatin n-CLL −0.3923104 Resistant 0.01 uM BIM Atorvastatin loss_10p12.2 −0.3447146 Resistant 0.01 uM BIM Azacitidine gain_19p13.3 −0.3533136 Resistant 0.01 uM BIM Azacitidine BIRC3 −0.3752794 Resistant 0.01 uM BIM Azacitidine FBXW7 −0.373508 Resistant 0.01 uM BIM Azacitidine ITIH2 −0.3752794 Resistant 0.01 uM BIM Azacitidine loss_17q11.2 −0.3752794 Resistant 0.01 uM BIM Azacitidine loss_9p21.3 −0.3752794 Resistant 0.01 uM BIM Carfilzomib loss_17q23.1 0.33868438 Sensitive 0.01 uM BIM Carfilzomib EC-m2 0.35202411 Sensitive 0.01 uM BIM Cerdulatinib loss_13q14.3 0.43528141 Sensitive 0.01 uM BIM Cerdulatinib KMT2D 0.33735471 Sensitive 0.01 uM BIM Cerdulatinib loss_13q14.13 0.43008695 Sensitive 0.01 uM BIM Crizotinib EC-i −0.3441979 Resistant 0.01 uM BIM Crizotinib NOTCHI −0.451415 Resistant 0.01 uM BIM Crizotinib IGLV321_R110 −0.451415 Resistant 0.01 uM BIM Crizotinib SF3B1 0.34464916 Sensitive 0.01 uM BIM Crizotinib loss_6q −0.451415 Resistant 0.01 uM BIM Crizotinib RPS15 −0.451415 Resistant 0.01 uM BIM Dasatinib NOTCH1 −0.5258126 Resistant 0.01 uM BIM Dasatinib IGLV321_R110 −0.5258126 Resistant 0.01 uM BIM Dasatinib loss_17q23.1 0.42597941 Sensitive 0.01 uM BIM Dasatinib loss_6q −0.5258126 Resistant 0.01 uM BIM Dasatinib RPS15 −0.5258126 Resistant 0.01 uM BIM Duvelisib NOTCH1 −0.3445454 Resistant 0.01 uM BIM Duvelisib EC-m2 0.35658655 Sensitive 0.01 uM BIM Entospletinib tri_12 0.43712522 Sensitive 0.01 uM BIM Erastin EC-m1 −0.3427213

Resistant 0.01 uM BIM Erastin ITIH2 −0.3466876 Resistant 0.01 uM BIM GSK690693 i-CLL 0.35522406 Sensitive 0.01 uM BIM Ibrutinib CNOT3 −0.4548753 Resistant 0.01 uM BIM Ibrutinib MYD88 −0.375696 Resistant 0.01 uM BIM JQ1 EC-i 0.38264049 Sensitive 0.01 uM BIM JQ1 priortrt_Post 0.36086327 Sensitive 0.01 uM BIM JQ1 loss_18q21.2 0.43654467 Sensitive 0.01 uM BIM JQ1 loss_3p13 0.33635134 Sensitive 0.01 uM BIM JQ1 loss_10q24.32 0.43654467 Sensitive 0.01 uM BIM JQ1 loss_1p36.11 0.43654467 Sensitive 0.01 uM BIM Lenalidomide loss_15q15.1b −0.3452293 Resistant 0.01 uM BIM Lenalidomide loss_20p11.22 −0.3452293 Resistant 0.01 uM BIM MK-2206 BIRC3 0.35328124 Sensitive 0.01 uM BIM MK-2206 ITIH2 0.35328124 Sensitive 0.01 uM BIM MK-2206 loss_17q11.2 0.35328124 Sensitive 0.01 uM BIM MK-2206 loss_9p21.3 0.35328124 Sensitive 0.01 uM BIM Navitoclax loss_8p 0.39508636 Sensitive 0.01 uM BIM Navitoclax loss_3p21.31 0.41463633 Sensitive 0.01 uM BIM Navitoclax loss_14q32.12 0.38220577 Sensitive 0.01 uM BIM Navitoclax loss_7q36.1 0.39508636 Sensitive 0.01 uM BIM Nirogacestat tri_12 0.33700499 Sensitive 0.01 uM BIM Nirogacestat EC-m2 0.41295567 Sensitive 0.01 uM BIM Onalespib loss_15q15.1b −0.3983134 Resistant 0.01 uM BIM Onalespib loss_20p11.22 −0.3983134 Resistant 0.01 uM BIM Osimertinib NOTCH1 −0.3688148 Resistant 0.01 uM BIM Osimertinib IGLV321_R110 −0.3688148 Resistant 0.01 uM BIM Osimertinib EC-u2 0.3685819 Sensitive 0.01 uM BIM Osimertinib BIRC3 0.40211777 Sensitive 0.01 uM BIM Osimertinib loss_6q −0.3688148 Resistant 0.01 uM BIM Osimertinib loss_17q11.2 0.40211777 Sensitive 0.01 uM BIM Osimertinib loss_9p21.3 0.40211777 Sensitive 0.01 uM BIM Osimertinib RPS15 −0.3688148 Resistant 0.01 uM BIM Ponatinib RFX7 −0.4535092 Resistant 0.01 uM BIM Ponatinib EC-i −0.4254906 Resistant 0.01 uM BIM Ponatinib NOTCH1 −0.4928195 Resistant 0.01 uM BIM Ponatinib IGLV321_R110 −0.4928195 Resistant 0.01 uM BIM Ponatinib i-CLL −0.3857108 Resistant 0.01 uM BIM Ponatinib priortrt_Post −0.3481786 Resistant 0.01 uM BIM Ponatinib U-CLL −0.3892403 Resistant 0.01 uM BIM Ponatinib loss_2q31.1 −0.3448179 Resistant 0.01 uM BIM Ponatinib ZMYM3 −0.4535092 Resistant 0.01 uM BIM Ponatinib loss_6q −0.4928195 Resistant 0.01 uM BIM Ponatinib M-CLL 0.38924027 Sensitive 0.01 uM BIM Ponatinib loss_10p12.2 −0.3448179 Resistant 0.01 uM BIM Ponatinib RPS15 −0.4928195 Resistant 0.01 uM BIM Ponatinib m-CLL 0.44631549 Sensitive 0.01 uM BIM Rapamycin RFX7 −0.590868 Resistant 0.01 uM BIM Rapamycin FBXW7 −0.4127982 Resistant 0.01 uM BIM Rapamycin ZMYM3 −0.590868 Resistant 0.01 uM BIM Rapamycin n-CLL −0.3681718 Resistant 0.01 uM BIM Ricolinostat loss_15q15.1b −0.3481502 Resistant 0.01 uM BIM Ricolinostat i-CLL −0.3799749 Resistant 0.01 uM BIM Ricolinostat loss_20p11.22 −0.3481502 Resistant 0.01 uM BIM Selinexor MYD88 0.3956828 Sensitive 0.01 uM BIM Sorafenib gain_19p13.3 −0.340448 Resistant 0.01 uM BIM Sorafenib RFX7 −0.6413425 Resistant 0.01 uM BIM Sorafenib FBXW7 −0.3976621 Resistant 0.01 uM BIM Sorafenib loss_2q31.1 −0.3972096 Resistant 0.01 uM BIM Sorafenib ZMYM3 −0.6413425 Resistant 0.01 uM BIM Sorafenib loss_10p12.2 −0.3972096 Resistant 0.01 uM BIM Sorafenib GNB1 0.39242101 Sensitive 0.01 uM BIM Sunitinib gain_19p13.3 −0.350848 Resistant 0.01 uM BIM Sunitinib RFX7 −0.7282073 Resistant 0.01 uM BIM Sunitinib loss_2q31.1 −0.4921048 Resistant 0.01 uM BIM Sunitinib ZMYM3 −0.7282073 Resistant 0.01 uM BIM Sunitinib n-CLL −0.3771445 Resistant 0.01 uM BIM Sunitinib loss_10p12.2 −0.4921048 Resistant 0.01 uM BIM Trametinib CHD2 0.56462878 Sensitive 0.01 uM BIM Trametinib loss_14q32.33 −0.3778377 Resistant 0.01 uM BIM Umbralisib NOTCH1 −0.3908108 Resistant 0.01 uM BIM Umbralisib U-CLL −0.410329 Resistant 0.01 uM BIM Umbralisib EC-m4 0.35101961 Sensitive 0.01 uM BIM Umbralisib M-CLL 0.41032904 Sensitive 0.01 uM BIM Venetoclax SAMHD1 0.33953418 Sensitive 0.01 uM BIM Vorinostat EC-m4 −0.4033298 Resistant 0.01 uM BIM Vorinostat EC-m2 0.35472621 Sensitive 0.01 uM BIM Voruciclib BCOR −0.4055636 Resistant 0.01 uM BIM Voruciclib n-CLL −0.337222 Resistant 0.01 uM BIM Zanubrutinib DICER1 −0.341418 Resistant 0.01 uM BIM Zanubrutinib loss_14q32.33 0.40971541 Sensitive 0.01 uM BIM Zanubrutinib MED1 −0.341418 Resistant 0.01 uM BIM Zanubrutinib EC-m4 −0.3738111 Resistant 0.01 uM BIM Zanubrutinib

tri_12 0.44636509 Sensitive 0.01 uM BIM Zanubrutinib EC-m2 0.33298948 Sensitive 0.01 uM
[0295] Agents of the present disclosure may be administered within a pharmaceutically-acceptable diluents, carrier, or excipient, in unit dosage form. Conventional pharmaceutical practice may be employed to provide suitable formulations or compositions to administer the compounds to patients suffering from a disease that is caused by excessive cell proliferation (e.g., CLL). Administration may begin before the patient is symptomatic. Any appropriate route of administration may be employed, for example, administration may be parenteral, intravenous, intraarterial, subcutaneous, intratumoral, intramuscular, intracranial, intraorbital, ophthalmic, intraventricular, intrahepatic, intracapsular, intrathecal, intracisternal, intraperitoneal, intranasal, aerosol, suppository, or oral administration. For example, therapeutic formulations may be in the form of liquid solutions or suspensions; for oral administration, formulations may be in the form of tablets or capsules; and for intranasal formulations, in the form of powders, nasal drops, or aerosols.
[0296] Methods well known in the art for making formulations are found, for example, in "Remington: The Science and Practice of Pharmacy" Ed. A. R. Gennaro, Lippincourt Williams & Wilkins, Philadelphia, Pa., 2000. Formulations for parenteral administration may, for example, contain excipients, sterile water, or saline, polyalkylene glycols such as polyethylene glycol, oils of vegetable origin, or hydrogenated napthalenes. Biocompatible, biodegradable lactide polymer, lactide/glycolide copolymer, or polyoxyethylene-polyoxypropylene copolymers may be used to control the release of the compounds. Other potentially useful parenteral delivery systems for agents of the present disclosure include ethylene-vinyl acetate copolymer particles, osmotic pumps, implantable infusion systems, and liposomes. Formulations for inhalation may contain excipients, for example, lactose, or may be aqueous solutions containing, for example, polyoxyethylene-9-lauryl ether, glycocholate and deoxycholate, or may be oily solutions for administration in the form of nasal drops, or as a gel.
[0297] The formulations can be administered to human patients in therapeutically effective amounts (e.g., amounts which prevent, eliminate, or reduce a pathological condition) to provide therapy for a neoplastic disease or condition (e.g., chronic lymphocytic leukemia). The preferred dosage of a nucleobase oligomer of the disclosure is likely to depend on such variables as the type and extent of the disorder, the overall health status of the particular patient, the formulation of the compound excipients, and its route of administration.
[0298] With respect to a subject having chronic lymphocytic leukemia (CLL), an effective amount is sufficient to stabilize, slow, or reduce the proliferation of CLL. Generally, doses of active polynucleotide compositions of the present disclosure would be from about 0.01 mg/kg per day to about 1000 mg/kg per day. It is expected that doses ranging from about 50 to about 2000 mg/kg will be suitable. Lower doses will result from certain forms of administration, such as intravenous administration. In the event that a response in a subject is insufficient at the initial doses applied, higher doses (or effectively higher doses by a different, more localized delivery route) may be employed to the extent that patient tolerance permits. Multiple doses per day are contemplated to achieve appropriate systemic levels of an agent and/or compositions of the present disclosure.
[0299] A variety of administration routes are available. The methods disclosed herein, generally speaking, may be practiced using any mode of administration that is medically acceptable, meaning any mode that produces effective levels of the active compounds without causing clinically unacceptable adverse effects. Other modes of administration include oral, rectal, topical, intraocular, buccal, intravaginal, intracisternal, intracerebroventricular, intratracheal, nasal, transdermal, within/on implants, e.g., fibers such as collagen, osmotic pumps, or grafts comprising appropriately transformed cells, etc., or parenteral routes.
Kits
[0300] In another aspect, the disclosure provides kits for aiding in patient selection for treatment and/or characterizing chronic lymphocytic leukemia (e.g., selecting a treatment method for a subject, selection of a subject for a clinical trial, predicting clinical outcome, and the like), which

kits are used to detect biomarkers according to the disclosure. In an embodiment, the kit comprises a drug for use in treatment of chronic lymphocytic leukemia (e.g., an agent listed in Table 1A or 2A). In some instances, the kit comprises reagents for collecting a sample from a patient and sequencing RNA from the sample (e.g., RNA-seq). In one embodiment, the kit comprises agents that specifically recognize the biomarkers identified in Tables 3A and 4, or a sub-set thereof. In another embodiment, the kit comprises agents for use in detecting the biomarkers identified in Tables 3A and 4, or a subset thereof. In related embodiments, the agents are antibodies or probes (e.g., oligonucleotides). The kit may contain about or at least about 1, 2, 3, 4, 5, 10, 50, 100, 110, 120, 130, 140, 150, 200 or more different antibodies and/or probes that each specifically recognize one of the biomarkers set forth in Tables 3A and 4.

[0301] In another embodiment, the kit comprises a solid support, such as a chip, a microtiter plate or a bead or resin having capture reagents attached thereon, wherein the capture reagents bind the biomarkers of the disclosure. In the case of biospecfic capture reagents, the kit can comprise a solid support with a reactive surface, and a container comprising the biospecific capture reagents.

[0302] The kit can also comprise a washing solution or instructions for making a washing solution, in which the combination of the capture reagent and the washing solution allows capture of the biomarker or biomarkers on the solid support for subsequent detection by, e.g., mass spectrometry. The kit may include more than one type of adsorbent, each present on a different solid support.

