



(51) International Patent Classification:

C12Q 1/6886 (2018.01) C12P 19/34 (2006.01)
G01N 33/574 (2006.01)

(21) International Application Number:

PCT/US2019/019413

(22) International Filing Date:

25 February 2019 (25.02.2019)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/634,508 23 February 2018 (23.02.2018) US

(71) Applicant: **BETH ISRAEL DEACONESS MEDICAL CENTER** [US/US]; 300 Brookline Ave, Boston, MA 02215 (US).(72) Inventors: **PANDOLFI, Pier Paulo**; c/o Beth Israel Deaconess Medical Center, 330 Brookline Ave, Boston, MA 02215 (US). **LEE, Jonathan**; c/o Beth Israel Deaconess Medical Center, 330 Brookline Ave, Boston, MA 02215 (US). **BESTER, Assaf**; c/o Beth Israel Deaconess Medical Center, 330 Brookline Ave, Boston, MA 02215 (US). **CLOHESSY, John G.**; c/o Beth Israel Deaconess Medical Center, 330 Brookline Ave, Boston, MA 02215 (US).(74) Agent: **ALTIERI, Stephen L.** et al.; Morgan, Lewis & Bockius LLP, 1111 Pennsylvania Avenue, NW, Washington, DC 20004 (US).

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

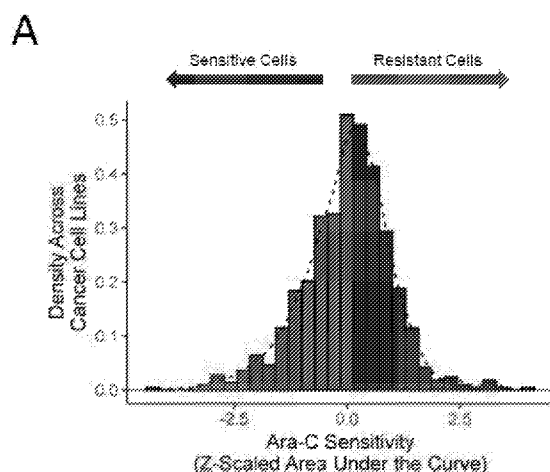
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,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

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: DRUG EFFICACY EVALUATIONS

FIGURE 1



(57) Abstract: The present disclosure provides materials and methods for predicting drug efficacy, for example, computational approaches for predicting drug efficacy or resistance in proliferative disorders, such as cancer.

DRUG EFFICACY EVALUATIONS**FIELD**

[0001] The disclosure relates, *inter alia*, to materials and methods for predicting drug efficacy, for example, computational approaches for predicting drug efficacy or resistance in proliferative disorders, (*e.g.*, cancer).

PRIORITY

[0002] This application claims the benefit of, and claims priority to, U.S. Provisional Application No. 62/634,508, filed February 23, 2018, the contents of which is hereby incorporated by reference in its entirety.

GOVERNMENT INTEREST

[0003] This invention was made with government support under R35 CA19752 awarded by the National Institute of Health. The government has certain rights in the invention.

DESCRIPTION OF THE TEXT FILE SUBMITTED ELECTRONICALLY

[0004] The contents of the text file submitted electronically herewith are incorporated herein by reference in its entirety: A computer readable format copy of the Sequence Listing (filename: BID-006PC_ST25.txt; date created: February 22, 2019; file size: 24,512 bytes).

BACKGROUND

[0005] Chemotherapy remains the initial, and last, line of defense for most cancer patients. However, approximately 30% to 50% of patients relapse with chemotherapy-resistant disease, representing a fatal prognosis for patients. Thus, there is a need to better understand the genetic and molecular mechanisms that contribute to chemotherapeutic resistance. Precision medicine and targeted therapies offer new hope for the specific elimination of genetically defined cancers. Forward genetic screening has proven to be a useful tool for the identification of protein coding genes affecting phenotypes *in vitro* and *in vivo*, however, technological barriers have limited the ability to study non-coding genes in a similar manner. Clustered Regularly Interspaced Short Palindromic Repeats (CRISPRa), which takes advantage of the CRISPR gene targeting strategy for gene activation, is a promising approach for the study of non-coding genes. Accordingly, the

development of an integrative and comprehensive CRISPR activation (CRISPRa) framework that would complement publicly available ‘Big Data’ databases to enable the discovery of functional human protein coding and long non-coding RNAs (lncRNA) genes that contribute to mechanisms of chemotherapeutic resistance is needed.

SUMMARY

[0006] The present disclosure is based, in part, on the discovery that a “priori” cancer subtype regression model that incorporates both protein-coding and noncoding gene expression biomarkers are useful for predicting drug efficacy. The disclosure further, provides a comprehensive and integrative genome-wide study of both the coding and non-coding genes that contribute to mechanisms of chemotherapeutic resistance.

[0007] In one aspect, the disclosure provides methods for determining an RNA signature indicative of sensitivity or resistance to a candidate cancer therapy, the method comprising: identifying one or more coding RNAs whose level of expression correlates with sensitivity or resistance to a candidate cancer therapy in cells of a cancer subtype, and identifying one or more noncoding RNAs whose level of expression correlates with sensitivity or resistance to the candidate cancer therapy in cells of the cancer subtype, thereby determining an RNA signature that is indicative of sensitivity or resistance to the candidate cancer therapy; and determining the presence or absence of the RNA signature in a cell or tissue of the cancer subtype, and classifying the cell or tissue for sensitivity or resistance to the candidate cancer therapy.

[0008] In some embodiments, the cancer subtype is defined by tissue of origin and histological subtype. In some embodiments, the cancer subtype is a solid tumor. In some embodiments, the cancer subtype is a hematological malignancy.

[0009] In some embodiments, cancer subtype is a sarcoma, which is optionally squamous cell carcinoma, fibrosarcoma, myosarcoma, osteogenic sarcoma, angiosarcoma, or endotheliosarcoma.

[0010] In some embodiments, the cancer subtype is a carcinoma, which is optionally adenocarcinoma.

[0011] In some embodiments, the cancer subtype is Small Cell Lung Cancer (CLC), Non-Small Cell Lung Cancer (NSCLC), or mesothelioma.

[0012] In some embodiments, the cancer subtype is a brain cancer or glioblastoma.

[0013] In some embodiments, the cancer subtype is a breast cancer, lymphoma, prostate cancer, pancreatic cancer, liver cancer, kidney cancer, colon or colorectal cancer, ovarian cancer, endometrial cancer, cervical cancer, testicular cancer, or melanoma.

[0014] In some embodiments, the cancer is a leukemia, which is optionally acute myeloid leukemia (AML), chronic myelogenous leukemia (CML), or acute lymphoblastic leukemia (ALL).

[0015] In some embodiments, the cells of the cancer subtype are cell lines.

[0016] In some embodiments, the cells of the cancer subtype are primary cell cultures.

[0017] In some embodiments, the cells of the cancer subtype comprise at least about 50 or at least about 100 or at least about 200 cell lines or primary cultures.

[0018] In some embodiments, the sensitivity or resistance of the cell lines or primary cell cultures to the candidate cancer therapy is determined by a chemosensitivity assay.

[0019] In some embodiments, one or more coding RNAs are involved in DNA replication, Cell cycle, Pyrimidine metabolism, Homologous recombination, p53 signaling pathway, Base excision repair, Nucleotide excision repair, Mismatch repair, One carbon pool by folate, Non-homologous end-joining, Citrate cycle (TCA cycle), Apoptosis, Cellular senescence, MAPK signaling pathway, PI3K-Akt signaling pathway, Jak-STAT signaling pathway, Hematopoietic cell lineage, Oxidative phosphorylation, Fatty acid degradation, Cytokine-cytokine receptor interaction, Ribosome, RNA transport, mRNA surveillance pathway, RNA degradation, Spliceosome, and/ or Purine metabolism.

[0020] In some embodiments, the non-coding RNAs are long non-coding RNAs.

[0021] In some embodiments, an RNA signature includes the expression levels of one or more coding and non-coding RNA gene pairs, and optionally at least 5 or at least 10 coding and non-coding gene pairs. In some embodiments, the RNA signature includes the expression levels for at least about 10 RNAs, or at least about 15 RNAs, or at least about 25 RNAs, or at least about 50 RNAs.

[0022] In some embodiments, the candidate cancer therapy is a chemotherapy. In some embodiments, the chemotherapy comprises cytarabine, paclitaxel, docetaxel, doxorubicin, platinum-based chemotherapeutics that is optionally cisplatin or carboplatin, mitomycin, mithramycin, methotrexate, 5-fluorouracil, vinorelbine, topotecan, irinotecan, bleomycin,

actinomycin, topoisomerase I and II inhibitors, anthracycline, epirubicin, idarubicin, mitoxantrone, valrubicin, etoposide, teniposide, and rubitecan.

[0023] In some embodiments, the candidate cancer therapy is a chemotherapy combination. In some embodiments, the combination is TFAC (paclitaxel, 5-fluorouracil, doxorubicin, and cyclophosphamide), EC (epirubicin and cyclophosphamide), or FEC (5-fluorouracil, epirubicin, and cyclophosphamide).

[0024] In some embodiments, the level of expression of the coding and non-coding RNAs is determined by a hybridization assay, RNA sequencing, or quantitative PCR. In some embodiments, the levels of expression of the coding and non-coding RNAs is determined in a biopsy sample from a subject, and the presence of the RNA signature determined. In some embodiments, the levels of expression of the coding and non-coding RNAs in the biopsy sample is determined using quantitative PCR, RNA sequencing, or hybridization assay. In some embodiments, the levels of expression of the coding and non-coding RNAs is determined in the cell lines or primary cell cultures using a hybridization assay or RNA sequencing, and the levels of expression of the coding and non-coding RNAs is determined in the biopsy sample using quantitative PCR.

[0025] In some embodiments, the cell or tissue is classified using one or more classification schemes selected from a Correlations analysis, Principal Components Analysis, Naïve Bayes, Support Vector Machines, Nearest Neighbors, Decision Trees, Logistic, Artificial Neural Networks, and Rule-based schemes.

[0026] In another aspect, the disclosure provides methods for treating cancer in a subject in need thereof. The method comprising determining an RNA signature indicative of sensitivity or resistance to a candidate cancer therapy according to the method of any one of the above embodiments or aspect and administering to the subject a cancer therapy that the cancer cell or tissue is classified as sensitive to.

[0027] In some embodiments, the cancer therapy is a chemotherapy, e.g., cytarabine, paclitaxel, docetaxel, doxorubicin, platinum-based chemotherapeutics that is optionally cisplatin or carboplatin, mitomycin, mithramycin, methotrexate, 5-fluorouracil, vinorelbine, topotecan, irinotecan, bleomycin, actinomycin, topoisomerase I and II inhibitors, anthracycline, epirubicin, idarubicin, mitoxantrone, valrubicin, etoposide, teniposide, or rubitecan.

[0028] In some embodiments, the candidate cancer therapy is a chemotherapy combination, e.g., TFAC (paclitaxel, 5-fluorouracil, doxorubicin, and cyclophosphamide), EC (epirubicin and cyclophosphamide), or FEC (5-fluorouracil, epirubicin, and cyclophosphamide).

[0029] In yet another aspect, the disclosure provides a composition for use in the method of any of the above aspects or embodiments.

[0030] Any aspect or embodiment described herein can be combined with any other aspect or embodiment as disclosed herein.

BRIEF DESCRIPTION OF THE DRAWINGS

[0031] The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawings will be provided by the Office upon request and payment of the necessary fee.

[0032] Figure 1A-F is a series of graphs illustrating that identification of protein-coding and noncoding gene biomarkers correlated with differential Cytarabine (1- β -darabinofuranosylcytosine, Ara-C) Response. Figure 1A shows the distribution of Ara-C drug sensitivities across 760 pan-cancer cell lines profiled by both CCLE and CTD2 studies, quantified by their Z-scaled area under the dose response curve values after regressing out lineage-specific effects. Figure 1B shows distribution of Z-scaled drug resistance-gene expression. Pearson correlation values of all analyzed genes. Representative protein-coding and noncoding gene symbols enriched beyond a Z-score threshold of ± 1.16 are demarcated. Figure 1C shows the summary of gene set enrichment analysis (GSEA) of protein-coding genes ranked by drug resistance-gene expression correlation values using annotated KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways. Figure 1D shows representative KEGG pathways from GSEA of protein-coding genes ranked by drug resistance-gene expression correlation values as shown in Figures 1B-1C. Figure 1E is a histogram showing the Pearson correlation distributions of gene pair expression levels in the cancer cell line panel across 997 sense-antisense cognate gene pairs and 5,000 random protein coding-lncRNA gene pairs. Wilcoxon rank-sum test: $p < 2.2e-16$. Figure 1F shows the relationship of drug resistance-gene expression correlation values between protein coding-lncRNA gene pairs across 997 sense-antisense cognate gene pairs (left panel: Pearson's $R = 0.552$, $p < 2.2e-16$) and 5,000 random gene pairs (right panel: Pearson's $R = 0.021$, $p = 0.1338$).

[0033] Figure 2A-I are a series of graphs illustrating CRISPRa functional screening of coding genes modulating Ara-C response. Figure 2A shows the distribution of Ara-C IC₅₀ values across a panel of AML cell lines. Figure 2B shows the Effect of BCL2 overexpression (middle data line) or DCK knockdown (right data line) on sensitivity to Ara-C in MOLM14 cells. Data are represented as mean $\pm$ SD, n = 3. Figure 2C shows a schematic of CRISPRa pooled screening for the identification of genes whose activation modulate sensitivity to Ara-C in MOLM14 cells. Figure 2D shows a Volcano plot summarizing the global changes in sgRNA representation of protein-coding genes before and after 14 days of treatment with Ara-C. A subset of genes validated herein (PI4K2A, MUL1, SETBP1, TUFT1, ZBP1, PXDC1, and CXCL17) or previously annotated (BCL2, GAS6, and DCK) to modulate Ara-C sensitivity are labeled. A false discovery rate threshold of 0.339 was determined by receiver operating characteristic analysis. Red data points to right of vertical dotted line – enrichment in the CRISPRa screening; blue data points to left of vertical dotted line – depletion in the CRISPRa screening; open black circles – genes previously associated with differential Ara-C sensitivity and above the significance threshold; filled black points – genes validated herein. Figure 2E shows a summary of gene set enrichment analysis (GSEA) of protein-coding genes ranked by CRISPRa screening using annotated KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways. Figure 2F shows a disease-free survival association with expression levels of ZBP1, MUL1, and PI4K2A, genes enriched in both protein-coding CRISPRa screening and drug resistance-gene expression correlation analyses, among patients treated with Ara-C therapy within the TCGA-LAML patient cohort. ZBP1: VST expression level cutoff = 6.13 (low, n = 42; high, n = 79), log-rank test: p-value = 0.0074. MUL1: VST expression level cutoff = 9.64 (low, n = 108; high, n = 13), log-rank test: p-value = 0.0033. PI4K2A: VST expression level cutoff = 7.23 (low, 36; high, n = 85), log-rank test: p-value = 0.038; in this figure, the high expression data lines (in red) are below the low expression data lines (in black). Figure 2G shows Ara-C efficacy measurements in MOLM14 cells expressing sgRNAs targeting ZBP1, MUL1, or PI4K2A based on normalized MTS reads following 48 hours of treatment. Data are represented as mean $\pm$ SD, n = 3; in this figure, the high expression data lines (in red) are above the low expression data lines (in black). Figure 2H shows Modulation of apoptotic response upon stable expression of sgRNAs targeting ZBP1, MUL1, or PI4K2A in MOLM14 cells. The percentage of apoptosis is determined by

annexin V and propidium iodide (PI) staining of cells treated with 0.25 μ M Ara-C for 72 hours. Figure 2I shows the Proliferation of unchallenged MOLM14 cells expressing sgRNAs targeting ZBP1, MUL1, or PI4K2A. Proliferation is quantified over four days (D1-D4). Figure 2I shows proliferation of unchallenged MOLM14 cells expressing sgRNAs targeting ZBP1, MUL1, or PI4K2A. Proliferation is quantified over four days (D1-D4).