[0303] In a further embodiment, such a kit can comprise instructions for use in any of the methods described herein. In some instances, the kit comprises drug sensitivity information for chronic lymphocytic leukemias (CLLs) having different expression subtypes. The drug sensitivity data is provided in some embodiments along with instructions for selecting a patient for administration of a drug (e.g., an agent listed in Table 1A or 2A) based upon an expression subtype of a chronic lymphocytic leukemia (CLL) in the subject. In embodiments, the instructions provide suitable operational parameters in the form of a label or separate insert. For example, the instructions may inform a consumer about how to collect the sample, how to wash the probe, and/or the particular biomarkers to be detected.

[0304] In yet another embodiment, the kit can comprise one or more containers with controls (e.g., biomarker samples) to be used as standard(s) for calibration.

[0305] The practice of the present disclosure employs, unless otherwise indicated, conventional techniques of molecular biology (including recombinant techniques), microbiology, cell biology, biochemistry and immunology, which are well within the purview of the skilled artisan. Such techniques are explained fully in the literature, such as, “Molecular Cloning: A Laboratory Manual”, second edition (Sambrook, 1989); “Oligonucleotide Synthesis” (Gait, 1984); “Animal Cell Culture” (Freshney, 1987); “Methods in Enzymology” “Handbook of Experimental Immunology” (Weir, 1996); “Gene Transfer Vectors for Mammalian Cells” (Miller and Calos, 1987); “Current Protocols in Molecular Biology” (Ausubel, 1987); “PCR: The Polymerase Chain Reaction”, (Mullis, 1994); “Current Protocols in Immunology” (Coligan, 1991). These techniques are applicable to the production of the polynucleotides and polypeptides disclosed herein, and, as such, may be considered in making and practicing the disclosure. Particularly useful techniques for particular embodiments will be discussed in the sections that follow.

[0306] The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how to make and use the assay, screening, and therapeutic methods of the disclosure, and are not intended to limit the scope of what the inventors regard as their invention.

EXAMPLES

[0307] By applying Bayesian non-negative matrix factorization for unsupervised clustering of RNA-seq data from 610 treatment-naive CLL samples, 8 robust expression clusters (ECs) were identified. The expression clusters (ECs) strongly associated with IGHV (heavy chain variable region of immunoglobulin genes) mutational status and/or epitype, revealed two subtypes of U-

CLL/n-CLL (EC-ul, EC-u2) and four subtypes of M-CLL/m-CLL (EC-m1, EC-m2, EC-m3, and EC-m4) (Tables 3A, 3B, and 4). EC-i was best defined by the i-CLL epitype whereas EC-o, the smallest cluster (n=24; 3.9%), was not significantly associated with any previously defined CLL group. Both EC-i and EC-o displayed borderline identity of somatic hypermutations in IGHV (heavy chain variable region of immunoglobulin genes) with germline, close to the 98% threshold distinguishing M-CLL (CLL with mutated IGHV) from U-CLL (CLL with unmutated IGHV) [0308] Tables 3, 3B, and 4 relate to the Expression cluster (EC) analysis.

TABLE-US-00006 TABLE 3A Expression cluster (EC) marker genes determined by non-negative matrix factorization ENTREZ GENE UNIPROT GENE GENE EC DIRECTION RANK SYNONYM ID ID TFEC EC-m1 UP 1 O14948 22797 COL18A1 EC-m1 UP 2 P39060 80781 SLC19A1 EC-m1 UP 3 P41440 6573 NRIP1 EC-m1 UP 4 P48552 8204 KCNH2 EC-m1 UP 5 Q12809 3757 SEPT10 EC-u1 UP 1 LDOC1 EC-u1 UP 2 O95751 23641 LPL EC-u1 UP 3 P06858 4023 KANK2 EC-u1 UP 4 Q63ZY3 25959 SOWAHC EC-u1 UP 5 Q53LP3 65124 DUSP26 EC-u1 UP 6 Q9BV47 78986 OSBPL5 EC-u1 UP 7 Q9H0X9 114879 WNT9A EC-u1 UP 8 O14904 7483 FGFR1 EC-u1 UP 9 P11362 2260 GTSF1L EC-u1 UP 10 Q9H1H1 149699 EML6 EC-m2 UP 1 Q6ZMW3 400954 HCK EC-m2 UP 2 P08631 3055 CD1C EC-m2 UP 3 P29017 911 VPS37B EC-m2 UP 4 Q9H9H4 79720 CYBB EC-m2 UP 5 P04839 1536 NXPH4 EC-m2 UP 6 O95158 11247 BTNL9 EC-m2 UP 7 Q6UXG8 153579 KLRK1 EC-m2 UP 8 P26718 100528032 IQSEC1 EC-m2 UP 9 Q6DN90 9922 BANK1 EC-m2 UP 10 Q8NDB2 55024 ACSM3 EC-o UP 1 Q53FZ2 6296 TOX2 EC-o UP 2 Q96NM4 84969 PHF16 EC-o UP 3 JADE3 SESN3 EC-o UP 4 P58005 143686 ITGB5 EC-u2 UP 1 P18084 3693 BCL7A EC-u2 UP 2 Q4VC05 605 PPP1R9A EC-u2 UP 3 Q9ULJ8 55607 TSPAN13 EC-u2 UP 4 O95857 27075 SLC12A7 EC-u2 UP 5 Q9Y666 10723 SSBP3 EC-u2 UP 6 Q9BWW4 23648 VASH1 EC-u2 UP 7 Q7L8A9 22846 SPG20 EC-u2 UP 8 SPART IL13RA1 EC-u2 UP 9 P78552 3597 NR3C2 EC-u2 UP 10 P08235 4306 MS4A4E EC-m3 UP 1 Q96PG1 643680 MYL9 EC-m3 UP 2 P24844 10398 NT5E EC-m3 UP 3 P21589 4907 MS4A6A EC-m3 UP 4 Q9H2W1 64231 PITPNC1 EC-m3 UP 5 Q9UKF7 26207 CNTNAP2 EC-m3 UP 6 Q9UHC6 26047 IGF2BP3 EC-m3 UP 7 O00425 10643 WNT3 EC-m3 UP 8 P56703 101929777 CLDN7 EC-m3 UP 9 O95471 1366 TCF7 EC-m3 UP 10 P36402 6932 MYBL1 EC-m4 UP 1 P10243 4603 NUGGC EC-m4 UP 2 Q68CJ6 389643 GNG8 EC-m4 UP 3 Q9UK08 94235 GRIK3 EC-i UP 1 Q13003 2899 IQGAP2 EC-i UP 2 Q13576 10788 FCER1G EC-i UP 3 P30273 2207 STK32B EC-i UP 4 Q9NY57 55351 GADD45A EC-i UP 5 P24522 1647 P2RX1 EC-m1 DOWN 1 P51575 5023 ARRDC5 EC-m1 DOWN 2 A6NEK1 645432 BEX4 EC-m1 DOWN 3 Q9NWD9 56271 APP EC-m1 DOWN 4 P05067 351 ADD3 EC-u1 DOWN 1 Q9UEY8 120 AKT3 EC-u1 DOWN 2 Q9Y243 10000 COBLL1 EC-u1 DOWN 3 Q53SF7 22837 MNDA EC-u1 DOWN 4 P41218 4332 FCRL3 EC-u1 DOWN 5 Q96P31 115352 FAM49A EC-u1 DOWN 6 CYRIA FCRL2 EC-u1 DOWN 7 Q96LA5 79368 SLC2A3 EC-u1 DOWN 8 P11169 6515 MARCKS EC-u1 DOWN 9 P29966 4082 LEF1 EC-m2 DOWN 1 Q9UJU2 51176 SH3D21 EC-m2 DOWN 2 A4FU49 79729 FMOD EC-m2 DOWN 3 Q06828 2331 SEMA4A EC-m2 DOWN 4 Q9H3S1 64218 CTLA4 EC-m2 DOWN 5 P16410 1493 ADTRP EC-m2 DOWN 6 Q96IZ2 84830 IGSF3 EC-m2 DOWN 7 O75054 3321 IGFBP4 EC-m2 DOWN 8 P22692 3487 PDGFD EC-m2 DOWN 9 Q9GZP0 80310 APOD EC-m2 DOWN 10 P05090 347 TBC1D9 EC-o DOWN 1 Q6ZT07 23158 PIP5K1B EC-o DOWN 2 O14986 8395 SIK1 EC-o DOWN 3 P57059 150094 DUSP5 EC-o DOWN 4 Q16690 1847 GNG7 EC-o DOWN 5 O60262 2788 HIVEP3 EC-o DOWN 6 Q5T1R4 59269 MARCKSL1 EC-o DOWN 7 P49006 65108 GPR183 EC-o DOWN 8 P32249 1880 HRK EC-o DOWN 9 O00198 8739 PITPNC1 EC-o DOWN 10 Q9UKF7 26207 TUBG2 EC-u2 DOWN 1 Q9NRH3 27175 ZNF804A EC-u2 DOWN 2 Q7Z570 91752 IL2RA EC-u2 DOWN 3 P01589 3559 BASP1 EC-m3 DOWN 1 P80723 10409 FLJ20373 EC-m3 DOWN 2 MAP4K4 EC-m3 DOWN 3 O95819 9448 LRRK2 EC-m3 DOWN 4 Q5S007 120892 SAMSN1 EC-m3 DOWN 5 Q9NSI8 388813 CEACAM1 EC-m3 DOWN 6 P13688 634 TNFRSF13B EC-m3 DOWN 7 O14836 23495 PHF16 EC-m3 DOWN 8 JADE3 MID1IP1 EC-m3 DOWN 9 Q9NPA3 58526

ABCA9 EC-m3 DOWN 10 Q8IUA7 10350 AEBP1 EC-m4 DOWN 1 Q8IUX7 165 HIP1R EC-m4 DOWN 2 O75146 9026 LATS2 EC-m4 DOWN 3 Q9NRM7 26524 RIMKLB EC-m4 DOWN 4 Q9ULI2 57494 EML6 EC-m4 DOWN 5 Q6ZMW3 400954 FADS3 EC-m4 DOWN 6 Q9Y5Q0 3995 MBOAT1 EC-m4 DOWN 7 Q6ZNC8 154141 LCN10 EC-m4 DOWN 8 Q6JVE6 414332 DCLK2 EC-m4 DOWN 9 Q8N568 166614 GLUL EC-m4 DOWN 10 P15104 2752 ITGAX EC-i DOWN 1 P20702 3687 KLF3 EC-i DOWN 2 P57682 51274 RFTN1 EC-i DOWN 3 Q14699 23180 PTK2 EC-i DOWN 4 Q05397 5747 DFNB31 EC-i DOWN 5 ZMAT1 EC-i DOWN 6 Q5H9K5 84460

TABLE-US-00007 TABLE 3B Legend for Table 3A Column Description GENE Marker gene EC Expression cluster RANK Rank of marker gene per this EC GENE Gene symbol synonym (more SYNONYM updated) UNIPROT ID Uniprot protein accession ID ENTREZ Entrez database gene accession GENE ID ID

TABLE-US-00008 TABLE 4 Biomarkers used in expression cluster (EC) classifier BATCH CORRECTED TRANSCRIPTS TRANSCRIPTS PER MILLION PER MILLION (TPM) (TPM) CLASSIFIER CLASSIFIER BIOMARKERS BIOMARKERS ABCA9 ACAP3 ACAP3 ACSM3 ACSM3 AEBP1 ADAP2 AKT3 AF127936.7 ARHGAP33 ARHGAP33 ARHGAP42 ARMC7 ARMC7 ARRDC5 ARRDC5 ARSD ATPIF1 ARSI BACH2 ASB2 BASP1 ATP1A3 BCL7A ATP2B1 C17orf100 ATPIF1 CBLB BASP1 CD72 BCL2A1 CD86 BCL7A CEACAM1 BCS1L CHPT1 CAMK2A CLDN7 CLDN23 CMTM7 CMTM7 CNTNAP1 COBLL1 COBLL1 CRELD2 COL18A1 CRY1 CRY1 CTAGE9 CTLA4 CTLA4 EGR3 DDR1 EML6 DKFZP761J1410 EZH2 DPF3 FADS3 EML6 FCER1G ERRFI1 FCRL2 ESPNL FGL2 EZH2 FLJ20373 FAHD2B FMOD FAM109A GADD45A FBXO27 GLIPR1 FGL2 GNB4 FLJ20373 GPR160 FMOD GPR34 GADD45A GRIK3 GNAO1 GUCD1 GPR160 HCK GPR34 HIP1R GUCD1 HIVEP3 HCK HMCES HDAC4 IGF2BP3 HIP1R IGSF3 HMCES IL21R IGSF3 INPP5F IQSEC1 IQGAP2 ITGAX IQSEC1 KCNH3 ITGAX KCNN3 ITGB5 KCTD3 JDP2 KDM1B KANK2 KLK1 KCNH2 KSR1 KDM1B LCN10 KLF3 LINC00865 LATS2 LPL LCN10 LRRK2 LEF1 LUZP1 LPL MAP4K4 LRRK2 MAPK4 LUZP1 MAST4 MAP4K4 MPRIP MID1IP1 MRO MMP14 MSI2 MPRIP MVB12B MSI2 MYBL1 MYBL1 MYC MYL9 MYL5 MYLIP MYL9 MZB1 MYO3A NBPF3 NEDD9 NRIP1 NFKBIZ NRSN2 NR2F6 NUGGC NRIP1 NXPH4 NRSN2 P2RX1 NUGGC P2RX5 P2RX1 P2RY14 PELI3 PDGFD PIGB PIP5K1B PIP5K1B PITPNC1 PITPNC1 PON2 PLD1 PRICKLE1 PTPN7 PTPN7 QDPR RCN3 REPS2 RDX RHBDF2 RHBDF2 RIMKLB RIMKLB RP11-134N1.2 RNF135 RP11-265P11.1 RP11-145M9.4 RP11-453F18.sub.——B.1 RP11-268J15.5 RP11-456H18.2 RP11-463O12.3 RP1-90J20.12 RP5-1028K7.2 SAMSN1 SAMSN1 SCPEP1 SCCPDH SH3D21 SCD SLC44A1 SCPEP1 SLC4A7 SDC3 SLC4A8 SECTM1 SMIM10 SESN3 SPN SH3BP2 SSBP3 SH3D21 STAM SLC16A5 STX5 SLC19A1 SYNGR3 SLC4A7 TAS1R3 SPN TBC1D2B SSBP3 TBC1D9 STX5 TFEC SUSD1 TIMELESS TBC1D2B TNFRSF13B TBC1D9 TNR TBKBP1 TOX2 TCF7 TRIM7 TFEC TUBG2 TGFBR3 VSIG10 TIGIT WNT5A TIMELESS ZMYND8 TMEM133 ZNF804A TNFRSF13B TOX2 TRAK2 TTC39C TUBG2 VPS37B VSIG10 WNT9A ZAP70 ZNF667-AS1 ZNF804A ZSWIM6

[0309] The top upregulated marker genes in EC-ul included SEPT10 and LPL. Another upregulated EC-ul gene, OSBPL5, is likely a top expression marker predicting shorter time to progression after treatment with fludarabine, cyclophosphamide, and rituximab.