[0034] Figure 3A-D are a series of graphs illustrating CRISPRa functional screening of noncoding genes modulating Ara-C response. Figure 3A shows a summary of the CaLR library design specifications, including lncRNA gene numbers, transcriptional start sites (TSS), and total sgRNA numbers and relationships between coding genes and lncRNA genes for corresponding lncRNA classifications; in this figure, the arrow representing the Coding Gene is to the right side of each line. Figure 3B is a volcano plot summarizing the global changes in sgRNA representation of noncoding genes before and after 14 days of treatment with Ara-C. A subset of genes previously annotated in various cancer related pathways (PVT1, HOTAIRM1, and TUG1) or validated herein to modulate Ara-C sensitivity (remaining listed genes) or are labeled. A false discovery rate threshold of $3.51\text{e-}5$ was determined by analysis of non-targeting sgRNA negative controls at the transcript level. Red data points to right of vertical dotted line – enrichment in the CRISPRa screening; blue data points to left of vertical dotted line points – depletion in the CRISPRa screening; filled black points – genes validated herein. Figure 3C shows the percentages of significantly enriched or depleted protein-coding or noncoding genes from CRISPRa screens detected in the TCGA-LAML patient samples. Chi-squared test: ***, $p=6.92\text{e-}3$. Figure 3D shows the gene expression level distributions of significantly enriched or depleted protein-coding or noncoding genes from CRISPRa screens detected in the TCGA-LAML patient samples. Wilcoxon rank-sum test: ***, $p=5.4\text{e-}7$. Figure 3E shows a guilt-by-association pathway annotation of enriched genes identified in the CaLR screen, KEGG pathway gene sets were used for this analysis.

[0035] Figure 4A-G are a series of graphs illustrating the validation of CaLR screening results. Figure 4A shows a Fold change (FC) of MOLM14 cell viability treated with 0.25 μ M Ara-C for 48 hours. Figure 4B shows a Fold change (FC) of expression levels of targeted lncRNAs upon overexpression of enriched sgRNAs versus endogenous levels. Figure 4C shows the Ara-C efficacy measurements in MOLM14 cells expressing sgRNAs targeting indicating genes based

on normalized MTS reads following 48 hours of treatment with the indicated concentrations of Ara-C (data in red upper lines are from the indicated genes). Figure 4D shows the proliferation of unchallenged MOLM14 cells expressing sgRNAs targeting indicating genes (data in red upper lines are from the indicated genes). Proliferation is quantified over four days (D1-D4). Figure 4E shows the modulation of apoptotic response upon stable expression of sgRNAs targeting a panel of significantly enriched sgRNAs as determined through CaLR screening in MOLM14 cells. The percentage of apoptosis is determined by annexin V and propidium iodide (PI) staining of MOLM14 cells stably infected with individual sgRNAs and treated with 0.25 μ M Ara-C for 72 hours. Right panel: representative flow cytometry plots of annexin V/PI staining intensities corresponding to two sgRNAs promoting survival versus non-targeting control. Figure 4F shows immunofluorescence images for DAPI and phospho- γ H2A.X staining in MOLM14 cells stably infected with sgRNAs targeting the lncRNA genes shown, and treated with 25 μ M Ara-C for 24 hours. Figure 4G shows the disease-free survival association with expression levels of GAS6-AS2 and AC008073.2, genes enriched in both noncoding CRISPRa screening and drug resistance-gene expression correlation analyses, among patients treated with Ara-C therapy within the TCGA-LAML patient cohort. GAS6-AS2: VST expression level cutoff = 3.38 (low, n = 92; high, n = 29), log-rank test: p-value = 0.035. AC008073.2: VST expression level cutoff = 4.39 (low, n = 93; high, n = 28), log-rank test: p-value = 0.0026; in this figure, the high expression data lines (in red) are below the low expression data lines (in black).

[0036] Figure 5A-I are a series of graphs illustrating that GAS6-AS2 promotes drug resistance *in vitro* and *in vivo*. Figure 5A shows the integration of drug resistance-gene expression correlative analysis and forward genetic screenings identifies seven sense-antisense gene pairs which pass all significance thresholds, a higher number than expected by chance alone (Chi-squared test: p = 9.85e-7). Figure 5B shows the fold change (FC) of MOLM14 cell viability treated with 0.25 μ M Ara-C for 48 hours. Cells expressing individual sgRNAs targeting GAS6-AS2. Figure 5C shows the Pearson correlation between cell viability versus GAS6-AS2 expression level for each of the 8 sgRNAs targeting GAS6-AS2. Figure 5D shows the Ara-C efficacy measurements in MOLM14 cells expressing sgRNAs #1 and #3 targeting GAS6-AS2 (top data lines in red) based on normalized MTS reads following 48 hours of treatment. Figure 5E (left panel) shows the representative flow cytometry data of MOLM14 cells expressing either control or GAS6-AS2-

targeting sgRNAs, treated with 25 μ M Ara-C for 24 hours and labeled with viability (propidium iodide (PI)) and apoptotic (annexin V) markers. Right panel: percentage of apoptosis determined from quantification of staining results. Figure 5F shows a competition assay between populations of MOLM14 Control-Blue and MOLM14 GAS6-AS2-Red following 25 μ M Ara-C treatment. Left panels: representative flow cytometry plots. Right panel: ratios between red and blue cells over time; in this figure, the GAS6-AS2 data lines (in red) are above the control data lines (in black). Figure 5G shows a schematic of an orthotopic xenograft competition assay between control (blue) and GAS6- AS2 (Red) MOLM14 cells with Ara-C treatment. Figure 5H shows the ratios of control (in blue and at top of columns) versus GAS6-AS2 (in red and base of columns) MOLM14 cells from bone marrow of mice treated and analyzed at day 17 as outlined in Figure 5G. Figure 5I shows a representative flow cytometry results of cells harvested from mouse bone marrow 17 days following transplantation, and treatment with vehicle or Ara-C for 5 days.

[0037] Figure 6A-H are a series of graphs that illustrates that GAS6-AS2 activates GAS6/TAM signaling. Figure 6A shows the Pearson correlation between GAS6-AS2 and GAS6 expression levels following GAS6-AS2 activation. Data are represented as mean of triplicate measurements. Figure 6B shows the Pearson correlation between GAS6-AS2 and GAS6 expression levels across the 760 cancer cell lines analyzed (Figure 1A-B). Figure 6C shows the Pearson correlation between GAS6-AS2 and GAS6 expression levels in AML patient samples. Figure 6D shows the western blot analysis of differential GAS6/TAM signaling activation in response to individual control or GAS6-AS2 sgRNA overexpression; in this figure, the GAS6-AS2 data (in red) are to the right of each data pair. Figure 6E shows the Pearson correlation between GAS6-AS2 and AXL expression levels in AML patient samples. Figure 6F shows the Pearson correlation between GAS6-AS2 and AXL expression levels across the 760 cancer cell lines analyzed (Figure 1A-B). Figure 6G shows the expression levels of GAS6-AS2, GAS6, and AXL in MOLM14 and K562 cell lines. Figure 6H shows the Ara-C efficacy measurements in MOLM14 and K562 cell lines, based on normalized MTS reads following 48 hours of treatment with the indicated concentrations of Ara-C; in this figure, the K562 data line (in black) is above the MOLM14 data line (in red).

[0038] Figure 7A-G are a series of graphs illustrating that GAS6-AS2 demonstrates trans-regulation of AXL. Figure 7A shows the Fold change (FC) of GAS6-AS2, GAS6, and AXL in

response to GAS6-AS2 knockdown *via* ASO in K562 cells. Figure 7B shows the Modulation of Ara-C response upon GAS6-AS2 knockdown *via* ASO in K562 cells. Figure 7C shows the Methylation of CpG islands in the HEK293T AXL promoter following modulation of GAS6-AS2 expression; in this figure, data to the left of each pair (in black) are from controls and data to the right of each pair (in red) is data from GAS6-AS2. Figure 7D shows the Gene ontology analysis of coding genes clustered with GAS6-AS2 as determined by k-means Clustering. Figure 7E shows the drug resistance-gene expression Pearson correlation values of DNA methyltransferases. Genes enriched beyond a Z-score threshold of ± 1.16 , i.e., DNMT1 and DNMT3A, are colored in red. *See also* Figure 1B. Figure 7F shows the distribution of FPKM-normalized transcript abundances associated with DNMT1 versus IgG. Figure 7G shows, without wishing to be bound by theory, a model summarizing the mechanism by which GAS6-AS2 regulates GAS6/TAM signaling.

[0039] Figure 8A-B show the identification of protein-coding and noncoding gene biomarkers correlated with differential Ara-C response. Figure 8A shows a summary of gene set enrichment analysis (GSEA) of protein-coding genes ranked by disease-free survival association strength using annotated KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways. Clinical and transcriptomic data from the TCGA-LAML patient cohort was used for this analysis. Disease-free survival association was quantified by the magnitude of the coefficient from a Cox proportional hazards model for each gene, with patient sex, age over 60, cytogenetic risk, and white blood cell count above 16 as covariates. Figure 8B shows representative KEGG pathways from GSEA of protein-coding genes ranked by disease-free survival association strength as shown in Figures 1B-1C. Clinical and transcriptomic data from the TCGA-LAML patient cohort was used for this analysis.

[0040] Figure 9 is a series of graphs showing the Fold change (FC) of expression levels modulated by CRISPRa for a cohort of single sgRNAs representing various protein-coding and lncRNA genes modulated in HEK293T (left-most, in black), MOLM14 (second to left, in red), K562 (second to right, in blue), or HL60 (right-most, in yellow) cells.

[0041] Figure 10A-C illustrate the CRISPRa functional Screening of noncoding genes modulating Ara-C Response. Figure 10A shows the fold change (FC) of expression levels modulated by CRISPRa for sgRNAs predicted to target the TUNA lncRNA in MOLM14 cells. Figure 10B

shows the fold change (FC) of expression levels by CRISPRa for sgRNAs predicted to target a previously annotated TSS (MIAT-01) of and an alternative, predicted TSS (MIAT-06) of the MIAT lncRNA gene in MOLM14 cells. *See* Figure 10C shows the chromosomal localizations and predicted transcriptional start sites of MIAT-01 and MIAT-06 transcript isoforms.

[0042] Figure 11 is a non-limiting schematic showing the integrated genome wide CRISPRa approach to functionalize lncRNAs in drug resistance.

DETAILED DESCRIPTION OF THE DISCLOSURE

[0043] The present disclosure is based, at least, in part, the discovery that “priori” cancer subtype regression model that incorporates both protein-coding and noncoding gene expression biomarkers are useful for predicting drug efficacy. The methods of the disclosure provide an improvement in prognostication and stratification of cancer patients with various drug treatments.

[0044] In one aspect, the disclosure provides methods for determining an RNA signature indicative of sensitivity or resistance to a candidate cancer therapy. Accordingly, the present methods find use, in various embodiments in directing a suitable treatment (*e.g.*, selection of a most beneficial candidate cancer therapy, informing increasing or decreasing dosing or frequency of administration, identifying a patient population that is most likely to respond to a candidate cancer therapy, and the like).

[0045] In various embodiments, the present methods can be used as a companion diagnostic to a candidate cancer therapy.

[0046] In one aspect, the disclosure provides a method for determining an RNA signature indicative of sensitivity or resistance to a candidate cancer therapy, comprising: identifying one or more coding RNAs whose level of expression correlates with sensitivity or resistance to a candidate cancer therapy in cells of a cancer subtype, and identifying one or more noncoding RNAs whose level of expression correlates with sensitivity or resistance to the candidate cancer therapy in cells of the cancer subtype, thereby determining an RNA signature that is indicative of sensitivity or resistance to the candidate cancer therapy; and determining the presence or absence of the RNA signature in a cell or tissue of the cancer subtype, and classifying the cell or tissue for sensitivity or resistance to the candidate cancer therapy.

[0047] In some embodiments, the present methods allow for a cell or tissue to be classified as

sensitive or resistant to a candidate cancer therapy, including those described herein, and this classification is used to determine a treatment.

[0048] In some embodiments, the classification comprises any one of diagnosis, prognosis, and response to treatment.

[0049] In some embodiments, the RNA signature-directed classification is predictive of a positive response to and/or benefit from a candidate cancer therapy, *e.g.*, chemotherapy. In such embodiments, such patients may receive treatment with the candidate cancer therapy. In such embodiments, such patients may receive an escalated dose of the candidate cancer therapy

[0050] In some embodiments, the RNA signature-directed classification is predictive of a non-responsiveness to and/or lack of benefit from a candidate cancer therapy, *e.g.*, chemotherapy. In such embodiments, such patients may receive an alternative candidate cancer therapy, *e.g.*, chemotherapy. Or, in such embodiments, such patients may receive palliative care.

[0051] In some embodiments, the RNA signature-directed classification is predictive of a positive response to and/or benefit from neoadjuvant and/or adjuvant chemotherapy. In such embodiments, such patients may receive treatment neoadjuvant and/or adjuvant chemotherapy.

[0052] In some embodiments, the RNA signature-directed classification is predictive of a non-responsiveness to and/or lack of benefit from neoadjuvant and/or adjuvant chemotherapy. In such embodiments, such patients may not receive treatment neoadjuvant and/or adjuvant chemotherapy.

[0053] In some embodiments, the RNA signature-directed classification comprises a high level of cancer aggressiveness, wherein the aggressiveness is characterizable by one or more of a high tumor grade, aggressive histological subtypes, low overall survival, high probability of metastasis, and the presence of a tumor marker indicative of aggressiveness.

[0054] In various embodiments, the methods of the invention can be used to determine whether or not chemotherapy is an appropriate method of treatment. If it is, methods of the invention can provide information useful in the design and/or optimization of chemotherapy regimens that are particularly safe and effective for specific individuals or groups of individuals. For example, the methods of the invention can be used to inform amounts and times suitable for the treatment of a disease and therefore, they can be used to minimize and/or avoid adverse events/side effects of, *e.g.*, chemotherapy drugs (such adverse events/side effects include, but are not limited to, early

and late-forming diarrhea, nausea, vomiting, anorexia, constipation, flatulence, leukopenia, anemia, neutropenia, asthenia, abdominal cramping, fever, pain, loss of body weight, dehydration, alopecia, dyspnea, insomnia, and dizziness).