[0310] Although EC-o was not associated with IGHV (heavy chain variable region of immunoglobulin genes) status or epitype, it was defined by enrichment in oxidative phosphorylation signaling relative to the other expression clusters (ECs) The EC-m clusters were distinguished by either upregulated or downregulated inflammatory signaling or antigen expression via nonclassical HLA class I. The EC-u clusters shared gene expression changes reflecting impaired protein translation, but were differentiated by TNFα signaling, which was low in EC-ul and high in EC-u2. EC-i was enriched for pathways regulating migration and the humoral immune response, possibly reflecting the autonomous BCR signaling of IGLV3-21.sup.R110. Finally, the

epiCMIT scores of the expression clusters (ECs) within each epitype were compared. In EC-m clusters, EC-m3 had a lower epiCMIT relative to the other ECs, consistent with a lower proliferative history and suggestive of better patient outcomes. Multivariable analysis that included clinical features and IGHV (heavy chain variable region of immunoglobulin genes) status confirmed independent prognostic impact of the expression clusters (ECs) on failure free survival (FFS) (n=609, p<0.001) and overall survival (OS) (p=0.012) The EC-u clusters had similarly short failure free survival (FFS) and EC-i displayed intermediate failure free survival outcomes in EC-m clusters were distinct where EC-m1, EC-m2, and EC-m4 demonstrated shorter failure free survival (FFS) relative to EC-m3,

Example 1: Drug Treatment Assignment for Chronic Lymphocytic Leukemia (CLL)

[0311] To address the challenge of selecting a proper treatment for a chronic lymphocytic leukemia, experiments were undertaken to optimize high-throughput dynamic BH3 profiling (HT-DBP) for evaluation of the drug sensitivities of CLL samples (FIGS. **1**, **13**, **5**A, **5**B, **6**A, **6**B, and **7**A-**7**C). HT-DBP was optimized as a functional assay that rapidly measured the initiation of apoptotic signaling after ex vivo exposure to drugs for interrogation of CLL samples. Some advantages of the optimized assay were: (i) rapidity—under 24 hours, which is especially important in CLL, where cell viability substantially decreases after 24 hours; (ii) miniaturization—a very limited number of primary cells were required; and (iii) scalability—allowing to conduct hundreds of drug response tests in parallel on one 384-well plate (FIG. **3**). These features collectively maximized the information yield from a given patient sample.

[0312] Prognostic genetic alterations and molecular subtypes of CLL, based on multiomic profiling of>1100 CLL samples have been characterized (Nature Genetics volume 54, pages1664-1674 (2022) the disclosure of which is incorporated herein by reference in its entirety for all purposes). To determine if these molecular findings were associated with novel therapeutic vulnerabilities in CLL, HT-DBP was performed on 65 primary CLL samples previously characterized by exome, transcriptome and methylome profiling (FIGS. **2**A and **2**B), were evaluated using 42 FDA approved drugs that were selected for potential relevance to CLL biology (see Tables 1A and 1B). Peripheral blood mononuclear cells (PBMCs) were isolated and cultured in conditioned media derived from stroma cells to reduce spontaneous apoptosis. Target cells were treated with a drug for 20 hours followed by BH3 peptide exposure. Mitochondrial outer membrane permeabilization (MOMP) was then measured on digitonin-permeabilized cells in response to BH3-only synthetic peptides that mimic pro-apoptotic BCL-2 family proteins. Mitochondrial cytochrome c release was quantified as a measure of MOMP by flow cytometry, gating on CD19+ and CD5+cells. This assay measured if a cell had moved closer to the threshold of apoptosis after drug treatment and thereby identified drugs that enhanced apoptosis priming. The peptides used in each experiment were derived from BIM or PUMA to measure increases in overall apoptotic priming, or BAD and MS1 peptides that identified BCL-2 and MCL-1 dependence, respectively (see FIG. **4**).

[0313] High-throughput dynamic BH3 priming (HT-DBP) results showed high quality and reproducibility (FIGS. **5**A and **5**B). The different BH3 peptides used in HT-DBP showed similar effects.

[0314] The HT-DBP screen revealed differential drug-induced apoptotic priming for various drugs (FIGS. **7**A-**7**C, **8**A, **8**B, **9**, **12**A, **12**B, and **12**C). Venetoclax and ibrutinib, current first-line treatments for CLL, were highly effective across CLL subtypes. Other drugs that demonstrated high priming included navitoclax (BCL-XL/BCL-2), nutlin-3 (MDM2), abexinostat (HDAC), gandotinib (JAK2), duvelisib (PI3Kδ/γ), idelalisib (PI3K6) and cerdulatinib (SYK/JAK). The assay was robust, as indicated by an 0.92 median Pearson correlation across replicates (FIGS. **5**A, **5**B, **6**A, and **6**B). Additionally, the majority of drugs had greater effect on CLL samples than on healthy PBMCs (p<0.001, paired t-test), supporting their specificity (FIGS. **7**A-**7**C).

[0315] The analysis of the HT-DBP screen data focused on the differential drug effects among molecular subtypes of CLL (FIGS. **9**, **12**A, **12**B, and **12**C). First, it was found that IGHV-mutated

CLLs (M-CLLs) became more primed to apoptosis than IGHV-unmutated CLLs (U-CLLs) across the panel of drugs ($p<0.001$, paired t-test) and that IGHV mutated CLLs had significant response to fludarabine and umbralisib (FDR<0.1, t-test). Second, drug-induced apoptotic priming was compared among 8 CLL subtypes (i.e., RNA expression clusters (ECs); Knisbacher et al. and PCT/US2021/045144, filed Aug. 9, 2021, the disclosures of which are incorporated herein by reference in their entireties for all purposes) (FIGS. **12**A, **12**B, and **12**C). Notable among the many drug priming-EC relationships observed was that within M-CLL ECs, venetoclax was most effective in EC-m3 (high IL-10 expression) and least effective in EC-m2 (low IL-10 and enriched in trisomy 12). Interestingly, EC-m1, which was associated with high TFEC expression and poor outcome, was most sensitive to nutlin-3. For U-CLLs, EC-ul was most sensitive to gandotinib and EC-u2 to navitoclax. EC-i, which was associated with the 20 intermediate methylation subtype of CLL, was the most resistant EC to ibrutinib but was more sensitive to navitoclax than any other drug. Additionally, tri(12) sample groups were observed to be sensitive to treatment via Zanubrutinib and Acalabrutinib, while FBXW7 sample groups exhibited resistance to Zanubrutinib (see FIG. **26**, which provides a table identifying kinase inhibitor drug sensitivities for different peptide concentrations and driver alterations).

[0316] Delta-priming was measured for the molecular features listed along the top of FIGS. **14**, **17**, and **32** using the HT-DBP screen. The molecular features included IGHV subtypes, epitypes, expression subtypes (i.e., expression clusters), mutations in driver genes, and recurrent copy-number events. A feature was included in the heatmap of FIG. **14** only if at least 2 patients in a DBP screen had the feature. Median delta-priming was computed across all BH3 peptides and across all patients within the feature. FIG. **17** shows median delta priming values for healthy donors per normal cell type, and FIG. **18** shows median delta priming values calculated by subtracting median delta priming values for normal cell types from the delta priming values. Plots showing molecular features for groups of 65 and 81 patients were also compiled (see FIGS. **19**, **20**).

[0317] Additionally, a published dataset of 136 CLL patients with RNA-seq whose samples were screened with 63 drugs was used to compile a plot showing differential drug sensitivity of several expression clusters of interest (see FIG. **23**). Furthermore, a heatmap and dendrogram were compiled in FIGS. **21**A-**21**B, showing significant peptide effect similarity for four peptide groups (PUMA 1 μM, BIM 0.01 μM, BAD 0.3 μM and MSI 2.5 μM).

[0318] A comparison of z-values for dynamic BH3 priming data and cell viability data gathered for chronic lymphocytic leukemia (CLL) cells treated with 13 drugs selected from the 42 listed in Table A2 (see FIGS. **15**, **16**) revealed that the BH3 priming data provided insights into drug efficacy that were not previously available using cell viability data alone. Because the different peptides used could be relatively more promiscuous or more selective for the anti-apoptotic proteins being targeted, use of each of the peptides provided different information with regard to a drug's impact on a CLL cell (see FIG. **29**).

[0319] FIG. **11** showed, together with FIGS. **10**A, **10**B, 1° C., and 1OD that the expression clusters were distinguished by molecular features and drivers.

[0320] Additionally, comparison of CLL sample groups and normal peripheral blood mononuclear cells (PBMCs) revealed that drugs such as Navitoclax and Venetoclax resulted in potentially therapeutically relevant priming in CLL groups, while generating minimal priming response in normal cell groups, indicating that these drugs were efficacious for eliciting an apoptotic response while avoiding potential side effects (see FIGS. **22**A-**22**B). Certain BCL2 inhibitors, such as Venetoclax, can exhibit similar priming responses when combined with peptides that have a BCL2 inhibiting effect, which can help identify new CLL therapies (see FIGS. **28**, **29**). Finally, MCL1 inhibitors were found to exhibit strong effects as part of combination therapies when used together with MS1 peptide, meaning that other MCL1 inhibitors that have a strong priming effect with MS1 peptide would likely be effective as a component in some combination CLL therapies (see FIG.

30).
[0321] Drug sensitivity results in M-CLL and U-CLL groups were compared by comparing delta-priming or DBP assay output and DKFZ viability assay output at the same concentration, the mean of the two closest (higher and lower) z-scored medians or z-scored means were plotted against each other (see FIGS. **24**A-**24**D). The plots of these results had low correlation, which likely means that the two assays could effectively reveal new information or previously unidentified targets for CLL therapies when used in comparison against each other.
[0322] DBP response with Venetoclax for different driver alterations at different peptide concentrations was also studied, by which the relative drug resistance or sensitivity for Venetoclax therapy was determined for different subtypes (see FIG. **25**). Similarly, the efficacy of Abexinostat was studied by analyzing DBP response under conditions, and was found to be effective in cases where patients would be likely to be resistant to drugs such as Nutlin 3, MK-2206, and Zanubrutinib (see FIG. **27**).
[0323] The level of association for features such as IGHV, epitype, EC subtype, and driver alterations were investigated using delta priming as a representative of drug response across all CLL samples in FIGS. **31**-**38**.
[0324] Median delta-priming for molecular features of interest, including median delta-priming for healthy donors per normal cell type, were shown in FIGS. **39**, **40**, and **41**.
[0325] Altogether, the above examples establish a framework that links ex-vivo drug response with molecular features including expression subtypes to highlight new therapeutic opportunities in CLL. Therefore, drug sensitivity experiment data can be used to inform differential effects among expression clusters (FIG. **13**) and inform treatment selection for a patient with a chronic lymphocytic leukemia.
[0326] The following methods and materials may be employed.
Data Availability Sequencing, expression, and genotyping is available at European Genome-Phenome Archive (EGA), which is hosted at the European Bioinformatics Institute (EBI), under accession numbers EGAS00000000092 and in dbGaP under accession numbers: phs001473, phs000922.v2.pl, phs001431, phs001091.v1.01, phs000435.v3.pl, phs002297.v1, phs000879.v1.pl. 450k array data is available at EGA under accession number EGAD00010001975.
Code Availability
[0327] Terra methods can be found at app.terra.bio/. The new epiCMIT suitable for Illumina arrays and NGS approaches can be found at github.com. The RFcaller pipeline is available at github.com. Additional code used for the project can be found at github.com.
Human Samples
[0328] The 1156 CLL/MBL samples (1010 CLL samples were used in the clinical analysis) included tumor and germline samples collected either during active surveillance (n=687), post-treatment (n=52), or at enrollment of a clinical trial prior to first cycle of therapy (n=417; treatment-naive n=371, relapsed/refractory n=46). Briefly, these trials included: (i) comparison of fludarabine and cyclophosphamide (FC) to FC-rituximab (FCR) in previously untreated patients (CLL8 trial, n=309); (ii) treatment-naive TP53 mutated patients within phase 2 CLL20 trial who all received alemtuzumab (n=31); (iii) ibrutinib or R-ibrutinib in relapsed/refractory (R/R) or untreated patients with 17p deletion, TP53 mutation, and/or 11q deletion (n=77; treatment-naive n=31; R/R n=46). If multiple samples were obtained from a patient, then the earliest collected sample was selected for analysis. Peripheral blood mononuclear cells were isolated and DNA and/or RNA were extracted and prepared as previously described (Stilgenbauer, S. et al. Gene mutations and treatment outcome in chronic lymphocytic leukemia: results from the CLL8 trial. Blood 123, 3247-3254 (2014). Landau, D. A. et al. Mutations driving CLL and their evolution in progression and relapse. Nature 526, 525 (2015); Puente, X. S. et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. Nature 526, 519 (2015); Gruber, M. et al. Growth dynamics in naturally progressing chronic lymphocytic leukaemia. Nature 570, 474-479 (2019); Landau, D. A. et al. The

evolutionary landscape of chronic lymphocytic leukemia treated with ibrutinib targeted therapy. Nat. Commun. 8, 2185 (2017); Kasar, S. et al. Whole-genome sequencing reveals activation-induced cytidine deaminase signatures during indolent chronic lymphocytic leukaemia evolution. Nat. Commun. 6, 8866 (2015); Burger, J. A. et al. Safety and activity of ibrutinib plus rituximab for patients with high-risk chronic lymphocytic leukaemia: a single-arm, phase 2 study. Lancet Oncol. 15, 1090-1099 (2014); Burger, J. A. et al. Clonal evolution in patients with chronic lymphocytic leukaemia developing resistance to BTK inhibition. Nat. Commun. 7, 11589 (2016)).

Molecular Data Retrieval and Assembly

[0329] Previously reported sequencing data was retrieved from CLL and MBL samples, including 984 whole-exome sequences, 177 whole-genome sequences, 453 RNA-seqs, 490 methylation 450k arrays, and 547 reduced-representation bisulfite sequencing. Additionally, 264 RNA-seq samples were sequenced, and targeted DNA sequencing of the NOTCH]3′ UTR was performed for 293 samples, as described below.

RNA-Seq Generation

[0330] For cDNA Library Construction, total RNA was quantified using the Quant-iT RiboGreen RNA Assay Kit and normalized to 5ng/ul. Following plating, 2 µL of ERCC controls (using a 1:1000 dilution) were spiked into each sample. An aliquot of 200ng for each sample underwent library preparation using an automated variant of the Illumina TruSeq Stranded mRNA Sample Preparation Kit, followed by heat fragmentation and cDNA synthesis from the RNA template. The resultant 400 bp cDNA then underwent dual-indexed library preparation, consisting of ‘A’ base addition, adapter ligation using P7 adapters, and PCR enrichment using P5 adapters. After enrichment, the libraries were quantified using Quant-iT PicoGreen (1:200 dilution). After normalizing samples to 5 ng/uL, the set was pooled and quantified using the KAPA Library Quantification Kit for Illumina Sequencing Platforms.

[0331] For Illumina sequencing, pooled libraries were normalized to 2 nM and denatured using 0.1 N NaOH prior to sequencing. Flowcell cluster amplification and sequencing were performed according to the manufacturer's protocols using the NovaSeq 6000, HiSeq 2000 or HiSeq 2500.

[0332] Each run was a 101 bp paired-end read with eight-base index barcodes. Raw data was analyzed using the Broad Picard Pipeline which includes de-multiplexing and data aggregation.

Sequence Data Processing and Analysis

[0333] All sequencing data (WES, WGS, RNA-seq, RRBS and targeted NOTCH] sequencing) were processed and analyzed using methods implemented in the Broad Institute's cloud-based Terra platform (app.terra.bio).

WES/WGS Alignment and Quality Control

[0334] All DNA sequence data was processed through the Broad Institute's data processing pipeline. For each sample, this pipeline combines data from multiple libraries and flow cell runs into a single BAM file. This file contains reads aligned to the human genome hg19 genome assembly (version b37) done by the Picard and Genome Analysis Toolkit (GATK) developed at the Broad Institute, a process that involves marking duplicate reads, recalibrating base qualities and realigning around indels. Reads were aligned to the hg19 genome assembly (version b37) using BWA-MEM (version 0.7.15-r1140).

Mutation Calling

[0335] Prior to variant calling, the impact of oxidative damage (oxoG) to DNA during sequencing was quantified using DeToxoG (Costello, M. et al. Discovery and characterization of artifactual mutations in deep coverage targeted capture sequencing data due to oxidative DNA damage during sample preparation. Nucleic Acids Res. 41, e67 (2013)). The cross-sample contamination was measured with ContEst based on the allele fraction of homozygous SNPs (Cibulskis, K. et al. ContEst: estimating cross-contamination of human samples in next-generation sequencing data. Bioinformatics 27, 2601-2602 (2011)), and this measurement was used in the downstream mutation calling pipeline. From the aligned BAM files, somatic alterations were identified using a set of

tools developed at the Broad Institute (broadinstitute.org/cancer/cga). The details of the sequencing data processing have been described elsewhere (Berger, M. F. et al. The genomic complexity of primary human prostate cancer. Nature 470, 214-220 (2011); Chapman, M. A. et al. Initial genome sequencing and analysis of multiple myeloma. Nature 471, 467-472 (2011)). Briefly, for sSNVs/indel detection, high-confidence somatic mutation calls were made by applying MuTect (Cibulskis, K. et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat. Biotechnol. 31, 213-219 (2013)), MuTect2 (Benjamin, D. et al. Calling Somatic SNVs and Indels with Mutect2. bioRxiv 861054 (2019) doi:10.1101/861054) and Strelka2 (Kim, S. et al. Strelka2: fast and accurate calling of germline and somatic variants. Nat. Methods 15, 591-594 (2018)) to WES/WGS sequencing data. Given that normal blood samples might also contain CLL cells, DeTiN (Taylor-Weiner, A. et al. DeTiN: overcoming tumor-in-normal contamination. Nat. Methods 15, 531-534 (2018)) was used to estimate tumor in normal (TiN) contamination in order to recover falsely rejected sSNVs/indels. Next, four types of filters were applied: (i) a realignment-based filter, which removes variants that can be attributed entirely to ambiguously mapped reads; (ii) an orientation bias filter, which removes possible oxoG and FFPE artifacts; (iii) a ContEst filter, which removes variants that might come from other samples due to contamination; and (iv) an allele fraction specific panel-of-normals filter, which compares the detected variants to a large panel of normal exomes or genomes and removes variants that were observed in the two panel-of-normals (PoNs): one consists of 8,334 normal samples in TCGA while the other consists of 481 CLL-matched normal samples with TiN estimates of 0. All four filters together contributed to the exclusion of potential false-positive events (e.g. commonly occurring germline variants or sequencing artifacts), which ultimately yielded the final list of mutations. All filtered events in candidate CLL driver genes were also manually reviewed using the Integrated Genomics Viewer (IGV) (Robinson, J. T. et al. Integrative genomics viewer. Nat. Biotechnol. 29, 24-26 (2011)).

[0336] In order to increase the sensitivity and precision of mutation calls in candidate driver genes, an additional variant calling step was performed for the candidate driver gene loci using Rfcaller (github.com/xa-lab/RFcaller), a pipeline that uses read-level features and extra trees/random forest algorithms for the detection of somatic mutations. This pipeline was run with default parameters for whole exome sequencing (WES) or whole genome sequencing (WGS) data, as well as for RNA-seq data for NOTCH], which has low coverage in hotspot regions in some samples due to high GC content. All candidate mutations that passed filters and were detected by both pipelines were considered positives. Mutations detected by only one of the callers were visually inspected by a set of at least four expert curators, considering the following exclusion criteria: (i) low evidence due to limited number of reads supporting the mutation in the tumor sample or excessive mutant reads in the normal sample; (ii) low depth of coverage to rule out germline variant; (iii) low base quality region; (iv) low mapping quality region leading to multi-mapped reads; (v) calls supported by reads with a strong strand bias.

Identification of Significantly Mutated Genes

[0337] To identify candidate cancer genes using the mutation calls from WES, SignatureAnalyzer (Kim, J. et al. Somatic ERCC2 mutations are associated with a distinct genomic signature in urothelial tumors. Nat. Genet. 48, 600-606 (2016)) was first used to identify mutational processes and potential artifact signatures. A signature likely due to the bleedthrough sequencing artifact was discovered and then mutations with greater than 95% chance attributed to that bleedthrough signature were filtered. Next, MutSig2CV (Lawrence, M. S. et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature 499, 214-218 (2013)) was run to identify driver genes from the filtered whole exome sequencing (WES) Mutation Annotation Format (MAF) file. A stringent manual review was conducted using the IGV (Robinson, J. T. et al. Integrative genomics viewer. Nat. Biotechnol. 29, 24-26 (2011)) to review the mutations in the driver genes and further exclude low evidence calls. Then MutSig2CV was rerun on the filtered set

of mutation calls from whole exome sequencing (WES) to identify the final candidate driver genes. In addition, CLUMPS (Kamburov, A. et al. Comprehensive assessment of cancer missense mutation clustering in protein structures. Proc. Natl. Acad. Sci. U.S.A 112, E5486-95 (2015)) was used to identify driver genes based on clustering of mutations in the 3D structure of the protein product. For CLUMPS, two FDR corrections were applied: one for all candidates and a second restricted hypothesis testing focused on genes in the COSMIC Cancer Gene Census (Sondka, Z. et al. The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. Nat. Rev. Cancer 18, 696-705 (2018)). Finally, for further stringency and to exclude candidates irrelevant to CLL biology, candidate genes that were not expressed in RNA-seq of 610 treatment-naive CLL samples were discarded using a one-sided t-test testing for difference from 0 in transcripts per million (TPM) space. This discarded 15 candidate genes.

UI g.3A>C Mutational Status

[0338] The U1 g.3A>C mutational status for 294 cases from the ICGC cohort was previously reported (Shuai, S. et al. The U1 spliceosomal RNA is recurrently mutated in multiple cancers. Nature 574, 712-716 (2019)). For the remaining 212 ICGC cases, U1 status was determined using a previously validated rhAMP SNP assay (Integrated DNA Technology) (Shuai, S. et al.). The U1 status of 425 patients from the DFCI/Broad cohort was inferred from RNA-seq data using a random forest classifier with 100 trees built from 3,174 differentially spliced introns between U1 mutated and wild-type cases, as previously described (Shuai, et al.). A cohort of 104 cases from the ICGC cohort (7 mutated, 97 wild-type) was used to train the model, while 54 cases (3 mutated, 51 wild-type) were used as a test (Shuai, et al.). Altogether, the U1 g.3A>C status was determined for 931 of 1156 cases.

NOTCH1 Mutation Calling

[0339] A subset of the whole exome sequencing (WES) data had reduced coverage in the GC-rich region of NOTCH], a common and clinically-relevant driver in CLL. The NOTCH] calls from WES/WGS were augmented by Sanger sequencing, targeted deep sequencing of NOTCH]3′ UTR (details below), and manual review of all WES, whole genome sequencing (WGS) and RNA-seq in IGV (Robinson, J. T. et al. Integrative genomics viewer. Nat. Biotechnol. 29, 24-26 (2011)). This was primarily focused on identifying NOTCH] hotspot CT deletion p.P2515Rfs*4 and NOTCH]3′ UTR mutational hotspot chr9:139390152T>C. RNA-seq review was based on the direct mutation and the splicing perturbation associated with the 3′ UTR mutation.

Targeted Sequencing of NOTCH1 3′ UTR

[0340] To amplify the region of the NOTCH]3′ UTR hotspot mutation at position chr9:139390152T>C and adjacent sequence from genomic DNA, the following PCR1 reaction mix was prepared including 1× PfX amplification buffer, 1× PfX enhancer solution (ThermoFisher, 11708039), 0.3 mM each dNTPs, 1 mM MgSO4, 0.6 μM of NOTCH]1.sup.st F-primer, 0.6 μM of Notchl 1.sup.stR-primer. To each well of a 96 well plate, 46pL of this mix was added and 2 μL of DNA sample (25ng/pL concentration), and then following PCR reaction was performed: 95° C. 5 min, 33 cycles of (95° C. 30s, 55° C. 30S, 68° C. 1 min), and then held at 4° C. Once the plate was heated to 95° C. for 1 min, the reaction was paused, and the plate was taken out and 2 μL Pfx polymerase mix (1:4 diluted Pfx Polymerase with water) was added into each well, and then the reaction program was continued. In order to add an identifier index onto each amplicon, the PCR2 was performed. First, the following reaction mix was prepared containing 1× Kapa HiFi Fidelity buffer (2 mM MgCl.sub.2), 0.41 mM of each dNTPs, 1 μL of Kapa HiFi hotstart polymerase (KapaBiosystems, KK2101), 0.82 μM of the forward primer, and 0.82 μM of each reverse primer (in a 96 well plate). Then 50 μL of the mix was added to anew 96 well plate and 10p L of the PCR1 mix was added to each well of the plate, and the following PCR reaction was performed: 98° C. 45s, 8 cycles of (98° C. 15s, 60° C. 30s, 72° C. 30s), 72° C. 1 min and then held at 4° C. After PCR2, 3 μL of each of the indexed PCR products was pooled and cleaned up using Ampure XP beads. After cleaning, the pooled libraries were quantified using a Bioanalyzer, and then sequenced

on a MiSeq using the following parameters: Read 1: 200nt, Read 2: 100nt, Index 1: 8nt, and index 2: 8nt.

Copy Number Analysis

[0341] For detecting somatic copy number alterations (sCNAs) the GATK4 CNV pipeline (github.com/gatk-workflows/gatk4-somatic-cnvs) was used, which involves the CalculateTargetCoverage, NormalizeSomaticReadCounts, and Circular Binary Segmentation (CBS) algorithms (Olshen, A. B., Venkatraman, E. S., Lucito, R. & Wigler, M. Circular binary segmentation for the analysis of array-based DNA copy number data. Biostatistics 5, 557-572 (2004)) for genome segmentation. In order to identify candidate somatic copy number alteration (sCNA) drivers (genomic regions that are significantly amplified or deleted), GISTIC 2.0 was then applied (Mermel, C. H. et al. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol. 12, R41 (2011)). To exclude potential germline CNAs, GISTIC 2.0 was first run on the matched normal samples and then the recurrent CNAs this outputted ($q<0.1$) was concatenated to the blacklisted regions. Then GISTIC 2.0 was run on the tumor samples to produce a list of candidate somatic copy number alteration (sCNA) driver regions. A force-calling process was applied to identify the presence/absence of each somatic copy number alteration (sCNA) driver event across tumor samples (github.com/getzlab/GISTIC2_postprocessing). To further filter the potential false positive drivers, only somatic copy number alteration (sCNA) drivers with population frequency greater than 1% were accepted. Finally, all filtered somatic copy number alteration (sCNA) drivers were manually reviewed using IGV (Robinson, J. T. et al. Integrative genomics viewer. Nat. Biotechnol. 29, 24-26 (2011)) to exclude drivers that are based on somatic copy number alteration (sCNA) events with low supporting evidence or that were localized close to centromeres. Somatic copy number alteration (sCNA) drivers were annotated by intersection with our list of CLL mutation driver genes and with genes in the COSMIC Cancer Gene Census (Sondka, Z. et al. The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. Nat. Rev. Cancer 18, 696-705 (2018)) (v90).