[0055] In some embodiments, the present methods provide for a companion diagnostic to a candidate cancer therapy. For instance, in some embodiments, the present methods: identify patients who are most likely to benefit from a candidate cancer therapy and/or identify patients likely to be at increased risk for serious side effects as a result of treatment with a candidate cancer therapy; and/or monitor response to treatment with a candidate cancer therapy for the purpose of adjusting treatment to achieve improved safety or effectiveness.

[0056] In some embodiments, the present methods inform patient inclusion or exclusion from a clinical trial of a candidate cancer therapy. For instance, in some embodiments, the present methods provide an RNA signature indicative of a likelihood to respond to a candidate cancer therapy and therefore directs inclusion of subjects bearing such RNA signature in a clinical trial of the candidate cancer therapy. Conversely, in some embodiments, the present methods provide an RNA signature indicative of a likelihood to non-responsiveness or inadequate responsiveness to a candidate cancer therapy and therefore directs exclusion of subjects bearing such RNA signature from a clinical trial of the candidate cancer therapy.

[0057] In some embodiments, the present methods provide information about a patient response to a candidate cancer therapy, which can be assessed using any endpoint indicating a benefit to the patient, including, without limitation, (1) inhibition, to some extent, of tumor growth, including slowing down and complete growth arrest; (2) reduction in the number of tumor cells; (3) reduction in tumor size; (4) inhibition (reduction, slowing down or complete stopping) of tumor cell infiltration into adjacent peripheral organs and/or tissues; (5) inhibition (reduction, slowing down or complete stopping) of metastasis; (6) enhancement of anti-tumor immune response, which may, but, does not have to, result in the regression or rejection of the tumor; (7) relief, to some extent, of one or more symptoms associated with the tumor; (8) increase in the length of survival following treatment; and (9) decreased mortality at a given point of time following treatment.

[0058] In another aspect, the disclosure provides methods for treating cancer in a subject in need thereof. The method comprising determining an RNA signature indicative of sensitivity or

resistance to a candidate cancer therapy according to the method of any one of the above embodiments or aspect and administering to the subject a cancer therapy that the cancer cell or tissue is classified as sensitive to.

[0059] In some embodiments, the cancer therapy is a chemotherapy, e.g., cytarabine, paclitaxel, docetaxel, doxorubicin, platinum-based chemotherapeutics that is optionally cisplatin or carboplatin, mitomycin, mithramycin, methotrexate, 5-fluorouracil, vinorelbine, topotecan, irinotecan, bleomycin, actinomycin, topoisomerase I and II inhibitors, anthracycline, epirubicin, idarubicin, mitoxantrone, valrubicin, etoposide, teniposide, or rubitecan.

[0060] In some embodiments, the candidate cancer therapy is a chemotherapy combination, e.g., TFAC (paclitaxel, 5-fluorouracil, doxorubicin, and cyclophosphamide), EC (epirubicin and cyclophosphamide), or FEC (5-fluorouracil, epirubicin, and cyclophosphamide).

[0061] In yet another aspect, the disclosure provides a composition for use in the method of any of the above aspects or embodiments.

[0062] Any aspect or embodiment described herein can be combined with any other aspect or embodiment as disclosed herein.

Predictive Biomarkers

[0063] Advances in the ability to profile cellular transcriptomes has provided a wealth of high-throughput data delineating expression of coding and non-coding genes across hundreds of cancer cell lines, however, there remains a critical lack of integrated high-throughput functional characterization and validation of these data in the context of disease progression and response to therapy. In some embodiments, the methods of the disclosure provide an integrative and comprehensive CRISPR activation (CRISPRa) framework that complements publicly available databases to enable the discovery of functional human protein coding and long non-coding RNA (lncRNA) genes that contribute to mechanisms of chemotherapeutic resistance. Accordingly, in some embodiments, the methods of the disclosure provide a dual coding and non-coding Integrated CRISPRa Screening (DICaS) platform and applied this integrative approach to identify genetic units and pathways that promote resistance to cancer treatments.

[0064] Provided herein are methods for predicting drug efficacy using a priori cancer subtype regression model that incorporates both protein-coding and noncoding gene expression biomarkers. Predictive RNA biomarkers are curated using this approach from *in vitro* cell line

data available from large pharmacogenomics databases such as the Cancer Therapeutics Response Portal (CTRP) or the Genomics of Drug Sensitivity in Cancer (GDSC), spanning a number of different cancer subtypes. These RNA biomarkers are then used to build a predictive model of drug response in murine models of cancer and in cancer patients. In some embodiments, the resulting analysis serves as a prognostic tool for clinicians and patients to estimate the expected response to a given therapy and facilitates appropriate therapeutic stratification of cancer patients for improved treatment outcomes. In some embodiments, the integration of pre-existing drug response data sets enables inference of drug resistance mechanisms and therapeutic outcomes for a drug uncharacterized in these screens *via* complementary small-scale pharmacological screening.

[0065] In some embodiments, “a priori” adjustment of drug response data for different cancer subtypes, and the use of both protein-coding and noncoding RNA biomarkers provide a more comprehensive and accurate prediction of response to drug therapy. Both of these facets are important in improving prognostication and stratification of cancer patients with various drug treatments. In some embodiments, the methods of the disclosure provide an improvement in prognostication and stratification of cancer patients with various drug treatments.

[0066] In some embodiment, this computational method described herein, leads to the development of a diagnostic software for clinicians and patients for predicting therapeutic outcomes given a set of patient RNA biomarkers as obtained from mRNA-seq. Additionally, the methods of the disclosure, in some embodiments, provide a computational platform to facilitate the discovery of drug resistance mechanisms of both included and excluded drugs in the analysis for downstream therapeutic modulation. In various embodiments, drug response data from, *e.g.*, the Cancer Therapeutics and Response Portal (CTRP) and/or mRNA seq data from the Cancer Cell Line Encyclopedia (CCLE) are interrogated in the present methods. In some embodiments, the present methods test a panel of small molecules for which drug response data are available. In some embodiments, the method is used to test the predictive capacity of the algorithm using available data sets and to test the efficacy of the approach in predicting drug response and resistance mechanisms for new drugs not included in existing data sets.

[0067] There are a large number of computational studies which focus on the predictive capacity of protein-coding genes and a smaller number which focus on noncoding genes. These studies

commonly use either a panel of cell lines or patient samples classified by a single cancer subtype; alternatively, analyses incorporating multiple cancer subtypes often stratify their analyses by cancer subtype and other similar features such as hematopoietic versus non-hematopoietic cell lines, incorporate cancer subtype as an additional predictive feature in addition to gene expression levels, or ignore this contribution entirely. Thus, the computational methods described herein are superior and more clinically relevant by removing the effect of cancer/tissue subtype prior to predictive modeling, thereby focusing the analysis on genetic features not associated with differences between cancer subtypes in the analyzed data sets.

[0068] In some embodiments, functional screening is carried out with CRISPRa based technologies using one or more of an established protein coding sgRNA library and a new genome-wide non-coding sgRNA (CaLR) library. In some embodiments, the methods of the disclosure are efficacious in predicting lncRNAs that facilitate resistance to cancer treatment. This demonstrates that many lncRNA genes are functionally relevant for cancer and modulate distinct cellular programs. In some embodiments, Guilt-by-association co-expression analysis is used to highlight cancer associations.

[0069] In one aspect, the disclosure provides methods for determining an RNA signature indicative of sensitivity or resistance to a candidate cancer therapy.

[0070] In one aspect, the disclosure provides a method for determining an RNA signature indicative of sensitivity or resistance to a candidate cancer therapy, comprising: identifying one or more coding RNAs whose level of expression correlates with sensitivity or resistance to a candidate cancer therapy in cells of a cancer subtype, and identifying one or more noncoding RNAs whose level of expression correlates with sensitivity or resistance to the candidate cancer therapy in cells of the cancer subtype, thereby determining an RNA signature that is indicative of sensitivity or resistance to the candidate cancer therapy; and determining the presence or absence of the RNA signature in a cell or tissue of the cancer subtype, and classifying the cell or tissue for sensitivity or resistance to the candidate cancer therapy.

[0071] In some embodiments, the method identifies an RNA signature that correlates with sensitivity or resistance. In some embodiments, the RNA signature comprises a coding and/or a non-coding RNA. In some embodiments the method identifies a coding RNA that correlates with sensitivity or resistance. In some embodiments, the methods identify gene pairs correlated with

Ara-C resistance. In some embodiment, the gene pairs correlated with Ara-C resistance are found in the list corresponding to Table 1. In some embodiments, the methods identify gene pairs correlated with Ara-C sensitivity. In some embodiments, the gene pairs correlated with Ara-C sensitivity are found in the list corresponding to Table 2. In some embodiments, the method identifies a coding RNA that correlate with sensitivity or resistance to Cytarabine (1- β -d-arabinofuranosylcytosine, Ara-C). In some embodiments, the gene comprises deoxycytidine kinase (DCK), equilibrative nucleoside transporter 1 (ENT1, SLC29A1), cytidine deaminase (CDA) and SAM Domain and HD Domain 1 (SAMHD1). In some embodiments, low expression of deoxycytidine kinase (DCK) and equilibrative nucleoside transporter 1 (ENT1, SLC29A1) may be correlated with increased resistance to Ara-C. In some embodiments, high expression of cytidine deaminase (CDA) and SAM Domain and HD Domain 1 (SAMHD1) correlated with increased resistance to Ara-C.

[0072] In some embodiments, the non-coding RNAs are long non-coding RNAs.

[0073] In some embodiments, an RNA signature includes the expression levels of one or more coding and non-coding RNA gene pairs, and optionally at least 5 or at least 10 coding and non-coding gene pairs.

[0074] In some embodiments, the RNA signature includes the expression levels for at least about 10 RNAs, or at least about 15 RNAs, or at least about 25 RNAs, or at least about 50 RNAs.

[0075] In some embodiments, the RNA signature includes the expression levels for at least about 10 RNAs, or at least about 15 RNAs, or at least about 20 RNAs, or at least about 25 RNAs, or at least about 30 RNAs, or at least about 35 RNAs, or at least about 40 RNAs, or at least about 45 RNAs, or at least about 50 RNAs, or at least about 55 RNAs, or at least about 60 RNAs, or at least about 65 RNAs, or at least about 70 RNAs, or at least about 75 RNAs, or at least about 80 RNAs, or at least about 85 RNAs, or at least about 90 RNAs, or at least about 95 RNAs, or at least about 100 RNAs, or at least about 105 RNAs, or at least about 110 RNAs, or at least about 115 RNAs, or at least about 120 RNAs, or at least about 125 RNAs, or at least about 130 RNAs, or at least about 135 RNAs, or at least about 140 RNAs, or at least about 145 RNAs, or at least about 150 RNAs, or at least about 155 RNAs, or at least about 160 RNAs, or at least about 165 RNAs, or at least about 170 RNAs, or at least about 175 RNAs, or at least about 180 RNAs, or at least about 185 RNAs, or at least about 190 RNAs, or at least about 195 RNAs, or at least about

200 RNAs, or at least about 225 RNAs, or at least about 250 RNAs, or at least about 275 RNAs, or at least about 300 RNAs, or at least about 325 RNAs, or at least about 350 RNAs, or at least about 375 RNAs, or at least about 400 RNAs, or at least about 425 RNAs, or at least about 450 RNAs, or at least about 475 RNAs, or at least about 500 RNAs, or at least about 525 RNAs, or at least about 550 RNAs, or at least about 575 RNAs, or at least about 600 RNAs, or at least about 625 RNAs, or at least about 650 RNAs, or at least about 675 RNAs, or at least about 700 RNAs, or at least about 725 RNAs, or at least about 750 RNAs, or at least about 775 RNAs, or at least about 800 RNAs, or at least about 825 RNAs, or at least about 850 RNAs, or at least about 875 RNAs, or at least about 900 RNAs, or at least about 925 RNAs, or at least about 950 RNAs, or at least about 975 RNAs, or at least about 1000 RNAs.

[0076] In some embodiments, the RNA signature includes the expression levels for 10 RNAs.

[0077] In some embodiments, the RNA signature includes the expression levels for not more than 10 RNA, or not more than 15 RNAs, or not more than 25 RNAs, or not more than 50 RNAs.

[0078] In some embodiments, the RNA signature includes the expression levels for not more than 10 RNA or not more than 15 RNAs, or not more than 20 RNAs, or not more than 25 RNAs, or not more than 30 RNAs, or not more than 35 RNAs, or not more than 40 RNAs, or not more than 45 RNAs, or not more than 50 RNAs, or not more than 55 RNAs, or not more than 60 RNAs, or not more than 65 RNAs, or not more than 70 RNAs, or not more than 75 RNAs, or not more than 80 RNAs, or not more than 85 RNAs, or not more than 90 RNAs, or not more than 95 RNAs, or not more than 100 RNAs, or not more than 105 RNAs, or not more than 110 RNAs, or not more than 115 RNAs, or not more than 120 RNAs, or not more than 125 RNAs, or not more than 130 RNAs, or not more than 135 RNAs, or not more than 140 RNAs, or not more than 145 RNAs, or not more than 150 RNAs, or not more than 155 RNAs, or not more than 160 RNAs, or not more than 165 RNAs, or not more than 170 RNAs, or not more than 175 RNAs, or not more than 180 RNAs, or not more than 185 RNAs, or not more than 190 RNAs, or not more than 195 RNAs, or not more than RNAs, or not more than 225 RNAs, or not more than 250 RNAs, or not more than 275 RNAs, or not more than 300 RNAs, or not more than 325 RNAs, or not more than 350 RNAs, or not more than 375 RNAs, or not more than 400 RNAs, or not more than 425 RNAs, or not more than 450 RNAs, or not more than 475 RNAs, or not more than 500 RNAs, or not more than 525 RNAs, or not more than 550 RNAs, or not more than 575 RNAs, or not more

than 600 RNAs, or not more than 625 RNAs, or not more than 650 RNAs, or not more than 675 RNAs, or not more than 700 RNAs, or not more than 725 RNAs, or not more than 750 RNAs, or not more than 775 RNAs, or not more than 800 RNAs, or not more than 825 RNAs, or not more than 850 RNAs, or not more than 875 RNAs, or not more than 900 RNAs, or not more than 925 RNAs, or not more than 950 RNAs, or not more than 975 RNAs, or not more than 1000 RNAs.

[0079] As used herein “gene signature” or “gene expression signature” is a single or combined group of genes in a cell with a uniquely characteristic pattern of gene expression that occurs as a result of an altered or unaltered biological process or pathogenic medical condition. Identification of pathway and gene expression signatures indicative of response and/or resistance to chemotherapeutic agents provides the ability to improve therapeutic efficacy by gene-expression analysis of patient tumors and/or malignant cells. The gene expression profile generally contains the expression levels for a sufficient number of genes to perform pathway analysis or evaluate for the presence of a gene expression signature as described herein. For example, the gene expression profile may contain the expression levels for at least about 10, 25, 50, 100, 500, 1000 genes or more, with these genes being associated with the enriched pathways disclosed herein. The profile may comprise the expression level of at least 10, 20, 30, 40, or 50 genes listed in any one of Tables 1-2. Where a significant number of genes associated with a pathway are differentially expressed, the pathway is deemed an “enriched pathway.”