Structural Variants Calling

[0342] For structural variation (SV) detection, the pipeline integrated evidence from three structural variation detection algorithms (Manta (Chen, X. et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. Bioinformatics 32, 1220-1222 (2016)), SvABA (Wala, J. A. et al. SvABA: genome-wide detection of structural variants and indels by local assembly. Genome Res. 28, 581-591 (2018)) and dRanger (Berger, M. F. et al. The genomic complexity of primary human prostate cancer. Nature 470, 214-220 (2011); Bass, A. J. et al. Genomic sequencing of colorectal adenocarcinomas identifies a recurrent VTITA-TCF7L2 fusion. Nat. Genet. 43, 964-968 (2011); Chapman, M. A. et al. Initial genome sequencing and analysis of multiple myeloma. Nature 471, 467-472 (2011)) to generate a list of structural variation events with high confidence. The three SV detection tools were followed with BreakPointer (Drier, Y. et al. Somatic rearrangements across cancer reveal classes of samples with distinct patterns of DNA breakage and rearrangement-induced hypermutability. Genome Res. 23, 228-235 (2013)) to pinpoint the exact breakpoint at base-level resolution. Breakpoint information was aggregated per sample to identify: (i) balanced translocations, which were defined as those with breakpoints on reverse strands within 1-kb of each other; (ii) inversions supported on both ends; (iii) complex events, based on the number of clustered events within 50-kb of each other. Breakpoints were annotated by intersection with the lists of CLL driver genes and significant somatic copy number alteration (sCNA) regions, as well as with 20 genes in the COSMIC Cancer Gene Census (v90) (Sondka, Z. et al. The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. Nat. Rev. Cancer 18, 696-705 (2018)).

Identification of Structural Variants Involving the Immunoglobulin (IG) Loci

[0343] Potentially oncogenic structural variants involving any of the IG loci were analyzed using

IgCaller (v1.1) (Nadeu, F. et al. IgCaller for reconstructing immunoglobulin gene rearrangements and oncogenic translocations from whole-genome sequencing in lymphoid neoplasms. Nat. Commun. 11, 3390 (2020)) and visually inspected in IGV (Robinson, J. T. et al. Integrative genomics viewer. Nat. Biotechnol. 29, 24-26 (2011)). The breakpoints of the IG loci were used to determine the underlying mechanisms leading to these events. To that end, a search was done for evidence of aberrant V(D)J recombination (i.e., breakpoints in any of the V(D)J genes and close to recombination-activation gene (RAG) signal sequences) or aberrant class switch recombination (CSR) (i.e., breakpoints located in any of the CSR regions). IG genes and CSR regions were annotated based on the annotations used by IgCaller. Of note, no evidence of IG structural variants mediated by somatic hypermutation (SHM) were identified (i.e., events with breakpoints within already rearranged V(D)J genes linked with the presence of SHM).

Estimation of Purity, Ploidy, and Cancer Cellfraction (CCF)

[0344] To estimate sample purity, ploidy, absolute allele-specific copy number and cancer cell fraction (CCF) of the filtered whole exome sequencing (WES) somatic coding mutations, ABSOLUTE (Carter, S. L. et al. Absolute quantification of somatic DNA alterations in human cancer. Nat. Biotechnol. 30, 413-421 (2012)) was used, which integrates allele fraction specific information from the sequencing data for sSNVs/indels and sCNAs. For each sample, manual review was conducted to determine the optimal ABSOLUTE solution. Using these ABSOLUTE solutions allowed for recovery of CCF estimates for 49,882 coding mutations of all 53,489 mutations (93.3%) identified in whole exome sequencing (WES) data.

[0345] Timing Analysis

[0346] To infer phylogenetic and evolutionary trajectories based on somatic mutations and copy number variation, PhylogicNDT Cluster, Timing, LeagueModel modules were applied (Leshchiner, I. et al. Comprehensive analysis of tumour initiation, spatial and temporal progression under multiple lines of treatment. bioRxiv 508127 (2019)) (github: github.com/broadinstitute/PhylogicNDT) on the filtered whole exome sequencing (WES) MAF with CCF annotated from the optimal ABSOLUTE solution. To determine if shared events had significantly different order of acquisition in M-CLL (CLL with mutated IGHV) and U-CLL, the timing score was randomly sampled 250,000 times for each shared event from the MCMC traces of M-CLL (CLL with mutated IGHV) and U-CLL (CLL with unmutated IGHV) respectively, and the difference between the two scores was calculated. Then the frequency of the differences being less than 0 was calculated. If the frequency was less than 0.5, then the p-value was assigned as two times the frequency to that event, i.e. p-value=2*freq; else, the p-value was assigned as two times one minus the frequency to that event, i.e. p-value=2*(1-freq). Then the Benjamini-Hochberg multiple hypothesis correction procedure was applied to all the p-values of shared driver events. The timing of a shared driver event was considered significantly different between the two subtypes if the corresponding q value was less than 0.1.

Gene Set Enrichment for Driver Genes

[0347] Gene set enrichment analysis was performed using g:profilerhttps://paperpile.com/c/CLn83y/PrfG (Reimand, J. et al. g:Profiler-a web server for functional interpretation of gene lists (2016 update). Nucleic Acids Res. 44, W83-9 (2016)) on the 97 driver genes, the total identified in the MutSig and CLUMPS analyses for “All,” M-CLL, and U-CLL (CLL with unmutated IGHV) (excluding genes detected only by CLUMPS restricted hypothesis testing for cancer genes, n=2; and excluding 5 genes not found in the gene set annotation). Gene sets from MSigDB v7.0 were used, aggregating Hallmark, C5:GO:BP and C2:CP:REACTOME collections. g:profiler results were filtered by $q<0.1$, restricted in size between 5 and 350 genes in the gene set, and required to include at least two drivers. To identify similar biological processes and remove redundancy in overlapping gene sets, significant gene sets were clustered using Louvain clustering (Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. arXiv [physics.soc-ph](2008)) (igraph R package

v1.2.5). To that end, a gene set network was constructed, where nodes were gene sets and edges are weighted based on shared gene membership by Jaccard index. Three cutoffs for the Jaccard index (0.9, 0.95, 0.99) were applied before clustering to produce different clustering resolutions. The clustering was repeated twice, considering membership by shared drivers or any shared genes between the gene sets. Results were reviewed and biological processes were generalized manually. Candidate genes that were not enriched in gene sets by this process were assigned to pathways by data curation.

Immunoglobulin (IG) Gene Characterization

[0348] The IG heavy (IGH) and light (IGL) chain gene rearrangements and mutational status were obtained from WGS/WES and RNA-seq using IgCaller (v1.1) (Nadeu, F. et al. IgCaller for reconstructing immunoglobulin gene rearrangements and oncogenic translocations from whole-genome sequencing in lymphoid neoplasms. Nat. Commun. 11, 3390 (2020)) and MiXCR (v.3.0.10) (Bolotin, D. A. et al. MiXCR: software for comprehensive adaptive immunity profiling. Nat. Methods 12, 380-381 (2015)), respectively. The rearrangements obtained were visually inspected in IGV (Robinson, J. T. et al. Integrative genomics viewer. Nat. Biotechnol. 29, 24-26 (2011)). IGH gene rearrangements were complemented with Sanger sequencing available for 1085 cases. The IGHV (heavy chain variable region of immunoglobulin genes) mutational status obtained by IgCaller (WGS/WES) and MiXCR were concordant in 506/516 (98%) cases with an IGH rearrangement identified by both methods. The 10 discordant cases were classified based on the IGHV (heavy chain variable region of immunoglobulin genes) mutational status determined by Sanger sequencing (concordant with MiXCR in 8 cases and with IgCaller in 2). IgCaller/MiXCR and Sanger sequencing were concordant in 903/925 (98%) of the cases with an IGH gene rearrangement obtained by both methodologies. The result obtained by IgCaller/MiXCR was used in the 22 discordant cases after careful examination of the sequences. Note that in 12/22 cases the results obtained by IgCaller and MiXCR were concordant. For the remaining 10 cases, only IgCaller or MiXCR results were available. The IGHV (heavy chain variable region of immunoglobulin genes) mutational status of 14 cases carrying a mix of mutated and unmutated IGH gene rearrangements was considered as “not available”. Similarly, the IGH genes in 43 cases carrying two IGH rearrangements (the previous 14 cases with mixed IGHV (heavy chain variable region of immunoglobulin genes) mutational status and 29 cases with two mutated or two unmutated IGH gene rearrangements) were considered as “not available”. Altogether, 1136/1154 (98%) cases were classified based on their IGHV (heavy chain variable region of immunoglobulin genes) mutational status. To study B-cell receptor (BCR) stereotypy, the 19 major stereotype subsets were annotated using the ARResT/AssignSubsets online tool (Bystry, V. et al. ARResT/AssignSubsets: a novel application for robust subclassification of chronic lymphocytic leukemia based on B cell receptor IG stereotypy. Bioinformatics 31, 3844-3846 (2015)).

[0349] IGL gene rearrangements obtained by IgCaller and MiXCR were concordant in all but five cases with both methods available (581/586, 99%). The output of MiXCR was accepted in the five discordant cases after manual revision. As performed for IGH gene rearrangements, cases carrying two IG populations with distinct IG gene rearrangements were blacklisted from the IGL gene annotation. To properly characterize the IGLV3-21R.sup.110, IGLV3-21 rearranged sequences reported by IgCaller were manually curated to phase single nucleotide polymorphisms with the rearranged allele, as previously described (Nadeu, F. et al. IGLV3-21R110 identifies an aggressive biological subtype of chronic lymphocytic leukemia with intermediate epigenetics. Blood (2020) doi: 10.1182/blood.2020008311). Curated IGLV3-21-rearranged sequences from IgCaller and original IGLV3-21-rearranged sequences from MiXCR (in which the manual phasing of polymorphisms is not needed) were used as input of IMGT/V-QUEST (v3.5.18; release 202018-4) (Brochet, X., Lefranc, M.-P. & Giudicelli, V. IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. Nucleic Acids Res. 36, W503-8 (2008)) to annotate the IGLV3-21 allele, the motifs involved in BCR-BCR interactions

[lysine (K) 16 and aspartates (D) 50 and 52], and the presence of the glycine to arginine mutation at position 110 (R110) (Nadeu, F. et al. IGLV3-21R110 identifies an aggressive biological subtype of chronic lymphocytic leukemia with intermediate epigenetics. Blood (2020) doi:10.1182/blood.2020008311). Overall, IGLV3-21.sup.R110 status was determined in 1128/1154 (97.7%) cases.

RNA-Seq Analysis

[0350] RNA-seq data was processed in Terra using the GTEx V7 pipeline (github.com/broadinstitute/gtex-pipeline). Briefly, reads were aligned with STAR (v2.6.1d) (Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15-21 (2013)) to hg19 (b37) using the GENCODE v19 annotation, and quality control metrics and gene expression were computed with RNA-SeQC (v2.3.6) (Graubert, A., Aguet, F., Ravi, A., Ardlie, K. G. & Getz, G. RNA-SeQC 2: Efficient RNA-seq quality control and quantification for large cohorts. Bioinformatics (2021) doi:10.1093/bioinformatics/btabl35). A collapsed version of the GENCODE annotation was used to quantify gene-level expression (available from gs://gtex-resources/GENCODE/gencode.v19.genes.v7.collapsed_only.patched_contigs.gtf). Transcripts per million (TPMs) were used for sample clustering, while gene counts were used for differential gene expression, as required.

RNA Expression Cluster Detection

[0351] Gene-level transcripts per million (TPMs) were estimated with RNA-SeQC (v2.3.6) for RNA-seq from 610 treatment-naive CLL. Genes expressed at less than 0.1 transcripts per million (TPM) in 10% of samples were discarded, retaining 11,119 genes, which were batch corrected (as described below), followed by selection of the top 2,500 most varying genes. The clustering methodology combined consensus hierarchical clustering and Bayesian non-negative matrix factorization (BayesNMF), as previously described (Robertson, A. G. et al. Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. Cell 171, 540-556.e25 (2017)). Briefly, the method computed a distance matrix 1-C, where element C.sub.ij represented the Spearman correlation between samples i and j across the 2,500 genes. It used the distance matrix to perform iterations of standard hierarchical clustering with 80% sample resampling for 250 iterations per value of parameter K, where K represents the cutoff for the number of clusters running from 2 to 20. The result was the cumulative consensus matrix M, where M.sub.ij represents the number of times samples i and j shared cluster membership, which was then normalized by the total number of iterations to create the matrix M*. Next, BayesNMF was performed on M* to identify the optimal number of clusters K* and computed the strength of association of each sample to each cluster. The maximum association determined final cluster assignment. By parallelization, the number of independent BayesNMF runs was increased from 20 to 1000, 77.4% of which converged to the dominant result of K*=8 clusters (20% K*=7, 1.8% K*=6).

RNA-Seq Batch Effect Correction

[0352] Preprocessing of RNA-seq data for expression cluster detection was undertaken to address batch effects between samples collected at different centers and processed by different protocols. To that end, a comprehensive set of covariates was assembled that allowed for adequate control for technical artifacts: (i) Quality metrics from RNA-SeQC v2.3.6 (Graubert, A., Aguet, F., Ravi, A., Ardlie, K. G. & Getz, G. RNA-SeQC 2: Efficient RNA-seq quality control and quantification for large cohorts. Bioinformatics (2021) doi:10.1093/bioinformatics/btabl35); (ii) CIBERSORT (Chen, B., Khodadoust, M. S., Liu, C. L., Newman, A. M. & Alizadeh, A. A. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. Methods Mol. Biol. 1711, 243-259 (2018)) relative immune cell composition estimates (cibersort.stanford.edu/) where B-cell estimates were excluded to prevent masking CLL-intrinsic signals; (iii) PEER factors (Stegle, O., Parts, L., Durbin, R. & Winn, J. A Bayesian framework to account for complex non-genetic factors in gene expression levels greatly increases power in eQTL studies. PLoS Comput. Biol. 6, e1000770 (2010)); (iv) Sex, which was

systematically inferred by KMeans clustering (sklearn v0.21.3) using XIST and RPS4Y1 gene expression; (v) explicit sequencing batch if available; (vi) sequencing center (Broad Institute or Barcelona); (vii) a metric devised to estimate the sample processing artifact described in Dvinge et al (Dvinge, H. et al. Sample processing obscures cancer-specific alterations in leukemic transcriptomes. Proceedings of the National Academy of Sciences 111, 16802-16807 (2014)). This metric was computed by Spearman correlation between a sample's expression profile to the genes reported by Dvinge et al to be differentially expressed after 48 hours of incubation at suboptimal temperatures. However, to reduce the potential contribution of CLL-related expression to this metric, the correlation was computed by focusing on 3,682 differentially expressed genes that have been previously defined as house-keeping genes (Eisenberg, E. & Levanon, E. Y. Human housekeeping genes, revisited. Trends Genet. 29, 569-574 (2013)). Of note, covariates from RNA-SeQC (Graubert, A., Aguet, F., Ravi, A., Ardlie, K. G. & Getz, G. RNA-SeQC 2: Efficient RNA-seq quality control and quantification for large cohorts. Bioinformatics (2021) doi:10.1093/bioinformatics/btabl35) and CIBERSORT were converted to PCA space. Top PCs and PEER factors were selected as appropriate. Batch correction for expression cluster (EC) detection was performed by including the covariates as fixed effects in a linear model to regress out effects they were associated with, and sample clustering was performed on the resulting residuals.

Marker Gene Detection and Differential Expression Analysis

[0353] To identify marker genes per expression cluster, a second non-negative matrix factorization step was applied, as previously described (Robertson, A. G. et al. Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. Cell 171, 540-556.e25 (2017)). However, in this study, batch-corrected transcripts per million (TPMs) were used and a fold-change of 1.5 was required between each cluster and all others. Markers selected were limited to the top 10 most up and down regulated genes per expression cluster (EC) (Tables 3A-3B and 4). Additionally, limma-voom (Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43, e47 (2015); Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol. 15, R29 (2014)) was used to identify differentially expressed genes between each expression cluster (EC) and all others. The same covariates used for RNA-seq batch effect correction for expression cluster discovery were included in the models, while using unmodified gene counts from RNA-SeQC (Graubert, A., Aguet, F., Ravi, A., Ardlie, K. G. & Getz, G. RNA-SeQC 2: Efficient RNA-seq quality control and quantification for large cohorts. Bioinformatics (2021) doi:10.1093/bioinformatics/btabl35). Genes with $q<0.05$ and absolute fold-change greater than 1.5 were considered differentially expressed (Tables 3A-3B and 4).

Gene Set Enrichment Analysis for Expression Clusters (ECs)

[0354] Gene set enrichment per each expression cluster was performed using fgsea (github.com/ctlab/fgsea) (Korotkevich, G. et al. Fast gene set enrichment analysis. bioRxiv 060012 (2021) doi:10.1101/060012), which was applied to the W matrix produced by the second BayesNMF step that detected marker genes associated with each expression cluster (EC) (see Robertson et al. (Robertson, A. G. et al. Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. Cell 171, 540-556.e25 (2017)) for details). In essence, this represents gene lists ranked by their association with each EC, ranging from most positively associated to most negatively associated. Gene sets from MSigDB v7.0 were used, aggregating Hallmark, C5:GO:BP and C2:CP:REACTOME collections. Analysis was restricted to gene sets of size 12 to 500, and $q<0.1$ was required. For further confidence, we applied Gene Set Variation Analysis (GSVA) from the gsva R package (Hanzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 14, 7 (2013)) using the top 2500 most varying genes. GSVA estimates were summarized per expression cluster (EC) and mean differences computed between each expression cluster (EC) and all others. The intersection of results from fgsea and GSVA was retained.

[0355] Next, to identify related biological processes and remove redundancy in overlapping gene sets, significant gene sets were clustered using Louvain clustering (Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. arXiv [physics.soc-ph](2008)) (igraph R package v1.2.5). To that end, a gene set network was constructed, where nodes were gene sets and edges were weighted based on shared gene membership by Jaccard index (using genes in the “leading edge” reported by fgsea). Three cutoffs for Jaccard index (0.8, 0.9, 0.95) were applied before clustering to produce different clustering resolutions. Finally, results were reviewed and biological processes were generalized manually. Only gene sets with absolute NES scores>2 fromfgsea and a>0.1 difference in mean GSVA score between the respective expression cluster (EC) and all other samples were considered.

Detection of Statistically Significant Pairwise Associations of Molecular Features

[0356] To identify statistically significant pairwise associations of molecular features (e.g., association of expression clusters (ECs) with candidate drivers), the curveball permutation algorithm (Strona, G., Nappo, D., Boccacci, F., Fattorini, S. & San-Miguel-Ayanz, J. A fast and unbiased procedure to randomize ecological binary matrices with fixed row and column totals. Nat. Commun. 5, 4114 (2014)) was applied to a comprehensive sample annotation table to generate the null distribution of the p-value from one-sided Fisher's Exact tests for each pair of features. The sample annotation table contained binary indicators for all sSNV/indel drivers and somatic copy number alteration (sCNA) drivers identified, in addition to U1 mutation, IGLV3-21.sup.R110 mutation, IGHV (heavy chain variable region of immunoglobulin genes) mutational status, expression clusters (ECs) and epitypes. Samples that had DNA, RNA and methylation data were focused upon, and they were also required to be treatment-naive (n=506). The goal of the curveball algorithm was to estimate an accurate null distribution through controlling the sample-level driver mutation rates, which reduced false positive associations caused by background mutation burdens. 5000 curveball permutation iterations were applied to generate this null distribution and then the observed p-value was compared against it to get the empirical p-value for co-occurring and mutual-exclusive patterns for each feature pair. The Benjamini-Hochberg procedure was then applied to the empirical p-values and the significant events were selected ($q<0.1$) (Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J. R. Stat. Soc. 57, 289-300 (1995)).

Expression Cluster Machine-Learning Classifier

[0357] The 610 treatment-naive RNA-seqs of the expression cluster (EC) discovery set were split into a training set (n=487, 80%) and test set (n=123, 20%). The latter was used to assess performance after final model selection. Features used in the model were derived from differential expression results between expression clusters (ECs) using limma-voom (Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43, e47 (2015)) on training set samples. Models were trained using the RandomForestClassifier class in the sklearn (v0.22.2) Python package (with parameter class weight=“balance subsample” to mitigate class imbalance in the models). Hyperparameters were optimized using 5-fold cross validation and model performance was evaluated by the harmonic mean of overall accuracy and macroFI (mean F1 across ECs). The final performance metric per hyper-parameter set was the mean of this value across cross-validation folds. Hyperparameters screened included forest size (500, 1000), number of most differentially expressed genes used from each comparison in limma-voom (5, 10, 20, 50) and oversampling method from the imblearn package (v0.6.2) used to improve performance (ADASYN, BorderlineSMOTE, SMOTE, SVMSMOTE or None). DESeq-normalized transcripts Per Million (TPMs) were used primarily and the process was repeated for batch-corrected transcripts Per Million (TPMs) to assess the impact of batch-correction on performance. Reported accuracy metrics were computed by applying the selected models to the test set.

Stability Assessment of Expression Clusters

[0358] CLL RNA-seq data generated across multiple timepoints was analyzed prior to treatment from 19 patients (Gruber, M. et al. Growth dynamics in naturally progressing chronic lymphocytic leukaemia. Nature 570, 474-479 (2019)), focusing on two time points per patient in 18 of 19 cases. For one patient, CRC-0019, all 6 samples available were analyzed prior to treatment. The machine learning expression cluster (EC) classifier was applied to these 42 samples to obtain predicted expression cluster (EC) assignments. Importantly, to avoid biases for these patient samples, the classifier was retrained while excluding these patients from the training process. Then, to test if the assignment of expression clusters (ECs) was consistent over time more than expected by chance, a permutation test was performed, randomizing all labels among the 42 samples 1,000,000 times. For each permutation a value H.sub.perm was computed by the sum of Shannon's entropy per patient. For example, a patient with consistent assignment in 2 samples contributed 0 bits to H.sub.perm, whereas a patient with two different labels contributed 1 bit. The mean H.sub.perm value was 10.47, compared to H.sub.real from the actual data that was 2.77. No randomizations were as low as this, providing a p-value<10.sup.−6 in support of expression cluster (EC) stability. This was based on stability in 15 of 19 patients, where 2/15 were classified differently than in the expression cluster (EC) discovery process. Considering 13/19 (68.4%), expression clusters (ECs) were consistent over time in most patients.

DNA Methylation Data Processing

[0359] DNA methylome data was analyzed for a total of 1,037 samples, including 490 samples profiled with Illumina 450k array previously analyzed (Duran-Ferrer, M. et al. The proliferative history shapes the DNA methylome of B-cell tumors and predicts clinical outcome. Nature Cancer 1, 1066-1081 (2020)) (EGA accession EGAD00010001975), and 547 samples profiled using reduced representation bisulfite sequencing (RRBS, with either single-end (SE), or paired-end (PE) approaches) (Landau, D. A. et al. Locally disordered methylation forms the basis of intratumor methylome variation in chronic lymphocytic leukemia. Cancer Cell 26, 813-825 (2014)). A pipeline in Terra was developed to obtain the CpG methylation estimates from RRBS data. First, FASTQC (bioinformatics.babraham.ac.uk/projects/fastqc/) and MultiQC (Ewels, P., Magnusson, M., Lundin, S. & Kdller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32, 3047-3048 (2016)) were used for quality control. Trimming was applied to the PE samples as appropriate for the RRBS protocol. Next, reads were aligned to hgl9 using BSMAP (Xi, Y. & Li, W. BSMAP: whole genome bisulfite sequence MAPping program. BMC Bioinformatics 10, 232 (2009)) (v2.90) and methylation was called with the mcall module from the MOABS package (Sun, D. et al. MOABS: model based analysis of bisulfite sequencing data. Genome Biol. 15, R38 (2014)) (v1.3.9.6). For SE samples, BSMAP was run with flags “-v 0.1-s 12-q20-w100-S1-u-R-DC-CGG-r 0”, and for PE samples with “-v0.1-s12-q20-w100-S1-u-R-r0”. mcall was run with flag “-F 256”, for primary alignments only. For downstream analysis, only CpGs covered by at least 5 reads were retained. 14 samples were then removed from the initial 1,037, since they did not pass the filtering criteria due to poor bisulfite conversion rates, poor alignment metrics, suspected contaminations from other samples, extremely low number of methylated CpGs, and/or very low number of CpGs with 5 reads compared to the general distribution. After all filtering criteria, a total of 1,023 samples were used for all downstream analyses. From these 1,023 samples, 24 were profiled twice with different platforms and were used to validate the robustness of the new epiCMIT (Duran-Ferrer, M. et al. The proliferative history shapes the DNA methylome of B-cell tumors and predicts clinical outcome. Nature Cancer 1, 1066-1081 (2020)) epigenetic mitotic clock across platforms (18 profiled with Illumina 450k vs RRBS-PE, and 6 profiled with RRBS-PE vs RRBS-SE). In these 24 cases, the platform with more CpGs covered across all samples was prioritized (from the highest to lowest priority, Illumina 450k>RRBS-PE>RRBS-SE). The remaining 999 unique samples included 490 profiled by Illumina 450k array, 390 by RRBS-SE and 119 by RRBS-PE (3 samples were not included in consensus matrices due to lower number of CpGs, including 2 RRBS-SE and 1 RRBS-PE samples). The

consensus matrices for each platform with shared CpGs across samples contained 447,800 CpGs and 490 samples for Illumina 450k data; 44,363 CpGs and 388 samples for RRBS-SE data; and 173,808 CpGs and 136 samples for RRBS-PE data [18 of these 136 samples were only used to test epiCMIT robustness across platforms, as they were already profiled with Illumina 450k; 6 of the remaining 118 RRBS-PE samples were also profiled with RRBS-SE to test epiCMIT robustness across platforms (analyzed separately and not included in the RRBS-SE consensus matrix), but were subsequently discarded and only their corresponding RRBS-PE samples were retained according to the aforementioned platform prioritization scheme]. These consensus matrices were used to perform principal component analyses (PCA) and in the case of RRBS data, also to assign CLL epitypes.

CLL Epitype Classification

[0360] The CLL epitypes were calculated for all 1,023 450k/RRBS samples. In the case of Illumina 450k data, a recently published algorithm was used which uses 4 CpGs and is suitable for both Illumina 450k and EPIC arrays (Duran-Ferrer, M. et al. The proliferative history shapes the DNA methylome of B-cell tumors and predicts clinical outcome. Nature Cancer 1, 1066-1081 (2020)). For RRBS data, the previously created consensus matrices created for RRBS-SE and RRBS-PE platforms were used separately and the following strategy was used: CLL patients with 100% and <95% IGHV (heavy chain variable region of immunoglobulin genes) identities were selected to perform differential DNA methylation analysis with mean methylation fraction differences between groups of at least 0.5. These IGHV (heavy chain variable region of immunoglobulin genes) cutoffs yielded 168 and 80 samples for RRBS-SE data, and 67 and 13 samples for RRBS-PE data with IGHV (heavy chain variable region of immunoglobulin genes) identities of 100% and <95%, respectively. These stringent cutoffs were imposed for both IGHV (heavy chain variable region of immunoglobulin genes) and DNA methylation differences to avoid borderline cases, compared with the traditional 98% IGHV (heavy chain variable region of immunoglobulin genes) and 0.25 methylation difference cutoffs. This filtering criteria translated into clearer signatures consisting of 32 and 153 differentially methylated CpGs for RRBS-SE and RRBS-PE data, respectively. These CpGs were then used to perform consensus clustering with ConsensusClusterPlus R package v.1.52.0 (Wilkerson, M. D. & Hayes, D. N. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics 26, 1572-1573 (2010)) with 10,000 permutations allowing from K=2 to K=7 groups, which robustly identified 3 consensus groups in both RRBS data types. Each sample was assigned a probability to belong to each of the groups (using the calcICL function). Samples where the maximum probability was below 0.5 or where 2 epitypes had a probability above 0.35 were considered as unclassified cases. In the 3 samples (2 RRBS-SE and 1 RRBS-PE) not included in the consensus matrices, the same strategy was used to find the CLL epitypes using the intersection of CpGs from both matrices used for consensus clustering (i.e., the 32-CpG and 153-CpG matrices for RRBS-SE and RRBS-PE data). In these cases, the epitype predictions were additionally verified using PCAs with all the shared CpGs with the rest of the samples, which further supported the assigned epitype.

Development of the epiCMIT Mitotic Clock for Next Generation Sequencing Data

[0361] The epigenetic mitotic clock, epiCMIT, was originally created with Illumina array data and thus is suitable for both 450k and EPIC arrays (Duran-Ferrer, M. et al. The proliferative history shapes the DNA methylome of B-cell tumors and predicts clinical outcome. Nature Cancer 1, 1066-1081 (2020)). The coverage of the original epiCMIT-CpGs based on Illumina 450k data in more targeted sequencing approaches like RRBS can be greatly compromised depending on the sequencing depth of samples or the enrichment towards particular regions of the genome. To overcome this, the epiCMIT-CpGs catalogue was expanded using high coverage whole genome bisulfite sequencing (WGBS) data from a previous publications including 15 samples covering the entire B-cell maturation spectrum (Kulis, M. et al. Whole-genome fingerprint of the DNA methylome during human B cell differentiation. Nat. Genet. 47, 746-756 (2015); Kretzmer, H. et al.