[0080] In some embodiments, the methods of the disclosure provide determining a prognosis of a subject having a proliferative disorder, for example, cancer (*e.g.*, acute myeloid leukemia (AML)).

[0081] Alternatively, prognosis may be based on, in addition to Ara-C, a gene expression signature of the cancer that is indicative of chemotherapy resistance, likelihood of cancer recurrence, or a high risk group for survival. Gene expression signatures are becoming increasingly available for predicting tumor response to therapy and/or other classification of tumors for prognosis.

[0082] The pathway and gene expression signatures (*e.g.*, data for pathway analysis) may be in a format consistent with any nucleic acid detection format, such as those described herein, and will generally be comparable to the format used for profiling patient samples. For example, the gene expression signatures and patient profiles may both be prepared by nucleic acid hybridization

method, and with the same hybridization platform and controls so as to facilitate comparisons. The gene expression signatures may further embody any number of statistical measures, including Mean or median expression levels and/or cut-off or threshold values.

[0083] Once the gene expression profile for patient samples are prepared, the profile is evaluated for the presence of one or more of the pathway or gene expression signatures, by scoring or classifying the patient profile against each pathway or gene expression signature. Illustrative pathway signatures for sensitivity or resistance to TFAC, EC, and FEC are disclosed herein in Figure 3. Illustrative gene expression signatures, derived from the identified enriched pathways, are disclosed in Tables 1-2.

[0084] Table 1: List of protein-coding/cognate antisense gene pairs identified to be highly correlated in expression levels with each other and have significant drug resistance-gene expression correlations. (Gene Pairs Correlated with Ara-C Resistance).

ACOXL	ACOXL-AS1	INHBA	INHBA-AS1	P4HA3	P4HA3-AS1
B4GALT1	B4GALT1-AS1	IQCJ-SCHIP1	IQCJ-SCHIP1-AS1	PIK3IP1	PIK3IP1-AS1
BDNF	BDNF-AS	ISPD	ISPD-AS1	PINK1	PINK1-AS
CCDC183	CCDC183-AS1	ITGB5	ITGB5-AS1	PRICKLE2	PRICKLE2-AS2
CDKN2A	CDKN2A-AS1	KIRREL3	KIRREL3-AS2	PRICKLE2	PRICKLE2-AS3
COL5A1	COL5A1-AS1	KRT7	KRT7-AS	PRKG1	PRKG1-AS1
CPEB2	CPEB2-AS1	LACTB2	LACTB2-AS1	RBPMS	RBPMS-AS1
CXXC5	CXXC5-AS1	LIFR	LIFR-AS1	RERG	RERG-AS1
DOCK9	DOCK9-AS2	LOXL1	LOXL1-AS1	RMDN2	RMDN2-AS1
EGFR	EGFR-AS1	LRP1	LRP1-AS	ROR1	ROR1-AS1
EXOC3	EXOC3-AS1	MAMDC2	MAMDC2-AS1	SEMA5A	SEMA5A-AS1
FAM83A	FAM83A-AS1	MKX	MKX-AS1	SLC2A1	SLC2A1-AS1
FLNC	FLNC-AS1	MRVI1	MRVI1-AS1	SPON1	SPON1-AS1
FRMD6	FRMD6-AS1	NALCN	NALCN-AS1	TGFB2	TGFB2-AS1
FRMD6	FRMD6-AS2	NDST1	NDST1-AS1	THSD4	THSD4-AS1
FTCD	FTCD-AS1	NEURL1	NEURL1-AS1	TM4SF1	TM4SF1-AS1
FUT8	FUT8-AS1	NEXN	NEXN-AS1	TM4SF19	TM4SF19-AS1
GAS6	GAS6-AS2	NTRK3	NTRK3-AS1	VLDLR	VLDLR-AS1
GRM5	GRM5-AS1	OSBPL10	OSBPL10-AS1	ZNF667	ZNF667-AS1
GSN	GSN-AS1	OSMR	OSMR-AS1	ZNF793	ZNF793-AS1
IGFBP7	IGFBP7-AS1	P3H2	P3H2-AS1		

[0085] Table 2: List of protein-coding/cognate antisense gene pairs identified to be highly correlated in expression levels with each other and have significant drug sensitivity-gene expression correlations. (Gene Pairs Correlated with Ara-C Sensitivity)

ADARB2	ADARB2-AS1	MYB	MYB-AS1
AP4B1	AP4B1-AS1	MYCBP2	MYCBP2-AS1

CCND2	CCND2-AS2	NOP53	NOP53-AS1
CCNT2	CCNT2-AS1	PABPC5	PABPC5-AS1
DCUN1D2	DCUN1D2-AS	PAXBP1	PAXBP1-AS1
DLGAP1	DLGAP1-AS4	PCCA	PCCA-AS1
DLGAP1	DLGAP1-AS5	PLCH1	PLCH1-AS1
ELFN1	ELFN1-AS1	PRC1	PRC1-AS1
ENTPD1	ENTPD1-AS1	RNF219	RNF219-AS1
FAM53B	FAM53B-AS1	SLC25A30	SLC25A30-AS1
FBXW7	FBXW7-AS1	SLCO4A1	SLCO4A1-AS1
FOXD2	FOXD2-AS1	SMAD1	SMAD1-AS2
GABPB1	GABPB1-AS1	SMC5	SMC5-AS1
GTF3C2	GTF3C2-AS1	SSBP3	SSBP3-AS1
HHATL	HHATL-AS1	THAP9	THAP9-AS1
INTS6	INTS6-AS1	TMPO	TMPO-AS1
INTS9	INTS9-AS1	USP3	USP3-AS1
KCNIP2	KCNIP2-AS1	ZBED3	ZBED3-AS1
LANCL1	LANCL1-AS1		
MAPKAPK5	MAPKAPK5-AS1		
MRGPRG	MRGPRG-AS1		

[0086] In some embodiments, the methods of the disclosure provide determining a prognosis of a subject having a proliferative disorder, for example, cancer (*e.g.*, Acute Lymphocytic Leukemia (ALL)).

[0087] As used herein, the term “cell proliferative disorder” refers to conditions in which unregulated or abnormal growth, or both, of cells can lead to the development of an unwanted condition or disease, which may or may not be cancerous. Illustrative cell proliferative disorders encompass a variety of conditions wherein cell division is deregulated.

[0088] Illustrative cell proliferative disorder includes, but are not limited to, neoplasms, benign tumors, malignant tumors, pre-cancerous conditions, *in situ* tumors, encapsulated tumors, metastatic tumors, liquid tumors, solid tumors, immunological tumors, hematological tumors, cancers, carcinomas, leukemias, lymphomas, sarcomas, and rapidly dividing cells. The term “rapidly dividing cell” as used herein is defined as any cell that divides at a rate that exceeds or is greater than what is expected or observed among neighboring or juxtaposed cells within the same tissue.

[0089] The present disclosure relates the treatment or prevention of cancers and/or tumors.

[0090] Cancers or tumors refer to an uncontrolled growth of cells and/or abnormal increased cell survival and/or inhibition of apoptosis which interferes with the normal functioning of the bodily organs and systems. Included are benign and malignant cancers, polyps, hyperplasia, as well as

dormant tumors or micrometastases. Also, included are cells having abnormal proliferation that is not impeded by the immune system (*e.g.*, virus infected cells). The cancer may be a primary cancer or a metastatic cancer. The primary cancer may be an area of cancer cells at an originating site that becomes clinically detectable, and may be a primary tumor. In contrast, the metastatic cancer may be the spread of a disease from one organ or part to another non-adjacent organ or part. The metastatic cancer may be caused by a cancer cell that acquires the ability to penetrate and infiltrate surrounding normal tissues in a local area, forming a new tumor, which may be a local metastasis. The cancer may also be caused by a cancer cell that acquires the ability to penetrate the walls of lymphatic and/or blood vessels, after which the cancer cell is able to circulate through the bloodstream (thereby being a circulating tumor cell) to other sites and tissues in the body. The cancer may be due to a process such as lymphatic or hematogeneous spread. The cancer may also be caused by a tumor cell that comes to rest at another site, re-penetrates through the vessel or walls, continues to multiply, and eventually forms another clinically detectable tumor. The cancer may be this new tumor, which may be a metastatic (or secondary) tumor.

[0091] The cancer may be caused by tumor cells that have metastasized, which may be a secondary or metastatic tumor. The cells of the tumor may be like those in the original tumor. As an example, if a breast cancer or colon cancer metastasizes to the liver, the secondary tumor, while present in the liver, is made up of abnormal breast or colon cells, not of abnormal liver cells. The tumor in the liver may thus be a metastatic breast cancer or a metastatic colon cancer, not liver cancer.

[0092] The cancer may have an origin from any tissue. The cancer may originate from melanoma, colon, breast, or prostate, and thus may be made up of cells that were originally skin, colon, breast, or prostate, respectively. The cancer may also be a hematological malignancy, which may be leukemia or lymphoma. The cancer may invade a tissue such as liver, lung, bladder, or intestinal.

[0093] Representative cancers and/or tumors of the present invention include, but are not limited to, a basal cell carcinoma, biliary tract cancer; bladder cancer; bone cancer; brain and central nervous system cancer; breast cancer; cancer of the peritoneum; cervical cancer; choriocarcinoma; colon and rectum cancer; connective tissue cancer; cancer of the digestive

system; endometrial cancer; esophageal cancer; eye cancer; cancer of the head and neck; gastric cancer (including gastrointestinal cancer); glioblastoma; hepatic carcinoma; hepatoma; intra-epithelial neoplasm; kidney or renal cancer; larynx cancer; leukemia; liver cancer; lung cancer (*e.g.*, small-cell lung cancer, non-small cell lung cancer, adenocarcinoma of the lung, and squamous carcinoma of the lung); melanoma; myeloma; neuroblastoma; oral cavity cancer (lip, tongue, mouth, and pharynx); ovarian cancer; pancreatic cancer; prostate cancer; retinoblastoma; rhabdomyosarcoma; rectal cancer; cancer of the respiratory system; salivary gland carcinoma; sarcoma; skin cancer; squamous cell cancer; stomach cancer; testicular cancer; thyroid cancer; uterine or endometrial cancer; cancer of the urinary system; vulval cancer; lymphoma including Hodgkin's and non-Hodgkin's lymphoma, as well as B-cell lymphoma (including low grade/follicular non-Hodgkin's lymphoma (NHL); small lymphocytic (SL) NHL; intermediate grade/follicular NHL; intermediate grade diffuse NHL; high grade immunoblastic NHL; high grade lymphoblastic NHL; high grade small non-cleaved cell NHL; bulky disease NHL; mantle cell lymphoma; AIDS-related lymphoma; and Waldenstrom's Macroglobulinemia; chronic lymphocytic leukemia (CLL); acute lymphoblastic leukemia (ALL); Hairy cell leukemia; chronic myeloblastic leukemia; as well as other carcinomas and sarcomas; and post-transplant lymphoproliferative disorder (PTLD), as well as abnormal vascular proliferation associated with phakomatoses, edema (such as that associated with brain tumors), and Meigs' syndrome.

[0094] In some embodiments, the cancer subtype is defined by tissue of origin and histological subtype.

[0095] In some embodiments, the cancer subtype is a solid tumor. In some embodiments, the cancer subtype is a hematological malignancy. In some embodiments, the cancer subtype is a sarcoma, which is optionally squamous cell carcinoma, fibrosarcoma, myosarcoma, osteogenic sarcoma, angiosarcoma, or endotheliosarcoma. In some embodiments, the cancer subtype is a carcinoma, which is optionally adenocarcinoma.

[0096] In some embodiments, the cancer subtype is Small Cell Lung Cancer (CLC), Non-Small Cell Lung Cancer (NSCLC), or mesothelioma.

[0097] In some embodiments, the cancer subtype is a brain cancer or glioblastoma.

[0098] In some embodiments, the cancer subtype is a breast cancer, lymphoma, prostate cancer, pancreatic cancer, liver cancer, kidney cancer, colon or colorectal cancer, ovarian cancer,

endometrial cancer, cervical cancer, testicular cancer, or melanoma.

[0099] In some embodiments, the cancer is a leukemia, which is optionally acute myeloid leukemia (AML), chronic myelogenous leukemia (CML), or acute lymphoblastic leukemia (ALL).

[00100] In some embodiments, the methods of the disclosure provide determining a prognosis of a subject having a proliferative disorder, for example, cancer (*e.g.*, acute myeloid leukemia (AML)). Such methods may include steps of isolating one or more lncRNA transcripts in a biological sample from the subject; measuring a test level of the one or more isolated lncRNA transcripts assigning the test level to a high expression level or a low expression level relative to a cutoff value (*e.g.*, a baseline, cutoff or threshold) level; and determining a prognosis for the subject having the cancer based on the test level relative to the cutoff value level. The prognosis may be, for example, a poor prognosis or a good prognosis, measured by a shortened survival or a prolonged survival, respectively. Further, the survival may be measured as an overall survival (OS), disease-free survival (DFS), or recurrence-free survival (RFS).

[00101] The cancer may be primary or recurrent, and may be of any type (as described above), stage (*e.g.*, Stage I, II, III, or IV or an equivalent of other staging system), and/or histology. The patient may be of any age, sex, performance status, and/or extent and duration of remission.

Coding and Non-coding RNAs

[00102] Elucidating mechanisms leading to therapy resistance has focused on protein coding genes, yet it is clear that cancer development and progression cannot be fully explained by the protein coding genome. Protein coding genes represent a small fraction of the transcribed and functional genome, however research and understanding related to the non-coding RNA (ncRNA) transcriptome has highlighted the importance of ncRNAs in biology. Functional validation of various ncRNA species highlights the fact that these RNAs may play important roles in the pathogenesis of diseases including cancer. One very large group of ncRNAs is represented by long non-coding RNAs (lncRNA). These RNAs can be either nuclear or cytoplasmic in localization and have been found to play roles in a diverse array of biological processes. Importantly, many lncRNAs with nuclear functions behave in a cis-acting manner to regulate the expression of neighboring genes. As a result, their study has been constrained by the

requirement for their expression to be driven from their endogenous loci, as they act in a proximal and localized manner. However, recent developments with CRISPR technologies now facilitate the modulation of gene expression directly from the endogenous promoter. Utilization of this approach has already been compellingly demonstrated using CRISPR interference (CRISPRi) to silence the expression of lncRNAs genome-wide.

Non-coding RNA

[00103] Non-coding RNAs make up the majority (98%) of the transcriptome, and several different classes of regulatory RNA with important functions are being discovered. Understanding the significance of this RNA world is one of the most important challenges facing biology today, and the non-coding RNAs within it represent a gold mine of potential new biomarkers and drug targets.