DNA methylome analysis in Burkitt and follicular lymphomas identifies differentially methylated regions linked to somatic mutation and transcriptional control. Nat. Genet. 47, 1316-1325 (2015); Kulis, M. et al. Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. Nat. Genet. 44, 1236-1242 (2012)). Briefly, the genome was segmented into 12 CHMM states with 200 bp resolution using the CHMM software (Ernst, J. & Kellis, M. ChromHMM: automating chromatin-state discovery and characterization. Nat. Methods 9, 215-216 (2012)) fed with 6 histone marks including H3K27ac, H3K4mel, H3K4me3, H3K36me3, H3K9me3 and H3K27me3 available for 15 normal and 16 neoplastic B cell samples. Normals included 6 naive B cells, 3 germinal-center B cells, 2 memory B cells and 3 tonsillar plasma cell samples. Neoplasia samples included 5 mantle cell lymphoma, 7 CLL and 4 multiple myeloma samples. These 12 chromatin states were ActProm (active promoter, with H3K27ac and H3K4me3), WkProm (weak promoter, with H3K4me1 and H3K4me3), PoisProm (poised promoter, with H3K27me3, H3K4me1 and H3K4me3), StrEnhl (strong enhancer 1, with H3K27ac, H3K4me1 and H3K4me3), StrEnh2 (strong enhancer 2, with H3K27ac and H3K4mel), WkEnh (weak enhancer, with H3K4mel), TxnTrans (transcription transition, with H3K36me3, H3K27ac and H3K4mel), TxnElong (transcription elongation, with H3K36me3), WkTxn (weak transcription, with low H3K36me3), H3K9me3 (H3K9me3-marked repressed heterochromatin), H3K27me3 (H3K27me3-marked repressed heterochromatin) and Het;LowSign (low-signal heterochromatin, with the absence of all six histone marks). Next, we selected CpGs located in repressive regions, including PoisProms, H3K27me3-repressed, H3K9me3 regions and Het;LowSign heterochromatin states. Afterwards, only CpGs showing extensive methylation differences (>0.5 difference in methylation fraction) between the lowly divided hematopoietic stem cell (HPC) and the highly divided bone-marrow plasma cells (bmPC) were retained, yielding 4,169 epiCMIT-hyper-CpGs (gaining methylation in H3K27me3 and PoisProm regions) and 808,872 epiCMIT-hypo-CpGs (CpGs losing methylation in H3K9me3 and Het;LowSign) in the hg38 genome assembly. Finally, the epiCMIT-hyper and epiCMIT-hypo scores were calculated as previously described (Duran-Ferrer, et al.) and the higher value in each sample was selected separately, which is different than the original strategy for Illumina array data where all samples shared the same epiCMIT-CpGs for the calculations (Duran-Ferrer, et al.) (only CpGs covered by at least 5 reads were used). This strategy was implemented to maximize the number of epiCMIT-CpGs in each sample, as only 124 and 311 epiCMIT-CpGs of the extended epiCMIT-CpGs catalogue were present in RRBS-SE and RRBS-PE consensus matrices, respectively. The new approach was validated using 24 samples profiled twice with different platforms, including 18 samples profiled with Illumina 450k and RRBS-PE, and 6 samples with RRBS-PE and RRBS-SE. In the samples profiled with Illumina 450k, the original epiCMIT-CpGs were used, whereas in RRBS data the available epiCMIT-CpGs was used in each sample of the extended catalogue of epiCMIT-CpGs based on WGBS data. These analyses showed that (i) the new epiCMIT approach was highly correlated with the original one, (ii) the epiCMIT could be calculated with varying numbers of epiCMIT-CpGs (with a minimum of around 800 epiCMIT-CpGs), and (iii) epiCMIT could be calculated with minimal impact due to different batches and platforms used. These statements were further supported by the PCA analyses with Illumina 450k data (ICGC cohort) and RRBS-SE data (DFCI and GCLLSG cohorts, n=93 and n=295, respectively) and RRBS-PE (data not shown), in which the epiCMIT gradient was similar in both platforms and unaffected by different cohorts.

H3K27Ac ChIP-Seq Analysis of Expression Clusters

[0362] To study the regulatory landscape of each ECs, previously analyzed cases with H3K27ac ChIP-seq were used (n=104), from which 70 cases had available RNA-seq and DNA methylation data. In these 70 samples, the number of cases for each expression cluster (EC) was: EC-m1=11, EC-ul=24, EC-m2=5, EC-o=2, EC-u2=5, EC-m3=10, EC-m4=12 and EC-i=1. From the 70 cases with available expression cluster (EC) classification, those expression clusters (ECs) with at least 5 cases (EC-o and EC-i were excluded) were selected and a differential analysis was performed using

DESeq2 (Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550 (2014)) with raw H3K27ac counts. Genome-wide analyses was performed comparing each expression cluster (EC) versus the others using a consensus matrix with 100,640 regions showing at least one H3K27ac peak in one of the 104 samples, and those regions with an FDR<0.05 in any of the comparisons were retained.

[0363] Additionally, differential analyses was performed focused on those regulatory regions associated with the marker genes of each expression cluster (EC). To do so, all expression cluster (EC) marker gene coordinates were selected and extended 2,000 bp upstream of their corresponding transcription start sites. These regions were then intersected with the consensus matrix (n=100,640) and a differential DESeq2 (Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550 (2014)) analysis was performed with each expression cluster (EC) versus all the others and identified regions with FDR<0.05. These results were used for the H3K27ac annotation of the marker genes.

Statistical Methods

[0364] Unless otherwise stated, two-sided t-test was used for mean comparison and multiple testing was corrected to compute false discovery rate (FDR, q) by the Benjamini-Hochberg procedure (Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J. R. Stat. Soc. 57, 289-300 (1995)). Categorical enrichments were computed using a two-sided Fisher's Exact test unless otherwise stated.

Clinical Outcome Modeling

[0365] Failure-free survival (FFS) was calculated for treatment-naive patients as the time from the date of the sequenced sample to the date of first treatment (“natural progression”), progression (if the patient was sampled at the time of enrollment on a clinical trial) or death, and censored at the last known event-free date. In the genetics-focused analysis (Tables 1A-1E and 2A-2E), the first event was defined as time to next treatment in patients who received therapy within 30 days. Subset analysis included patients who were treatment-naive at the time of the sequenced sample and not enrolled on a therapeutic clinical trial; in this analysis, time between sample and date of first treatment was used. Overall survival (OS) was calculated as the time from the date of the sequenced sample to the date of death and censored at the date last known alive. Univariate and multivariable Cox regression models were constructed for each subset of data. Final models were selected using the glmnet function for regularized Cox regression using an elastic net penalty within the Coxnet package in R. Ten-fold cross-validation using the cv.glmnet function with a partial-likelihood deviance metric to minimize k was performed and the minimum CV-error model was used. The alpha was set to 1 corresponding to a Lasso penalty. The maximum iterations (maxit) parameter was set to 1000. Features identified as having non-zero coefficient values using elastic net and selected in the final model were then included in a Cox regression model to obtain the hazard ratios. These hazard ratios estimated the magnitude of effect but p-values and confidence intervals are not readily interpretable in the elastic net model and are therefore not reported. For the integrated analysis of all available datatypes variables including expression cluster and epitype categories were dummy coded. Prognostic significance of expression cluster and IGHV (heavy chain variable region of immunoglobulin genes) status were also considered using a chi-squared test with the difference in −2 log likelihood (−2logL) between models including somatic single nucleotide variants (sSNVs) and somatic copy number alterations (sCNAs). The Breslow approximation was used for handling ties in survival time.

Non-Coding Driver Discovery Procedure

[0366] MutSig2CV-NC (Rheinbay, E. et al. Analyses of non-coding somatic drivers in 2,658 cancer whole genomes. Nature 578, 102-111 (2020)) (github.com/broadinstitute/getzlab-PCAWG-MutSig2CV NC.git) was first used to identify candidate non-coding drivers in different genomic regions including enhancers, 3′ UTRs, 5′ UTRs, promoters and lncRNA genes. Then the stringent

post-filtering steps described in detail in the Pan-cancer Analysis of Whole Genomes (PCAWG) Project's non-coding drivers paper (Bailey, et al) was followed on the candidate targets (q<0.5). In summary, the post-filters required: [0367] 1). at least three mutations are present in the candidate driver; [0368] 2). at least three patients have mutations in the candidate driver; [0369] 3). less than 50% of mutations are in palindromic DNA; [0370] 4). more than 50% of mutations are in mappable regions; [0371] 5). less than 35% of mutations have Activation-induced cytidine deaminase (AID)-related signatures attribution greater than 50%; or [0372] 6). mutations pass manual review in IGV. [0373] For candidate targets failing any of the above filters, their p-values were re-assigned to be 1. Finally, Benjamini-Hochberg multiple hypothesis correction was applied on the corrected p-values to get the post-filtered q-values. This provided 1 candidate (q<0.1): WDR74 which was reported in the aforementioned PCAWG paper (Rheinbay, et al). Additionally, RNA-seq analysis of mutated versus unmutated samples did not reveal a notable effect on gene expression of mutations in an extended list of candidate genes. Thus, novel non-coding drivers were not reported.

Mutational Signatures Review

[0374] By applying SignatureAnalyzer (Kim, J. et al. Somatic ERCC2 mutations are associated with a distinct genomic signature in urothelial tumors. Nat. Genet. 48, 600-606 (2016)) to 177 WGS, 8 mutational signatures were observed acting in these samples. A careful review suggested that three signatures (S5, S7, S8) might correspond to possible sequencing artifacts, and thus were removed from the main signatures plot depicting the 5 biological mutational processes identified by SignatureAnalyzer. Specifically, the cosine similarity between S5 and SBS51 (per COSMIC v3.1) is 0.82, while the cosine similarity between S8 and SBS50 (per COSMIC v3.1) is 0.74. S7 only contains one striking peak at G(T>G)G motif and thus it is assumed to be a bleed-through artifact.

OTHER EMBODIMENTS

[0375] From the foregoing description, it will be apparent that variations and modifications may be made to the invention described herein to adapt it to various usages and conditions. Such embodiments are also within the scope of the following claims.

[0376] The recitation of a listing of elements in any definition of a variable herein includes definitions of that variable as any single element or combination (or subcombination) of listed elements. The recitation of an embodiment herein includes that embodiment as any single embodiment or in combination with any other embodiments or portions thereof.

[0377] All patents and publications mentioned in this specification are herein incorporated by reference to the same extent as if each independent patent and publication was specifically and individually indicated to be incorporated by reference. The disclosure may be related to PCT/US2021/045144, filed Aug. 9, 2021, the disclosure of which is incorporated herein by reference in its entirety for all purposes.

Claims

1. A method of treating a selected subject having chronic lymphocytic leukemia (CLL), the method comprising administering one of the following agents to the subject, wherein the subject is selected as sensitive to the agent by having a corresponding feature selected from EC-i, EC-m1, EC-m2, EC-m3, EC-m4, EC-o, EC-ul, EC-u2, M-CLL, U-CLL or from one of the following: TABLE-US-00009 Agent Feature Direction AZD5991 EC-i Sensitive Azacitidine CARD11 Sensitive Cerdulatinib loss_13q14.13 Sensitive GSK690693 i-CLL Sensitive Rapamycin i-CLL Sensitive Rapamycin EC-m4 Sensitive Rapamycin M-CLL Sensitive Umbralisib M-CLL Sensitive Trametinib CHD2 Sensitive Bendamustine loss_12p13.31a Sensitive Bendamustine loss_5p15.33 Sensitive Bendamustine loss_7p22.2 Sensitive Bendamustine loss_14q32.12 Sensitive Cerdulatinib loss_13q14.3 Sensitive Cerdulatinib loss_13q14.13 Sensitive JQ1 FBXW7 Sensitive Navitoclax M-CLL Sensitive Rapamycin M-CLL Sensitive Ruxolitinib loss_13q14.3 Sensitive Venetoclax loss_13q14.3 Sensitive AZD5991 loss_3p13 Sensitive Cerdulatinib loss_13q14.3 Sensitive

Cerdulatinib loss_13q14.13 Sensitive Entospletinib tri_12 Sensitive GSK690693 i-CLL Sensitive JQ1 EC-i Sensitive Selinexor MYD88 Sensitive Trametinib CHD2 Sensitive Vorinostat EC-m2 Sensitive