[00104] Long non-coding RNAs (lncRNAs) are a large and diverse class of transcribed RNA molecules with a length of more than 200 nucleotides that do not encode proteins (or lack > 100 amino acid open reading frame). lncRNAs are thought to encompass nearly 30,000 different transcripts in humans, hence lncRNA transcripts account for the major part of the non-coding transcriptome. lncRNAs can be transcribed as whole or partial natural antisense transcripts (NAT) to coding genes, or located between genes or within introns. Some lncRNAs originate from pseudogenes. lncRNAs may be classified into different subtypes (Antisense, Intergenic, Overlapping, Intronic, Bidirectional, and Processed) according to the position and direction of transcription in relation to other genes. Gene expression profiling and in situ hybridization studies have revealed that lncRNA expression is developmentally regulated, can be tissue- and cell-type specific, and can vary spatially, temporally, or in response to stimuli. Many lncRNAs are expressed in a more tissue-specific fashion and with greater variation between tissues compared to protein-coding genes

[00105] In general, a CRISPR system is characterized by elements that promote the formation of a CRISPR complex at the site of a target sequence (also referred to as a protospacer in the context of an endogenous CRISPR system). In the context of formation of a CRISPR complex, “target sequence” refers to a sequence to which a guide sequence is designed to target, *e.g.*, have complementarity, where hybridization between a target sequence and a guide sequence promotes the formation of a CRISPR complex. The section of the guide sequence through which

complementarity to the target sequence is important for cleavage activity is referred to herein as the seed sequence. A target sequence may comprise any polynucleotide, such as DNA or RNA polynucleotides and is comprised within a target locus of interest. In some embodiments, a target sequence is located in the nucleus or cytoplasm of a cell. The herein described invention encompasses novel effector proteins of Class 2 CRISPR-Cas systems, of which Cas9 is an illustrative effector protein and hence terms used in this application to describe novel effector proteins, may correlate to the terms used to describe the CRISPR-Cas9 system.

[00106] As used herein, a Cas protein or a CRISPR enzyme refers to any of the proteins presented in the new classification of CRISPR-Cas systems.

[00107] As used herein, the term “crRNA” or “guide RNA” or “single guide RNA” or “sgRNA” or “one or more nucleic acid components” of a Type V or Type VI CRISPR-Cas locus effector protein comprises any polynucleotide sequence having sufficient complementarity with a target nucleic acid sequence to hybridize with the target nucleic acid sequence and direct sequence-specific binding of a nucleic acid-targeting complex to the target nucleic acid sequence. In some embodiments, the degree of complementarity, when optimally aligned using a suitable alignment algorithm, is about or more than about 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97.5%, 99%, or more.

[00108] A guide sequence, and hence a nucleic acid-targeting guide RNA may be selected to target any target nucleic acid sequence. The target sequence may be DNA. The target sequence may be any RNA sequence. In some embodiments, the target sequence may be a sequence within a RNA molecule selected from the group consisting of messenger RNA (mRNA), pre-mRNA, ribosomal RNA (rRNA), transfer RNA (tRNA), micro-RNA (miRNA), small interfering RNA (siRNA), small nuclear RNA (snRNA), small nucleolar RNA (snoRNA), double stranded RNA (dsRNA), non-coding RNA (ncRNA), long non-coding RNA (lncRNA), and small cytoplasmic RNA (scRNA). In some preferred embodiments, the target sequence may be a sequence within a RNA molecule selected from the group consisting of mRNA, pre-mRNA, and rRNA. In some embodiments, the target sequence may be a sequence within a RNA molecule selected from the group consisting of ncRNA, and lncRNA. In some more preferred embodiments, the target sequence may be a sequence within an mRNA molecule or a pre-mRNA molecule.

[00109] In some embodiments, the non-coding RNAs are long non-coding RNAs.

[00110] In some embodiments, an RNA signature includes the expression levels of one or more coding and non-coding RNA gene pairs, and optionally at least 5 or at least 10 coding and non-coding gene pairs.

[00111] In some embodiments, the RNA signature includes the expression levels for at least about 10 RNAs, or at least about 15 RNAs, or at least about 25 RNAs, or at least about 50 RNAs.

[00112] In some embodiments, one or more coding RNAs are involved in DNA replication, Cell cycle, Pyrimidine metabolism, Homologous recombination, p53 signaling pathway, Base excision repair, Nucleotide excision repair, Mismatch repair, One carbon pool by folate, Non-homologous end-joining, Citrate cycle (TCA cycle), Apoptosis - multiple species, Cellular senescence, MAPK signaling pathway, PI3K-Akt signaling pathway, Jak-STAT signaling pathway, Hematopoietic cell lineage, Oxidative phosphorylation, Fatty acid degradation, Cytokine-cytokine receptor interaction, Ribosome, RNA transport, mRNA surveillance pathway, RNA degradation, Spliceosome, or Purine metabolism, or other relevant cellular processes identified by pathway analysis algorithms.

[00113] Illustrative sequences encoding validated sgRNAs are shown in Table 3.

Table 3: GAS6-AS2 sgRNAs

GAS6-AS2 sgRNAs	Target Sequence
#1	GCCGCCGCGATGTGACCTTC (SEQ ID NO: 1)
#2	CGCCGCCGCCGCCGCCGCCG (SEQ ID NO: 2)
#3	GGCGGCGGCGGCGGCGGCGG (SEQ ID NO: 3)
#4	GCACCTCAAGCGCTCGGTCT (SEQ ID NO: 4)
#5	ACGCCCAGACCGAGCGCTTG (SEQ ID NO: 5)
#6	GCGCGGGGAGGCGGCGGGGG (SEQ ID NO: 6)
#7	GCCCCGTCCGGGAGGGAGGT (SEQ ID NO: 7)
#8	TGACCCCCCACCTCCCTCC (SEQ ID NO: 8)

Table 4: Illustrative sequences encoding control sgRNAs are shown in Table 4.

List of sgRNAs for Control and Validations	
Control 1	GCAGCTCGACCTCAAGCCGT (SEQ ID NO: 9)
Control 2	GTACTCCAATCCGCGATGAC (SEQ ID NO: 10)
Control 3	GTCGGTGATCCAACGTATCG (SEQ ID NO: 11)
Control 4	GCGCCTTAAGAGTACTCATC (SEQ ID NO: 12)
Control 5	GTATGGTGAGTAGTCGCTTG (SEQ ID NO: 13)
MIAT-006	CTGTCCCACGACACCTTCTG (SEQ ID NO: 14)

MIAT-006	GTCAGGGAGAGGGCTCTATG (SEQ ID NO: 15)
MIAT-006	AACAATCTGCATCCAGATCC (SEQ ID NO: 16)
MIAT-006	AGATTGTTGAGACTGAAATT (SEQ ID NO: 17)
MIAT-001	CTCCCGCATTAATAATTCAT (SEQ ID NO: 18)
MIAT-001	ATGAAATTTTAATGCGGGAG (SEQ ID NO: 19)
MIAT-001	GCGCCCATGAAATTTTAATG (SEQ ID NO: 20)
MIAT-001	GCGCCCATGAAATTTTAATG (SEQ ID NO: 21)
MIAT-001	ACGACGCCGGCTGCGCTCGC (SEQ ID NO: 22)
TUNAR	GCGCGCTGCGAGCCGGGCTG (SEQ ID NO: 23)
TUNAR	GCGCGTCAGCGGGGGCTCCC (SEQ ID NO: 24)
TUNAR	CGTTACCTAGCGACTCGCTC (SEQ ID NO: 25)
TUNAR	TGGCTCCGACGACGCGGGG (SEQ ID NO: 26)
DCK	AAGAGGTTCCGGAGTCGGGT (SEQ ID NO: 27)
TUFT1	CAGAGAACTAACGATAGGCG (SEQ ID NO: 28)
ZBP1	GGCTGTGCCTGGAACGGGGA (SEQ ID NO: 29)
CYB561	CCACGCACCCCCGATCATGG (SEQ ID NO: 30)
MUL1	GCGCCGTGTTTCGAAGGGGA (SEQ ID NO: 31)
PI4K2A	TGACGGGGAGAAAAATGCGG (SEQ ID NO: 32)
PXDC1	GGCAACTACCCGGGGTCCGG (SEQ ID NO: 33)
AC008073.2-201	CCGTTTTCTGTGAAATCAAA (SEQ ID NO: 34)
AL359182.1-201	ACCGAGTGAAGACTCAGTGC (SEQ ID NO: 35)
AC106897.1-201	GCGCGCTGGGCGGGCGGCGG (SEQ ID NO: 36)
AC091982.2-201	AAATCGAGCTGCTTAATTCT (SEQ ID NO: 37)
LINC00923	TGCATGGTTTGATTAAAATA (SEQ ID NO: 38)
AC097359.2	GACTAACTCGTGACAGCCTC (SEQ ID NO: 39)
RMDN2-AS1	GCTAACAATTCTTCTTTATT (SEQ ID NO: 40)
AL353148.1	TTGATTTTCTCTCTGTTTTT (SEQ ID NO: 41)
AC012150.1	TTTGAATCTCTCTCCTGTCT (SEQ ID NO: 42)
AL157688.1	CTCATTTGTTTTCTCTCTCTC (SEQ ID NO: 43)
AC08477.5	ATTTCTAACTTTCTCCTCTT (SEQ ID NO: 44)
LINC02426	GATAATCTTATTTTCTCTCT (SEQ ID NO: 45)
TTN	CCTTGGTGAAGTCTCCTTTG (SEQ ID NO: 46)
MIAT	ATGCGGGAGGCTGAGCGCAC (SEQ ID NO: 47)
RHOXF2	ACGCGTGCTCTCCCTCATC (SEQ ID NO: 48)
ASCL1	CGGGAGAAAGGAACGGGAGG (SEQ ID NO: 49)

Table 5: Panel of validated sgRNAs targeting the promoters of both coding and noncoding genes.

	FORWORD	REVERSE
HUPO	GCTTCCTGGAGGGTGTCC (SEQ ID NO: 50)	GGACTCGTTTGTACCCGTTG (SEQ ID NO: 51)
GAS6	ATGTGAGCCACGACTTCTA (SEQ ID NO: 52)	TTCTACAGCCTGGACTACA (SEQ ID NO: 53)
AXL	GTTTGGAGCTGTGATGGAAGGC (SEQ ID NO: 54)	CGCTTCACTCAGGAAATCCTCC (SEQ ID NO: 55)
TYRO 3	GCAAGCCTTTGACAGTGTCTATGG (SEQ ID NO: 56)	GTTTCATCGCTGATGCCCCAAGCT (SEQ ID NO: 57)
MER	CAGGAAGATGGGACCTCTCTGA (SEQ ID NO: 58)	GGCTGAAGTCTTTCATGCACGC (SEQ ID NO: 59)
MIAT	TGGCTGGGGTTTGAACCTTT	AGGAAGCTGTTCCAGACTGC

	(SEQ ID NO: 60)	(SEQ ID NO: 61)
TTN	TGTTGCCACTGGTGCTAAAG (SEQ ID NO: 62)	ACAGCAGTCTTCTCCGCTTC (SEQ ID NO: 63)
RHOXF2	GGCAAGAAGCATGAATGTGA (SEQ ID NO: 64)	TGTCTCCTCCATTTGGCTCT (SEQ ID NO: 65)
ASCL1	CGCGGCCAACAAGAAGATG (SEQ ID NO: 66)	CGACGAGTAGGATGAGACCG (SEQ ID NO: 67)
HBG1	AGATGCCACAAAGCACCTG (SEQ ID NO: 68)	CTGCAGTCACCATCTTCTGC (SEQ ID NO: 69)
TUNA (1)	AGAACAAGGGGAAAGCTCG (SEQ ID NO: 70)	ATACCCACCCGCTTTTGAG (SEQ ID NO: 71)
TUNA (2)	TCTTGGAATGCTGGGGATT (SEQ ID NO: 72)	GCCATTACAGGGTGGTGTA (SEQ ID NO: 73)
AL353148.1	ACACCTGGCTGCTTTCTGA (SEQ ID NO: 74)	CCATCTTCTCACTGTGTCTCA (SEQ ID NO: 75)
AC012150.1	CTCTCCTAGCGGTGCTTTTG (SEQ ID NO: 76)	GATTTTTCATTTGCCAATCTCAT (SEQ ID NO: 77)
LINC02426	GCCCCAGAGAAAGACTGAGA (SEQ ID NO: 78)	AACCCAGTTGCTGGAAATACA (SEQ ID NO: 79)
AL157688.1	TGGATTCAAGCCTAAGTCACTG (SEQ ID NO: 80)	ATGAGCAATGAGCTGCAGAA (SEQ ID NO: 81)
AC087477.5	CCATGAACCTTCTGGAAGAGAAAC (SEQ ID NO: 82)	GCTTCTGCTCCCAAATGAAA (SEQ ID NO: 83)
AC005307.3	CAGATGAAGCCCAGAGATTTG (SEQ ID NO: 84)	TCCTAGAGCTTGAACAGTGTCT (SEQ ID NO: 85)
RMDN2-AS1	GGAAGTCAAGATGGCTGAC (SEQ ID NO: 86)	CCTGCCTTCTGTGTTGGTCT (SEQ ID NO: 87)
LINC02523	GCAGAAGGGTGGAGAGAGG (SEQ ID NO: 88)	TCAATTGGTCACCCCAGAAT (SEQ ID NO: 89)
GAS6-AS2	TGGTTCAACTTGAGCAGACG (SEQ ID NO: 90)	AGCAGAGATCACGCCACTG (SEQ ID NO: 91)
DCK	AAATATGAAAGTCTGGTTGAAAAGG (SEQ ID NO: 92)	AAAGCTGAAGTATCTGGAACCATT (SEQ ID NO: 93)
TUFT1	GTGCACCCAGAGGACCAG (SEQ ID NO: 94)	GCTGAGTGGCTGGACACC (SEQ ID NO: 95)
ZBP1	CAAAAGACGAACAGAGAAGTTTCA (SEQ ID NO: 96)	TGCTTACAGCCCGGAAATAG (SEQ ID NO: 97)
ZBP1	AGGAAAGTGACTTCTGGATTG (SEQ ID NO: 98)	GGAGCAGTCCATACTGAGA (SEQ ID NO: 99)
ZNF532	GCTTCAAGATCCTGGGTAGA (SEQ ID NO: 100)	CCAACACCCTTTAGTAGCTT (SEQ ID NO: 101)
PXDC1	CAGGCTTAATGAGGTGGAGAAG (SEQ ID NO: 102)	GGAGATCTTTCGAAGAAGGTGAG (SEQ ID NO: 103)
ZBP1	CAAAAGACGAACAGAGAAGTTTCA (SEQ ID NO: 104)	TGCTTACAGCCCGGAAATAG (SEQ ID NO: 105)
ZBP1	AGGAAAGTGACTTCTGGATTG (SEQ ID NO: 106)	GGAGCAGTCCATACTGAGA (SEQ ID NO: 107)
MUL1	TCCAGGAAAATGCGTGCCTTA (SEQ ID NO: 108)	ATTCCACACCATCTTGTGCTC (SEQ ID NO: 109)

MUL1	GTTTCATCTCTCGTCCTCCTTTG (SEQ ID NO: 110)	CACTGGACACTGCATCATACT (SEQ ID NO: 111)
PI4K2A	ATAGGCCCTGCCTCCAGT (SEQ ID NO: 112)	AGGGGGCAGCAACACTCT (SEQ ID NO: 113)
AC091982.2	GAGAGGGAGCATGGTGTCC (SEQ ID NO: 114)	TTAGAAGTCCTCACTTTCTGTGGTT (SEQ ID NO: 115)
AC008073.2	TCCGCCTTCCAGATTCAA (SEQ ID NO: 116)	GCATGGTGATGGAGTCCTGT (SEQ ID NO: 117)
AL359182.1	CCGAAACCGAGCACTTC (SEQ ID NO: 118)	AATCTCCAAGGCAGTAGGT (SEQ ID NO: 119)
AC106897.1	AACTCTTACCCGGGCTCATC (SEQ ID NO: 120)	TCAAATCCCCAAATGCTGA (SEQ ID NO: 121)
LINC00923	TGAAGCGCATTTATCTCCACT (SEQ ID NO: 122)	ATGGAGGCTCATCTGGTGT (SEQ ID NO: 123)

Gene Expression Assay Formats

[00114] Gene expression profiles, including patient gene expression profiles and the drug-sensitive and drug-resistant signatures as described herein, may be prepared according to any suitable method for measuring gene expression. That is, the profiles may be prepared using any quantitative or semi-quantitative method for determining RNA transcript levels in samples. Such methods include polymerase-based assays, such as RT-PCR, TaqMan™, hybridization-based assays, for example using DNA microarrays or other solid support (*e.g.*, Whole Genome DASL™ Assay, Illumina, Inc.), nucleic acid sequence based amplification (NASBA), flap endonuclease-based assays, as well as direct mRNA capture with branched DNA (QuantiGene™) or Hybrid Capture™ (Digene). The assay format, in addition to determining the gene expression profiles, will also allow for the control of, *inter alia*, intrinsic signal intensity variation between tests. Such controls may include, for example, controls for background signal intensity and/or sample processing, and/or other desirable controls for gene expression quantification across samples. For example, expression levels between samples may be controlled by testing for the expression level of one or more genes that are not associated with enriched pathways or differentially expressed between drug-sensitive and drug-resistant cells, or which are generally expressed at similar levels across the population. Such genes may include constitutively expressed genes, many of which are known in the art. Illustrative assay formats for determining gene expression levels, and thus for preparing gene expression profiles and drug-sensitive and drug-resistant signatures are described in this section.

[00115] In some embodiments, the level of expression of the coding and non-coding RNAs is

determined by a hybridization assay, RNA sequencing, or quantitative PCR. In some embodiments, the levels of expression of the coding and non-coding RNAs is determined in a biopsy sample from a subject, and the presence of the RNA signature determined.

[00116] In some embodiments, the levels of expression of the coding and non-coding RNAs in the biopsy sample is determined using quantitative PCR, RNA sequencing, or hybridization assay.

[00117] In some embodiments, the levels of expression of the coding and non-coding RNAs is determined in the cell lines or primary cell cultures using a hybridization assay or RNA sequencing, and the levels of expression of the coding and non-coding RNAs is determined in the biopsy sample using quantitative PCR.

[00118] In some embodiments, the primary cell lines include cell lines from healthy tissues, and cancer cell lines.

[00119] In some embodiments, the human tumor specimen is a biopsy and/or is any one of a fresh tissue sample, frozen tumor tissue specimen, cultured cells (*e.g.*, primary cultures from tumor specimens, circulating tumor cells), and a formalin-fixed paraffin-embedded tumor tissue specimen.

[00120] In some embodiments, the tumor specimen may be a biopsy sample, such as cultured cells. These cells may be processed using the usual cell culture techniques that are known in the art. These cells may be circulating tumor cells. In some embodiments, the primary cell cultures include primary cultures from tumor specimens and/or circulating tumor cells.

[00121] In some embodiments, the tumor specimen contains less than 100 mg of tissue, or in certain embodiments, contains about 50 mg of tissue or less. The tumor specimen (or biopsy) may contain from about 20 mg to about 50 mgs of tissue, such as about 35 mg of tissue.

[00122] The tissue may be obtained, for example, as one or more (*e.g.*, 1, 2, 3, 4, or 5) needle biopsies (*e.g.*, using a 14-gauge needle or other suitable size). In some embodiments, the biopsy is a fine-needle aspiration in which a long, thin needle is inserted into a suspicious area and a syringe is used to draw out fluid and cells for analysis. In some embodiments, the biopsy is a core needle biopsy in which a large needle with a cutting tip is used during core needle biopsy to draw a column of tissue out of a suspicious area. In some embodiments, the biopsy is a vacuum-assisted biopsy in which a suction device increases the amount of fluid and cells that is extracted

through the needle. In some embodiments, the biopsy is an image-guided biopsy in which a needle biopsy is combined with an imaging procedure, such as, for example, X ray, computerized tomography (CT), magnetic resonance imaging (MRI) or ultrasound. In other embodiments, the sample may be obtained *via* a device such as the MAMMOTOME® biopsy system, which is a laser guided, vacuum-assisted biopsy system for breast biopsy.

[00123] In some embodiments, the malignant cells are enriched or expanded in culture by forming a monolayer culture from tumor sample explants. For example, cohesive multicellular particulates (explants) are prepared from a patient's tissue sample (*e.g.*, a biopsy sample or surgical specimen) using mechanical fragmentation. This mechanical fragmentation of the explant may take place in a medium substantially free of enzymes that are capable of digesting the explant.

[00124] In some embodiment, the biopsy sample is from healthy tissues, malignant tissues, other diseases.

[00125] In some embodiments, the tumor specimen may be a biological fluid and that is indicative of a state of a cancer. In some embodiment, the biological fluids include is blood, serum, plasma, urine, saliva, mucus, tears, amniotic fluid, breast milk, sputum, cerebrospinal fluid, peritoneal fluid, pleural fluid, seminal fluid, a fraction thereof or a combination thereof.

[00126] In some embodiment, the biological fluids include fluids are peripheral blood, serum, plasma, ascites, urine, sputum, saliva, bronchoalveolar lavage fluid, cyst fluid, pleural fluid, peritoneal fluid, lymph, pus, lavage fluids from sinus cavities, bronchopulmonary aspirates, and bone marrow aspirates.

[00127] A gene expression profile is determined for the tumor tissue or cell sample, such as a tumor sample removed from the patient by surgery or biopsy. The tumor sample may be "fresh," in that it was removed from the patient within about five days of processing, and remains suitable or amenable to culture. In some embodiments, the tumor sample is not "fresh," in that the sample is not suitable or amenable to culture. Tumor samples are generally not fresh after from 3 to 7 days (*e.g.*, about five days) of removal from the patient. The sample may be frozen after removal from the patient, and preserved for later RNA isolation. The sample for RNA isolation may be a formalin-fixed paraffin-embedded (FFPE) tissue.

[00128] In determining a tumor's gene expression profile, or in determining a drug-sensitive or

drug-resistant profile in accordance with the invention, a hybridization-based assay may be employed. Nucleic acid hybridization involves contacting a probe and a target sample under conditions where the probe and its complementary target sequence (if present) in the sample can form stable hybrid duplexes through complementary base pairing. The nucleic acids that do not form hybrid duplexes may be washed away leaving the hybridized nucleic acids to be detected, typically through detection of an attached detectable label. It is generally recognized that nucleic acids may be denatured by increasing the temperature or decreasing the salt concentration of the buffer containing the nucleic acids. Under low stringency conditions (*e.g.*, low temperature and/or high salt) hybrid duplexes (*e.g.*, DNA: DNA, RNA: RNA, or RNA: DNA) will form even where the annealed sequences are not perfectly complementary. Thus, specificity of hybridization is reduced at lower stringency. Conversely, at higher stringency (*e.g.*, higher temperature or lower salt) successful hybridization tolerates fewer mismatches. One of skill in the art will appreciate that hybridization conditions may be selected to provide any degree of stringency.

[00129] In some embodiments, the levels of expression of the coding and non-coding RNAs in the biopsy sample is determined using quantitative PCR, RNA sequencing, or hybridization assay. In some embodiment, the levels of expression of the coding and non-coding RNAs is determined in the cell lines or primary cell cultures using a hybridization assay or RNA sequencing, and the levels of expression of the coding and non-coding RNAs is determined in the biopsy sample using quantitative PCR.

[00130] In some embodiments, the disclosure may employ reverse transcription polymerase chain reaction (RT-PCR), which is a sensitive method for the detection of mRNA, including low abundant mRNAs present in clinical samples. The application of fluorescence techniques to RT-PCR combined with suitable instrumentation has led to quantitative RT-PCR methods that combine amplification, detection and quantification in a closed system. Two commonly used quantitative RT-PCR techniques are the TaqMan RT-PCR assay (ABI, Foster City, USA) and the Lightcycler assay (Roche, USA).

[00131] In some embodiments, the preparation of patient gene expression profiles or the preparation of drug-sensitive and drug-resistant profiles comprises conducting real-time quantitative PCR (TaqMan) with sample- derived RNA and control RNA. Holland, *et al.*, PNAS

88:7276-7280 (1991) describe an assay known as a TaqMan assay. The 5' to 3' exonuclease activity of Taq polymerase is employed in a polymerase chain reaction product detection system to generate a specific detectable signal concomitantly with amplification. An oligonucleotide probe, non-extendable at the 3' end, labeled at the 5' end, and designed to hybridize within the target sequence, is introduced into the polymerase chain reaction assay. Annealing of the probe to one of the polymerase chain reaction product strands during the course of amplification generates a substrate suitable for exonuclease activity. During amplification, the 5' to 3' exonuclease activity of Taq polymerase degrades the probe into smaller fragments that can be differentiated from undegraded probe.

[00132] Numerous hybridization assay formats are known, and which may be used in accordance with the methods of the disclosure. In some embodiments, such hybridization-based formats include solution-based and solid support-based assay formats. Solid supports containing oligonucleotide probes designed to detect differentially expressed genes can be filters, polyvinyl chloride dishes, particles, beads, microparticles or silicon or glass based chips, etc. Any solid surface to which oligonucleotides can be bound, either directly or indirectly, either covalently or non-covalently, may be used. In some embodiments, bead-based assays and/or chip-based assays are used.

[00133] In some embodiments, hybridization is performed at low stringency, such as 6x SSPET at 37° C (0.005% Triton X-100), to ensure hybridization, and then subsequent washes are performed at higher stringency (*e.g.*, 1x SSPET at 37° C) to eliminate mismatched hybrid duplexes. Successive washes may be performed at increasingly higher stringency (*e.g.*, down to as low as 0.25x SSPET at 37° C to 50° C) until a desired level of hybridization specificity is obtained. Stringency can also be increased by addition of agents such as formamide. Hybridization specificity may be evaluated by comparison of hybridization to the test probes with hybridization to the various controls that may be present, as described below (*e.g.*, expression level control, normalization control, mismatch controls, etc.).

[00134] Probes based on the sequences of the genes described herein for preparing expression profiles may be prepared by any suitable method.

[00135] A “probe” is a nucleic acid capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base

pairing, usually through hydrogen bond formation. A probe may include natural (*e.g.*, A, G, U, C, or T) or modified bases (7-deazaguanosine, inosine, etc.), or locked nucleic acid (LNA). In addition, the nucleotide bases in probes may be joined by a linkage other than a phosphodiester bond, so long as the bond does not interfere with hybridization. Thus, probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages. The hybridization-based assay will be generally conducted under conditions in which the probe(s) will hybridize to their intended target subsequence, but with only insubstantial hybridization to other sequences or to other sequences, such that the difference may be identified. Such conditions are sometimes called “stringent conditions.” Stringent conditions are sequence-dependent and can vary under different circumstances. For example, longer probe sequences generally hybridize to perfectly complementary sequences (over less than fully complementary sequences) at higher temperatures. Generally, stringent conditions may be selected to be about 5° C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. Illustrative stringent conditions may include those in which the salt concentration is at least about 0.01 to 1.0 M Na⁺ ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30° C for short probes (*e.g.*, 10 to 50 nucleotides). Desired hybridization conditions may also be achieved with the addition of agents such as formamide or tetramethyl ammonium chloride (TMAC). The hybridization-based assay may employ expression level controls, for example, probes that hybridize specifically with constitutively expressed genes in the biological sample. Typically, expression level control probes have sequences complementary to subsequences of constitutively expressed “housekeeping genes” including, but not limited to the actin gene, the transferrin receptor gene, the GAPDH gene, and the like.

[00136] Various classification schemes are known for classifying samples between two or more classes or groups, and these include, without limitation: Principal Components Analysis, Na’ive Bayes, Support Vector Machines, Nearest Neighbors, Decision Trees, Logistic, Artificial Neural Networks, and Rule-based schemes. In addition, the predictions from multiple models can be combined to generate an overall prediction. For example, a “majority rules” prediction may be generated from the outputs of a Na’ive Bayes model, a Support Vector Machine model, and a Nearest Neighbor model.

[00137] In some embodiments, the cell or tissue is classified using one or more classification analysis that incorporates one or more machine learning algorithms.

[00138] In some embodiments, the cell or tissue is classified using one or more classification schemes selected from a correlation analysis, Principal Components Analysis, Naïve Bayes, Support Vector Machines, Nearest Neighbors, Decision Trees, Logistic, Artificial Neural Networks, and Rule-based schemes.

[00139] In some embodiments, a Pearson's correlation analysis followed by a receiver operating characteristic analysis is used to determine optimal thresholds.

[00140] Thus, a classification algorithm or "class predictor" may be constructed to classify samples. The process for preparing a suitable class predictor is reviewed in R. Simon, Diagnostic and prognostic prediction using gene expression profiles in high- dimensional microarray data, British Journal of Cancer (2003) 89, 1599-1604, which review is hereby incorporated by reference in its entirety. Generally, the gene expression profiles for patient specimens are scored or classified as drug-sensitive signatures or drug-resistant signatures using the pathway analysis or gene expression signatures, including with stratified or continuous intermediate classifications or scores reflective of drug resistance or sensitivity. Such signatures may be assembled from gene expression data disclosed herein, or prepared from independent data sets. The signatures may be stored in a database and correlated to patient tumor gene expression profiles in response to user inputs.

[00141] After comparing the patient's gene expression profile to the drug- sensitive and/or drug-resistant signature, the sample is classified as, or for example, given a probability of being, a drug-sensitive profile or a drug-resistant profile. The classification may be determined computationally based upon known methods as described above. The result of the computation may be displayed on a computer screen or presented in a tangible form, for example, as a probability (*e.g.*, from 0 to 100%) of the patient responding to a given treatment. The report will aid a physician in selecting a course of treatment for the cancer patient. For example, in certain embodiments of the invention, the patient's gene expression profile will be determined to be a drug- sensitive profile on the basis of a probability, and the patient will be subsequently treated with that drug or combination. In other embodiments, the patient's profile will be determined to be a drug-resistant profile, thereby allowing the physician to exclude one or more candidate

treatments for the patient, thereby sparing the patient the unnecessary toxicity.

[00142] The method according to this aspect may lend additional or alternative predictive value over standard methods, such as for example, gene expression tests known in the art, or chemoresponse testing.