2. A method of treating a sensitive subject having chronic lymphocytic leukemia (CLL), the method comprising: administering an agent to the sensitive subject, wherein the subject's sensitivity is determined by identifying the presence of a feature selected from EC-i, EC-m, EC-m2, EC-m3, EC-m4, EC-o, EC-ul, EC-u2, or from among the following expression subtypes, drives, genetic alterations, or CLL subtypes, or electing not to administer an agent to a resistant subject wherein the subject's resistance is determined by identifying the presence of a feature selected from EC-i, EC-m, EC-m2, EC-m3, EC-m4, EC-o, EC-ul, EC-u2, or from among the following expression subtypes, drives, genetic alterations, or CLL subtypes, wherein the agent, feature, and sensitivity or resistance includes: TABLE-US-00010 Agent Feature Direction A-1331852 EC-m2 Resistant AZD5991 EC-i Sensitive Azacitidine CARD11 Sensitive Cerdulatinib loss_13q14.13 Sensitive GSK690693 i-CLL Sensitive Gandotinib EC-i Resistant Navitoclax EC-m2 Resistant Nutlin-3 priortrt_Post Resistant Nutlin-3 FBXW7 Resistant Rapamycin EC-u1 Resistant Rapamycin i-CLL Sensitive Rapamycin U-CLL Resistant Rapamycin EC-m4 Sensitive Rapamycin n-CLL Resistant Rapamycin M-CLL Sensitive Umbralisib U-CLL Resistant Umbralisib n-CLL Resistant Umbralisib M-CLL Sensitive Venetoclax EC-m2 Resistant Rapamycin n-CLL Resistant Trametinib CHD2 Sensitive Bendamustine loss_12p13.31a Sensitive Bendamustine loss_5p15.33 Sensitive Bendamustine loss_7p22.2 Sensitive Bendamustine loss_14q32.12 Sensitive Cerdulatinib loss_13q14.3 Sensitive Cerdulatinib loss_13q14.13 Sensitive Gandotinib EC-i Resistant JQ1 FBXW7 Sensitive MK-2206 priortrt_Post Resistant Navitoclax U-CLL Resistant Navitoclax n-CLL Resistant Navitoclax M-CLL Sensitive Nutlin-3 priortrt_Post Resistant Rapamycin U-CLL Resistant Rapamycin M-CLL Sensitive Ruxolitinib loss_13q14.3 Sensitive Venetoclax loss_13q14.3 Sensitive AZD5991 loss_3p13 Sensitive Cerdulatinib loss_13q14.3 Sensitive Cerdulatinib loss_13q14.13 Sensitive Entospletinib tri_12 Sensitive GSK690693 i-CLL Sensitive JQ1 EC-i Sensitive Rapamycin n-CLL Resistant Selinexor MYD88 Sensitive Trametinib CHD2 Sensitive Vorinostat EC-m4 Resistant Vorinostat EC-m2 Sensitive drug feature direction A-1331852 EC-m2 Resistant AZD5991 EC-i Sensitive Azacitidine CARD11 Sensitive Cerdulatinib loss_13q14.13 Sensitive GSK690693 i-CLL Sensitive Gandotinib EC-i Resistant Navitoclax EC-m2 Resistant Nutlin-3 priortrt_Post Resistant Nutlin-3 FBXW7 Resistant Rapamycin EC-u1 Resistant Rapamycin i-CLL Sensitive Rapamycin U-CLL Resistant Rapamycin EC-m4 Sensitive Rapamycin n-CLL Resistant Rapamycin M-CLL Sensitive Umbralisib U-CLL Resistant Umbralisib n-CLL Resistant Umbralisib M-CLL Sensitive Venetoclax EC-m2 Resistant Rapamycin n-CLL Resistant Trametinib CHD2 Sensitive Bendamustine loss_12p13.31a Sensitive Bendamustine loss_5p15.33 Sensitive Bendamustine loss_7p22.2 Sensitive Bendamustine loss_14q32.12 Sensitive Cerdulatinib loss_13q14.3 Sensitive Cerdulatinib loss_13q14.13 Sensitive Gandotinib EC-i Resistant JQ1 FBXW7 Sensitive MK-2206 priortrt_Post Resistant Navitoclax U-CLL Resistant Navitoclax n-CLL Resistant Navitoclax M-CLL Sensitive Nutlin-3 priortrt_Post Resistant Rapamycin U-CLL Resistant Rapamycin M-CLL Sensitive Ruxolitinib loss_13q14.3 Sensitive Venetoclax loss_13q14.3 Sensitive AZD5991 loss_3p13 Sensitive Cerdulatinib loss_13q14.3 Sensitive Cerdulatinib loss_13q14.13 Sensitive Entospletinib tri_12 Sensitive GSK690693 i-CLL Sensitive JQ1 EC-i Sensitive Rapamycin n-CLL Resistant Selinexor MYD88 Sensitive Trametinib CHD2 Sensitive Vorinostat EC-m4 Resistant Vorinostat EC-m2 Sensitive A-1331852 EC-m2 Resistant AZD5991 EC-i Sensitive Azacitidine CARD11 Sensitive Cerdulatinib loss_13q14.13 Sensitive GSK690693 i-CLL Sensitive Gandotinib EC-i Resistant Navitoclax EC-m2 Resistant Nutlin-3 priortrt_Post Resistant Nutlin-3 FBXW7 Resistant Rapamycin EC-u1 Resistant Rapamycin i-CLL Sensitive Rapamycin U-CLL Resistant Rapamycin EC-m4 Sensitive Rapamycin n-CLL Resistant Rapamycin M-CLL Sensitive Umbralisib U-CLL Resistant Umbralisib n-CLL Resistant Umbralisib M-CLL Sensitive Venetoclax EC-m2 Resistant Rapamycin

n-CLL Resistant Trametinib CHD2 Sensitive Bendamustine loss_12p13.31a Sensitive Bendamustine loss_5p15.33 Sensitive Bendamustine loss_7p22.2 Sensitive Bendamustine loss_14q32.12 Sensitive Cerdulatinib loss_13q14.3 Sensitive Cerdulatinib loss_13q14.13 Sensitive Gandotinib EC-i Resistant JQ1 FBXW7 Sensitive MK-2206 priortrt_Post Resistant Navitoclax U-CLL Resistant Navitoclax n-CLL Resistant Navitoclax M-CLL Sensitive Nutlin-3 priortrt_Post Resistant Rapamycin U-CLL Resistant Rapamycin M-CLL Sensitive Ruxolitinib loss_13q14.3 Sensitive Venetoclax loss_13q14.3 Sensitive AZD5991 loss_3p13 Sensitive Cerdulatinib loss_13q14.3 Sensitive Cerdulatinib loss_13q14.13 Sensitive Entospletinib tri_12 Sensitive GSK690693 i-CLL Sensitive JQ1 EC-i Sensitive Rapamycin n-CLL Resistant Selinexor MYD88 Sensitive Trametinib CHD2 Sensitive Vorinostat EC-m4 Resistant Vorinostat EC-m2 Sensitive

3. The method of claim 1, wherein Ec-i comprises an increase in one or more of GRIK3, IQGAP2, FCER1G, STK32B, GADD45A, ITGAX, KLF3, RFTN1, PTK2, DFNB31, and ZMAT1 polypeptides, or nucleic acid molecules encoding said polypeptides; wherein EC-m1 comprises an increase in one or more of TFEC, COL18A1, SLC19A1, NRIP1, KCNH2, P2RX1, ARRDC5, BEX4, and APP polypeptides, or nucleic acid molecules encoding said polypeptide; wherein EC-m2 comprises an increase in one or more of EML6, HCK, CD1C, VPS37B, CYBB, NXPH4, BTNL9, KLRK1, IQSEC1, BANKI, LEF1, SH3D21, FMOD, SEMA4A, CTLA4, ADTRP, IGSF3, IGFBP4, PDGFD, and APOD polypeptides, or nucleic acid molecules encoding said polypeptide; wherein EC-m3 comprises an increase in MS4A4E, MYL9, NT5E, MS4A6A, PITPNC1, CNTNAP2, IGF2BP3, WNT3, CLDN7, TCF7, BASP1, FLJ20373, MAP4K4, LRRK2, SAMSN1, CEACAMI, TNFRSF13B, PHF16, MID1IPI, and ABCA9 polypeptides or nucleic acid molecules encoding said polypeptides; wherein EC-m4 comprises an increase in MYBL1, NUGGC, GNG8, AEBP1, HIP1R, LATS2, RIMKLB, EML6, FADS3, MBOAT1, LCN10, DCLK2, and GLUL polypeptides or nucleic acid molecules encoding said polypeptide; wherein EC-o comprises ACSM3, TOX2, PHF16, SESN3, TBC1D9, PIP5K1B, SIK1, DUSP5, GNG7, HIVEP3, MARCKSL1, GPR183, HRK, and PITPNC1, or nucleic acid molecules encoding said polypeptides; wherein EC-ul comprises an increase in SEPT10, LDOC1, LPL, KANK2, SOWAHC, DUSP26, OSBPL5, WNT9A, FGFR1, GTSF1L, ADD3, AKT3, COBLL1, MNDA, FCRL3, FAM49A, FCRL2, SLC2A3, and MARCKS polypeptides, or nucleic acid molecules encoding said polypeptide; or wherein EC-u2 comprises ITGB5, BCL7A, PPP1R9A, TSPAN13, SLC12A7, SSBP3, VASH1, SPG20, IL13RA1, NR3C2, TUBG2, ZNF804A, and IL2RA polypeptides, or nucleic acid molecules encoding said polypeptides.

4. The method of claim 3, wherein levels of the polypeptide and polynucleotide are increased.

5. The method of claim 1, wherein a subject having a characterized CLL is treated as follows: M-CLL is treated with navitoclax, nutlin-3, duvelisib, ibrutinib, or venetoclax; or U-CLL is treated with navitoclax, nutlin-3, duvelisib, ibrutinib, dasatinib, venetoclax, or idelasib.

6. The method of claim 1, wherein venetoclax is administered in combination with an MCL1 inhibitor.

7. The method of claim 1, wherein a subject having a characterized CLL is treated as follows: EC-m3 is treated with venetoclax in combination with an MCL1 inhibitor; EC-m2, M-CLL, and having a trisomy-12 driver is administered zanubrutinib; or EC-i is administered abexinostat.

8. The method of claim 1, wherein a subject receiving venetoclax is also administered one or more of the following: abexinostat, navitoclax, cerdulatinib, AZD5991, atorvastatin, zanubrutinib, GSK690693, trametinib, ponatinib, bendamustine, nutlin-3, and rapamycin.

9. The method of claim 1, wherein a subject receiving venetoclax is administered two or more of the following: navitoclax, abexinostat, dasatinib, idelaslisib, duvelisib, cerdulatinib, bendamustine, GSK690693, nirogacestat, trametinib, and rapamycin.

10. The method of claim 1, wherein a subject having the following expression subtype is treated as follows: EC-i expression subtype, the method comprising administering to the subject navitoclax; EC-m1 expression subtype, the method comprising administering to the subject nutlin-3,

navitoclax, or cerdulatinib; EC-m2 expression subtype, the method comprising administering to the subject abexinostat, duvelisib, idelalisib, entospletinib, or vorinostat; EC-m3 expression subtype, the method comprising administering to the subject venetoclax, navitoclax, or Abexinostat; EC-m4 expression subtype, the method comprising administering to the subject navitoclax, nutlin-3, or gandotinib; or EC-o expression subtype, the method comprising administering to the subject gandotinib, abexinostat, or cerdulatinib; EC-ul expression subtype, the method comprising administering to the subject gandotinib; EC-u2 expression subtype, the method comprising administering to the subject ibrutinib, A-1331852, navitoclax, or rapamycin; M-CLL subtype, the method comprising administering to the subject navitoclax or abexinostat; or U-CLL subtype, the method comprising administering to the subject A-1331852, 25 atorvastatin, AZD5991, bendamustine, onalespib, trametinib, voruciclib, or zanubrutinib.

11. The method of claim 1, wherein venetoclax is administered in combination with an MCL1 inhibitor selected from the group consisting of AZD5991, tapotoclax, MIK665, A-1210477, ANJ810, PRT1419, AS00491, APG-3526, CT-03, and CPT-628.

12. The method of claim 1, wherein the subject is selected as comprising a driving alteration, wherein the driving alteration is: (a) in a gene encoding a polypeptide selected from the group consisting of ATM, CARD11, CHD2, FBXW7, ITIH2, NOTCH1, NRAS, POT1, SF3B1, TP53, and ZMYM3; or (b) in a genomic region selected from the group consisting of 7922.1, 15q24.2, 16pl1.2, 19pl3.3, 1921.3, 1942.13, 2p11.2, 2q31.1, 3p21.31, 3pl3, 5pl5.33, 7p22.2, 9q34.3, 1Op12.2, 10q24.2, 10q24.32, 11q22.3, 12pl3.31a, 13q14.13, 13q14.3, 14q32.12, 16q22.1, 17pl3.3, 17p13.1, and chromosome 12, and/or 2p.

13. The method of claim 1, wherein CLL is further characterized as having: i. a mutated (M-CLL) or unmutated IGHV (U-CLL) subtype; ii. an expression subtype selected from EC-i, EC-m1, EC-m2, EC-m3, EC-m4, EC-o, EC-ul, or EC-u2; and/or iii. a driving alteration in a gene encoding a polypeptide selected from the group consisting of ATM, CARD11, CHD2, FBXW7, ITIH2, NOTCHI, NRAS, POT1, SF3B1, TP53, and ZMYM3; and/or iv. a driving alteration in a genomic region selected from the group consisting of 7q22.1, 15q24.2, 16pl1.2, 19pl3.3, 1921.3, 1942.13, 2pl1.2, 2q31.1, 3p21.31, 3pl3, 5p15.33, 7p22.2, 9q34.3, 1Op12.2, 10q24.2, 10q24.32, 11q22.3, 12p13.31a, 13q14.13, 13q14.3, 14q32.12, 16q22.1, 17pl3.3, 17pl3.1, chromosome 12, and 2p; and wherein the agent has a delta priming value listed in FIG. **14** greater than 15 associated with the CLL subtype or driving alteration.

14. The method of claim 1, further comprising characterizing the CLL as having: i. a mutated (M-CLL) or unmutated IGHV (U-CLL) subtype; ii. an expression subtype selected from EC-i, EC-m1, EC-m2, EC-m3, EC-m4, EC-o, EC-ul, or EC-u2; and/or iii. a driving alteration in a gene encoding a polypeptide selected from the group consisting of ATM, CARD11, CHD2, FBXW7, ITIH2, NOTCHI, NRAS, POT1, SF3B1, TP53, and ZMYM3; and/or iv. a driving alteration in a genomic region selected from the group consisting of 7q22.1, 15q24.2, 16pl1.2, 19pl3.3, 1921.3, 1942.13, 2pl1.2, 2q31.1, 3p21.31, 3pl3, 5pl5.33, 7p22.2, 9q34.3, 1Op12.2, 10q24.2, 10q24.32, 11q22.3, 12p13.31a, 13q14.13, 13q14.3, 14q32.12, 16q22.1, 17pl3.3, 17pl3.1, chromosome 12, and 2p; and (b) selecting the subject for inclusion in the clinical trial if the agent has a positive delta priming value of greater than 15 listed in FIG. **14** for the subtype and/or driving alteration of the CLL, and otherwise excluding the subject from the clinical trial.

15. The method of claim 14, wherein the driving alteration to the genomic region is a duplication or a deletion.

16. A combination therapeutic comprising venetoclax and an agent having a delta priming value listed in FIG. **14** greater than 15, abexinostat, navitoclax, cerdulatinib, AZD5991, atorvastatin, zanubrutinib, GSK690693, trametinib, ponatinib, bendamustine, nutlin-3, and rapamycin, or a MCL1 inhibitor.

17. The combination therapeutic of claim 16, wherein venetoclax and the agent are formulated separately.

18. The combination therapeutic of claim 16, wherein venetoclax and the agent are administered concurrently.
19. The combination therapeutic of claim 16, wherein venetoclax and the agent are administered sequentially within at least about 1, 3, 6, 9, 12, or 24 hours of one another.
20. The combination therapeutic of claim 16, wherein the MCL1 inhibitor is selected from the group consisting of AZD5991, tapotoclax, MIK665, A-1210477, ANJ810, PRT1419, AS00491, APG-3526, CT-03, and CPT-628.