[00143] The methods of the disclosure aid the prediction of an outcome of treatment. That is, the gene expression signatures are each predictive of an outcome upon treatment with a candidate agent or combination. The outcome may be quantified in a number of ways. For example, the outcome may be an objective response, a clinical response, or a pathological response to a candidate treatment. The outcome may be determined based upon the techniques for evaluating response to treatment of solid tumors as described in Therasse *et al.*, New Guidelines to Evaluate the Response to Treatment in Solid Tumors, J. of the National Cancer Institute 92(3):205-207 (2000), which is hereby incorporated by reference in its entirety. For example, the outcome may be survival (including overall survival or the duration of survival), progression-free interval, or survival after recurrence. The timing or duration of such events may be determined from about the time of diagnosis or from about the time treatment (*e.g.*, chemotherapy) is initiated. Alternatively, the outcome may be based upon a reduction in tumor size, tumor volume, or tumor metabolism, or based upon overall tumor burden, or based upon levels of serum markers especially where elevated in the disease state. The outcome in some embodiments may be characterized as a complete response, a partial response, stable disease, and progressive disease, as these terms are understood in the art.

[00144] In some embodiments, the gene signature is indicative of a pathological complete response upon treatment with a particular candidate agent or combination (as already described). A pathological complete response, *e.g.*, as determined by a pathologist following examination of tissue removed at the time of surgery, generally refers to an absence of histological evidence of invasive tumor cells in the surgical specimen.

Chemoresponse Assay

[00145] The methods of the present disclosure may further comprise conducting chemoresponse testing with a panel of chemotherapeutic agents on cultured cells from the cancer patient, to thereby add additional predictive value. That is, the presence of one or more pathway or gene expression signatures in tumor cells, and the *in vitro* chemoresponse results for the tumor

specimen, are used to predict an outcome of treatment (*e.g.*, survival, pCR, etc.). For example, where the gene expression profile and chemoresponse test both indicate that a tumor is sensitive or resistant to a particular treatment, the predictive value of the method may be particularly high.

[00146] In some embodiments, the sensitivity or resistance of the cell lines or primary cell cultures to the candidate cancer therapy is determined by a chemosensitivity assay. In some embodiments, the chemosensitivity assay is an IC₅₀ determining assay.

[00147] As used herein “IC₅₀” is the half maximal inhibitory concentration of a drug/small molecule for its effectiveness in inhibiting a specific biological or biochemical function. In some embodiments, the IC₅₀s is used to determine the ability of the chemotherapeutic agent to induce death of the cells/cultures.

[00148] In some embodiments, the methods further comprise conducting chemoresponse testing with a panel of chemotherapeutic agents on cultured cells from a cancer patient, to thereby add additional predictive value. That is, the presence of one or more indicative pathway signatures, and the *in vitro* chemoresponse results for the tumor specimen, are used to predict an outcome of treatment (*e.g.*, survival, pCR, etc.). For example, where the gene expression profile and chemoresponse test both indicate that a tumor is sensitive or resistant to a particular treatment, the predictive value of the method may be particularly high. Chemoresponse testing may be performed *via* the CHEMOFX test, as known in the art.

[00149] In some embodiments, the disclosure provides a method for identifying a pathway signature indicative of a cancer cell, such as an ALL cell line or cell line's sensitivity or resistance against a chemotherapeutic agent. The method comprises determining the level of sensitivity of a panel of ALL cancer cell lines for the chemotherapeutic agent *in vitro*, and evaluating the gene expression levels of ALL cancer cell lines to identify biochemical pathways associated with the level of sensitivity. In some embodiments, the panel of ALL cancer cell lines are immortalized cell lines, and may comprise the panel described herein or a subset thereof. In some embodiments, the panel of ALL cancer cell lines are derived from explants of patient tumor specimens as described herein (*e.g.*, *via* ChemoFx), and are useful for identifying a population response rate, or patient sub-population likely to respond to the drug candidate.

[00150] In some embodiments, cohesive multicellular particulates (explants) are prepared from a patient's tissue sample (*e.g.*, a biopsy sample or surgical specimen) using mechanical

fragmentation. This mechanical fragmentation of the explant may take place in a medium substantially free of enzymes that are capable of digesting the explant. Prior to the chemotherapy assay, the growth of the cells may be monitored, and data from periodic counting may be used to determine growth rates which may or may not be considered parallel to growth rates of the same cells *in vivo* in the patient. If growth rate cycles can be documented, for example, then dosing of certain active agents can be customized for the patient. Monolayer growth rate and/or cellular morphology may be monitored using, for example, a phase-contrast inverted microscope. Generally, the cells of the monolayer should be actively growing at the time the cells are suspended and plated for drug exposure. The epithelial character of the cells may be confirmed by any number of methods. Thus, the monolayers will generally be non-confluent monolayers at the time the cells are suspended for drug exposure.

[00151] A panel of active agents may then be screened using the cultured cells. Generally, the agents are tested against the cultured cells using plates such as microtiter plates. For the chemosensitivity assay, a reproducible number of cells is delivered to a plurality of wells on one or more plates, preferably with an even distribution of cells throughout the wells. For example, cell suspensions are generally formed from the monolayer cells before substantial phenotypic drift of the tumor cell population occurs. The cell suspensions may be, without limitation, about 4,000 to 12,000 cells/ml, or may be about 4,000 to 9,000 cells/ml, or about 7,000 to 9,000 cells/ml. The individual wells for chemoresponse testing are inoculated with the cell suspension, with each well or "segregated site" containing about 102 to 104 cells. The cells are generally cultured in the segregated sites for about 4 to about 30 hours prior to contact with an agent. Each test well is then contacted with at least one pharmaceutical agent, for example, an agent for which a gene expression signature is available. In some embodiments, such agents include TFAC (paclitaxel, 5-fluorouracil, doxorubicin, and cyclophosphamide), EC (epirubicin and cyclophosphamide), or FEC (5-fluorouracil, epirubicin, and cyclophosphamide). In some embodiments, the agents include the combination of cyclophosphamide, doxorubicin, fluorouracil, and paclitaxel ("TFAC"), the combination of cyclophosphamide and epirubicin ("EC"), or the combination of cyclophosphamide, epirubicin, fluorouracil ("TFEC").

[00152] In some embodiments, the agents include Folinic acid, fluorouracil and oxaliplatin (FOLFOX), Leucovorin Calcium (Folinic Acid), Fluorouracil, Irinotecan Hydrochloride,

(FOLFIRI), Irinotecan Hydrochloride (IFL), Folinic Acid (FL), and QUASAR. In some embodiments, the chemotherapy regimen comprises a CAF regimen, which comprises cyclophosphamide, doxorubicin hydrochloride (Adriamycin), and fluorouracil, which may be used with adjuvant. In some embodiments, the chemotherapy regimen is a Machover schedule. In some embodiments, the chemotherapy regimen comprises a CMF regimen, which comprises cyclophosphamide, methotrexate and/or 5 fluorouracil. In some embodiments, the chemotherapy regimen comprises a ECF regimen, which comprises Epirubicin, cisplatin and continuous 5-fluorouracil (5-FU) infusion. In some embodiments, the chemotherapy regimen comprises a FEC regimen, which comprises 5-fluorouracil, epirubicin. Cyclophosphamide.

[00153] In some embodiments, the candidate cancer therapy is a chemotherapy. In some embodiments, the chemotherapy comprises cytarabine, paclitaxel, docetaxel, doxorubicin, platinum-based chemotherapeutics that is optionally cisplatin or carboplatin, mitomycin, mithramycin, methotrexate, 5-fluorouracil, vinorelbine, topotecan, irinotecan, bleomycin, actinomycin, topoisomerase I and II inhibitors, anthracycline, epirubicin, idarubicin, mitoxantrone, valrubicin, etoposide, teniposide, and rubitecan.

[00154] In some embodiments, the candidate cancer therapy is a chemotherapy combination. In some embodiments, the combination is TFAC (paclitaxel, 5-fluorouracil, doxorubicin, and cyclophosphamide), EC (epirubicin and cyclophosphamide), or FEC (5-fluorouracil, epirubicin, and cyclophosphamide).

[00155] Examples of candidate cancer therapy chemotherapeutic agents include one or more of 5-FU (Fluorouracil), Abemaciclib, Abiraterone Acetate, Abitrexate (Methotrexate), Abraxane (Paclitaxel Albumin-stabilized Nanoparticle Formulation), ABVD, ABVE, ABVE-PC, AC, Acalabrutinib, AC-T, ADE, Adriamycin (Doxorubicin), Afatinib Dimaleate, Afinitor (Everolimus), Afinitor Difsperz (Everolimus), Akynzeo (Netupitant and Palonosetron), Aldara (Imiquimod), Aldesleukin, Alecensa (Alectinib), Alectinib, Alimta (PEMETREXED), Aliqopa (Copanlisib Hydrochloride), Alkeran (Melphalan), Aloxi (Palonosetron Hydrochloride), Alunbrig (Brigatinib), Ambochlorin (Chlorambucil), Amboclorin (Chlorambucil), Amifostine, Aminolevulinic Acid, Anastrozole, Aprepitant, Aredia (Pamidronate), Arimidex (Anastrozole), Aromasin (Exemestane), Arranon (Nelarabine), Arsenic Trioxide, Asparaginase Erwinia chrysanthemi, Axicabtagene Ciloleucel, Axitinib, Azacitidine, BEACOPP, Becenum

(Carmustine), Beleodaq (Belinostat), Belinostat, Bendamustine Hydrochloride, BEP, Bexarotene, Bicalutamide, BiCNU (Carmustine), Blenoxane (Bleomycin), Bortezomib, Bosulif (Bosutinib), Bosutinib, Brigatinib, BuMel, Busulfan, Busulfex (Busulfan)C, Cabazitaxel, Cabometyx (Cabozantinib), Cabozantinib-S-Malate, CAF, Calquence (Acalabrutinib), Camptosar (Irinotecan Hydrochloride), Capecitabine, CAPOX, Caprelsa (Vandetanib), Carac (Fluorouracil--Topical), Carboplatin, CARBOPLATIN-TAXOL, Carfilzomib, Carmubris (Carmustine), Carmustine, Casodex (Bicalutamide), CeeNU (Lomustine), CEM, Ceritinib, Cerubidine (Daunorubicin), Cervarix (Recombinant HPV Bivalent Vaccine), CEV, Chlorambucil, CHLORAMBUCIL-PREDNISONE, CHOP, Cisplatin, Cladribine, Clafen (Cyclophosphamide), Clofarabine, Clofarex (Clofarabine), Clolar (Clofarabine), CMF, Cobimetinib, Cometriq (Cabozantinib), Copanlisib Hydrochloride, COPDAC, COPP, COPP-ABV, Cosmegen (Dactinomycin), Cotellic (Cobimetinib), Crizotinib, CVP, Cyclophosphamide, Cyfos (Ifosfamide), Cytarabine, Cytarabine Liposome, Cytosar-U (Cytarabine), Cytosan (Cyclophosphamide), Cytosan (Cytosan), Dabrafenib, Dacarbazine, Dacogen (Decitabine), Dactinomycin, Dasatinib, Daunorubicin Hydrochloride, Daunorubicin Hydrochloride and Cytarabine Liposome, DaunoXome (Daunorubicin Lipid Complex), Decadron (Dexamethasone), Decitabine, Defibrotide Sodium, Defitelio (Defibrotide Sodium), Degarelix, Denileukin Diftitox, DepoCyt (Cytarabine Liposome), Dexamethasone, Dexamethasone Intensol (Dexamethasone), Dexpak Taperpak (Dexamethasone), Dexrazoxane Hydrochloride, Docefrez (Docetaxel), Docetaxel, Doxil (Doxorubicin Hydrochloride Liposome), Doxorubicin Hydrochloride, Doxorubicin Hydrochloride Liposome, Dox-SL (Doxorubicin Hydrochloride Liposome), Droxia (Hydroxyurea), DTIC (Decarbazine), DTIC-Dome (Dacarbazine), Efudex (Fluorouracil--Topical), Eligard (Leuprolide), Elitek (Rasburicase), Ellence (Ellence (epirubicin)), Eloxatin (Oxaliplatin), Elspar (Asparaginase), Eltrombopag Olamine, Emcyt (Estramustine), Emend (Aprepitant), Enasidenib Mesylate, Enzalutamide, Epirubicin Hydrochloride, EPOCH, Eribulin Mesylate, Erivedge (Vismodegib), Erlotinib Hydrochloride, Erwinaze (Asparaginase Erwinia chrysanthemi), Ethyol (Amifostine), Etopophos (Etoposide Phosphate), Etoposide, Etoposide Phosphate, Eulexin (Flutamide), Evacet (Doxorubicin Hydrochloride Liposome), Everolimus, Evista (Raloxifene Hydrochloride), Evomela (Melfalan Hydrochloride), Exemestane, Fareston (Toremifene), Farydak (Panobinostat), Faslodex (Fulvestrant), FEC, Femara (Letrozole),

Filgrastim, Firmagon (Degarelix), FloPred (Prednisolone), Fludara (Fludarabine), Fludarabine Phosphate, Fluoroplex (Fluorouracil), Fluorouracil, Flutamide, Folex (Methotrexate), Folex PFS (Methotrexate), FOLFIRI, FOLFIRINOX, FOLFOX, Folutyn (Pralatrexate), FUDR (FUDR (floxuridine)), FU-LV, Fulvestrant, Gardasil (Recombinant HPV Quadrivalent Vaccine), Gardasil 9 (Recombinant HPV Nonavalent Vaccine), Gefitinib, Gemcitabine Hydrochloride, GEMCITABINE-CISPLATIN, GEMCITABINE-OXALIPLATIN, Gemzar (Gemcitabine), Gilotrif (Afatinib Dimaleate), Gilotrif (Afatinib), Gleevec (Imatinib Mesylate), Gliadel (Carmustine), Glucarpidase, Goserelin Acetate, Halaven (Eribulin Mesylate), Hemangeol (Propranolol Hydrochloride), Hexalen (Altretamine), HPV Bivalent Vaccine, Recombinant, HPV Nonavalent Vaccine, Recombinant, HPV Quadrivalent Vaccine, Recombinant, Hycamtin (Topotecan Hydrochloride), Hycamtin (Topotecan), Hydrea (Hydroxyurea), Hydroxyurea, Hyper-CVAD, Ibrance (Palbociclib), Ibrutinib, ICE, Iclusig (Ponatinib), Idamycin PFS (Idarubicin), Idarubicin Hydrochloride, Idelalisib, Idhifa (Enasidenib), Ifex (Ifosfamide), Ifosfamide, Ifosfamidum (Ifosfamide), Imatinib Mesylate, Imbruvica (Ibrutinib), Imiquimod, Imlygic (Talimogene Laherparepvec), Inlyta (Axitinib), Iressa (Gefitinib), Irinotecan Hydrochloride, Irinotecan Hydrochloride Liposome, Istodax (Romidepsin), Ixabepilone, Ixazomib Citrate, Ixempra (Ixabepilone), Jakafi (Ruxolitinib Phosphate), Jakafi (Ruxolitinib), JEB, Jevtana (Cabazitaxel), Keoxifene (Raloxifene Hydrochloride), Kepivance (Palifermin), Kisqali (Ribociclib), Kyprolis (Carfilzomib), Lanreotide Acetate, Lanvima (Lenvatinib), Lapatinib Ditosylate, Lenalidomide, Lenvatinib Mesylate, Lenvima (Lenvatinib Mesylate), Letrozole, Leucovorin Calcium, Leukeran (Chlorambucil), Leukine (Sargramostim), Leuprolide Acetate, Leustatin (Cladribine), Levulan (Aminolevulinic Acid), Linfolizin (Chlorambucil), LipoDox (Doxorubicin Hydrochloride Liposome), Lomustine, Lonsurf (Trifluridine and Tipiracil), Lupron (Leuprolide), Lynparza (Olaparib), Lysodren (Mitotane), Marqibo (Vincristine Sulfate Liposome), Marqibo Kit (Vincristine Lipid Complex), Matulane (Procarbazine), Mechlorethamine Hydrochloride, Megace (Megestrol), Megestrol Acetate, Mekinist (Trametinib), Melphalan, Melphalan Hydrochloride, Mercaptopurine, Mesnex (Mesna), Metastron (Strontium-89 Chloride), Methazolastone (Temozolomide), Methotrexate, Methotrexate LPF (Methotrexate), Methylnaltrexone Bromide, Mexate (Methotrexate), Mexate-AQ (Methotrexate), Midostaurin, Mitomycin C, Mitoxantrone Hydrochloride, Mitozytrex

(Mitomycin C), MOPP, Mostarina (Prednimustine), Mozobil (Plerixafor), Mustargen (Mechlorethamine), Mutamycin (Mitomycin), Myleran (Busulfan), Mylosar (Azacitidine), Nanoparticle Paclitaxel (Paclitaxel Albumin-stabilized Nanoparticle Formulation), Navelbine (Vinorelbine), Nelarabine, Neosar (Cyclophosphamide), Neratinib Maleate, Nerlynx (Neratinib), Netupitant and Palonosetron Hydrochloride, Neulasta (filgrastim), Neulasta (pegfilgrastim), Neupogen (filgrastim), Nexavar (Sorafenib), Nilandron (Nilutamide), Nilotinib, Nilutamide, Ninlaro (Ixazomib), Nipent (Pentostatin), Niraparib Tosylate Monohydrate, Nolvadex (Tamoxifen), Novantrone (Mitoxantrone), Nplate (Romiplostim), Odomzo (Sonidegib), OEPA, OFF, Olaparib, Omacetaxine Mepesuccinate, Oncaspar (Pegaspargase), Oncovin (Vincristine), Ondansetron Hydrochloride, Onivyde (Irinotecan Hydrochloride Liposome), Ontak (Denileukin Diftitox), Onxol (Paclitaxel), OPPA, Orapred (Prednisolone), Osimertinib, Oxaliplatin, Paclitaxel, Paclitaxel Albumin-stabilized Nanoparticle Formulation, PAD, Palbociclib, Palifermin, Palonosetron Hydrochloride, Palonosetron Hydrochloride and Netupitant, Pamidronate Disodium, Panobinostat, Panretin (Alitretinoin), Paraplat (Carboplatin), Pazopanib Hydrochloride, PCV, PEB, Pediapred (Prednisolone), Pegaspargase, Pegfilgrastim, Pemetrexed Disodium, Platinol (Cisplatin), PlatinolAQ (Cisplatin), Plerixafor, Pomalyst (Pomalidomide), Ponatinib Hydrochloride, Pralatrexate, Prednisone, Procarbazine Hydrochloride, Proleukin (Aldesleukin), Promacta (Eltrombopag Olamine), Propranolol Hydrochloride, Purinethol (Mercaptopurine), Purixan (Mercaptopurine), Radium 223 Dichloride, Raloxifene Hydrochloride, Rasburicase, R-CHOP, R-CVP, Reclast (Zoledronic acid), Recombinant Human Papillomavirus (HPV) Bivalent Vaccine, Recombinant Human Papillomavirus (HPV) Nonavalent Vaccine, Recombinant Human Papillomavirus (HPV) Quadrivalent Vaccine, Regorafenib, Relistor (Methylnaltrexone Bromide), R-EPOCH, Revlimid (Lenalidomide), Rheumatrex (Methotrexate), Ribociclib, R-ICE, Rolapitant Hydrochloride, Romidepsin, Romiplostim, Rubex (Doxorubicin), Rubidomycin (Daunorubicin Hydrochloride), Rubraca (Rucaparib), Rucaparib Camsylate, Ruxolitinib Phosphate, Rydapt (Midostaurin), Sandostatin (Octreotide), Sandostatin LAR Depot (Octreotide), Sclerosol Intrapleural Aerosol (Talc), Soltamox (Tamoxifen), Somatuline Depot (Lanreotide Acetate), Sonidegib, Sorafenib Tosylate, Sprycel (Dasatinib), STANFORD V, Sterapred (Prednisone), Sterapred DS (Prednisone), Sterile Talc Powder (Talc), Steritalc (Talc), Sterecyst (Prednimustine), Stivarga (Regorafenib),

Sunitinib Malate, Supprelin LA (Histrelin), Sutent (Sunitinib Malate), Sutent (Sunitinib), Synribo (Omacetaxine Mepesuccinate), Tabloid (Thioguanine), TAC, Tafinlar (Dabrafenib), Tagrisso (Osimertinib), Talc, Talimogene Laherparepvec, Tamoxifen Citrate, Tarabine PFS (Cytarabine), Tarceva (Erlotinib), Targretin (Bexarotene), Tassigna (Decarbazine), Tassigna (Nilotinib), Taxol (Paclitaxel), Taxotere (Docetaxel), Temodar (Temozolomide), Temozolomide, Temsirolimus, Tepadina (Thiotepa), Thalidomide, Thalomid (Thalidomide), TheraCys BCG (BCG), Thioguanine, Thioplex (Thiotepa), Thiotepa, TICE BCG (BCG), Tisagenlecleucel, Tolak (Fluorouracil--Topical), Toposar (Etoposide), Topotecan Hydrochloride, Toremifene, Torisel (Temozolomide), Totect (Dexrazoxane Hydrochloride), TPF, Trabectedin, Trametinib, Treanda (Bendamustine hydrochloride), Trelstar (Triptorelin), Trexall (Methotrexate), Trifluridine and Tipiracil Hydrochloride, Trisenox (Arsenic trioxide), Tykerb (lapatinib), Uridine Triacetate, VAC, Valrubicin, Valstar (Valrubicin Intravesical), Valstar (Valrubicin), VAMP, Vandetanib, Vantas (Histrelin), Varubi (Rolapitant), VeIP, Velban (Vinblastine), Velcade (Bortezomib), Velsar (Vinblastine Sulfate), Vemurafenib, Venclexta (Venetoclax), Vepesid (Etoposide), Verzenio (Abemaciclib), Vesanoid (Tretinoin), Viadur (Leuprolide Acetate), Vidaza (Azacitidine), Vinblastine Sulfate, Vincasar PFS (Vincristine), Vincrex (Vincristine), Vincristine Sulfate, Vincristine Sulfate Liposome, Vinorelbine Tartrate, VIP, Vismodegib, Vistogard (Uridine Triacetate), Voraxaze (Glucarpidase), Vorinostat, Votrient (Pazopanib), Vumon (Teniposide), Vyxeos (Daunorubicin Hydrochloride and Cytarabine Liposome), W, Wellcovorin (Leucovorin Calcium), Wellcovorin IV (Leucovorin), Xalkori (Crizotinib), XELIRI, Xeloda (Capecitabine), XELOX, Xofigo (Radium 223 Dichloride), Xtandi (Enzalutamide), Yescarta (Axicabtagene Ciloleucel), Yondelis (Trabectedin), Zaltrap (Ziv-Aflibercept), Zanosar (Streptozocin), Zarxio (Filgrastim), Zejula (Niraparib), Zelboraf (Vemurafenib), Zinecard (Dexrazoxane Hydrochloride), Ziv-Aflibercept, Zofran (Ondansetron Hydrochloride), Zoladex (Goserelin), Zoledronic Acid, Zolinza (Vorinostat), Zometa (Zoledronic acid), Zortress (Everolimus), Zydelig (Idelalisib), Zykadia (Ceritinib), Zytiga (Abiraterone Acetate), and Zytiga (Abiraterone).

[00156] The efficacy of each agent in the panel is determined against the patient's cultured cells, by determining the viability of the cells (*e.g.*, number of viable cells). For example, at predetermined intervals before, simultaneously with, or beginning immediately after, contact

with each agent or combination, an automated cell imaging system may take images of the cells using one or more of visible light, UV light and fluorescent light. Alternatively, the cells may be imaged after about 25 to about 200 hours of contact with each treatment. The cells may be imaged once or multiple times, prior to or during contact with each treatment. Of course, any method for determining the viability of the cells may be used to assess the efficacy of each treatment *in vitro*.

[00157] In some embodiments, the *in vitro* efficacy grade for each agent in the panel may be determined. While any grading system may be employed (including continuous or stratified), in certain embodiments the grading system is stratified, having from 2 or 3, to 10 response levels, *e.g.*, about 3, 4, or 5 response levels. For example, when using three levels, the three grades may correspond to a responsive grade (*e.g.*, sensitive), an intermediate responsive grade, and a non-responsive grade (*e.g.*, resistant), as discussed more fully herein. In some embodiments, the patient's cells show a heterogeneous response across the panel of agents, making the selection of an agent particularly crucial for the patient's treatment. The output of the assay is a series of dose-response curves for tumor cell survivals under the pressure of a single or combination of drugs, with multiple dose settings each (*e.g.*, ten dose settings). To better quantify the assay results, the invention employs in some embodiments a scoring algorithm accommodating a dose-response curve. Specifically, the chemoresponse data are applied to an algorithm to quantify the chemoresponse assay results by determining an adjusted area under curve (aAUC).

[00158] Since a dose-response curve only reflects the cell survival pattern in the presence of a certain tested drug, assays for different drugs and/or different cell types have their own specific cell survival pattern. Thus, dose response curves that share the same aAUC value may represent different drug effects on cell survival. Additional information may therefore be incorporated into the scoring of the assay. In particular, a factor or variable for a particular drug or drug class (such as those drugs and drug classes described) and/or reference scores may be incorporated into the algorithm. For example, in certain embodiments, the invention quantifies and/or compares the *in vitro* sensitivity/resistance of cells to drugs having varying mechanisms of action, and thus, in some cases, different dose-response curve shapes. In these embodiments, the invention compares the sensitivity of the patient's cultured cells to a plurality of agents that show some effect on the patient's cells *in vitro* (*e.g.*, all score sensitive to some degree), so that the most effective agent

may be selected for therapy. In such embodiments, an aAUC can be calculated to take into account the shape of a dose response curve for any particular drug or drug class. The aAUC takes into account changes in cytotoxicity between dose points along a dose-response curve, and assigns weights relative to the degree of changes in cytotoxicity between dose points. For example, changes in cytotoxicity between dose points along a dose-response curve may be quantified by a local slope, and the local slopes weighted along the dose-response curve to emphasize cytotoxicity.

[00159] For example, aAUC may be calculated as follows.

[00160] Step 1: Calculate Cytotoxicity Index (CI) for each dose, where $CI = \text{Mean}_{\text{drug}} / \text{Mean}^{\text{control}}$.

[00161] Step 2: Calculate local slope (Sd) at each dose point, for example, as $Sd = (CI_d - CI_{d-1}) / \text{Unit of Dose}$, or $Sd = (CI_{d-i} - CI_d) / \text{Unit of Dose}$.

[00162] Step 3: Calculate a slope weight at each dose point, *e.g.*, $W_d = 1 - Sd$.

[00163] Step 4: Compute aAUC, where $\text{aAUC} = \sum W_d CI_d$, and where, $d = 1, 2, 10$; $\text{aAUC} \sim (0, 10)$; And at $d = 1$, then $CI_{d-1} = 1$. Equation 4 is the summary metric of a dose response curve and may be used for subsequent regression over reference outcomes.

[00164] Usually, the dose-response curves vary dramatically around middle doses, not in lower or higher dose ranges. Thus, the algorithm in some embodiments need only determine the aAUC for a middle dose range, such as for example (where from 8 to 12 doses are experimentally determined, *e.g.*, about 10 doses), the middle 4, 5, 6, or 8 doses are used to calculate aAUC. In this manner, a truncated dose-response curve might be more informative in outcome prediction by eliminating background noise.

[00165] The numerical aAUC value (*e.g.*, test value) may then be evaluated for its effect on the patient's cells. For example, a plurality of drugs may be tested, and aAUC determined as above for each, to determine whether the patient's cells have a sensitive response, intermediate response, or resistant response to each drug. In some embodiments, each drug is designated as, for example, sensitive, or resistant, or intermediate, by comparing the aAUC test value to one or more cut-off values for the particular drug (*e.g.*, representing sensitive, resistant, and/or intermediate aAUC scores for that drug). The cut-off values for any particular drug may be set or determined in a variety of ways, for example, by determining the distribution of a clinical outcome within a range of corresponding aAUC reference scores. That is, a number of patient

tumor specimens are tested for chemosensitivity/resistance (as described herein) to a particular drug prior to treatment, and aAUC quantified for each specimen. Then after clinical treatment with that drug, aAUC values that correspond to a clinical response (*e.g.*, sensitive) and the absence of significant clinical response (*e.g.*, resistant) are determined. Cut-off values may alternatively be determined from population response rates. For example, where a patient population is known to have a response rate of 30% for the tested drug, the cutoff values may be determined by assigning the top 30% of aAUC scores for that drug as sensitive. Further still, cut-off values may be determined by statistical measures.

[00166] In some embodiments, the aAUC scores may be adjusted for drug or drug class. For example, aAUC values for dose response curves may be regressed over a reference scoring algorithm adjusted for test drugs. The reference scoring algorithm may provide a categorical outcome, for example, sensitive (s), intermediate sensitive (i) and resistant (r), as already described. Logistic regression may be used to incorporate the different information, *e.g.*, three outcome categories, into the scoring algorithm. However, regression can be extended to other forms, such as linear or generalized linear regression, depending on reference outcomes. The regression model may be fitted as the following: $\text{Logit}(\text{Pref}) = \alpha + \beta (\text{aAUC}) + \gamma (\text{drugs})$, where γ is a covariate vector and the vector can be extended to clinical and genomic features. The score may be calculated as $\text{Score} = \beta (\text{aAUC}) + \gamma (\text{drugs})$. Since the score is a continuous variable, results may be classified into clinically relevant categories, *e.g.*, sensitive (S), intermediate sensitive (I), and resistant (R), based on the distribution of a reference scoring category or maximized sensitivity and specificity relative to the reference. As stated, the chemoresponse score for cultures derived from patient specimens may provide additional predictive or prognostic value in connection with the gene expression profile analysis.

[00167] Alternatively, where applied to immortalized cell line collections or patient-derived cultures, the *in vitro* chemoresponse assay may be used to supervise or train pathway and gene expression signatures. Once gene expression signatures are identified in cultured cells, *e.g.*, by correlating the level of *in vitro* chemosensitivity with gene expression levels, the resulting gene expression signatures may be independently validated in patient test populations having available gene expression data and corresponding clinical data, including information regarding the treatment regimen and outcome of treatment. This aspect of the invention reduces the length of

time and quantity of patient samples needed for identifying and validating such gene expression signatures.

Definitions

[00168] As used herein, the term “subject,” refers to an individual organism such as a human or an animal. In embodiments, the subject is a mammal (e.g., a human, a non-human primate, or a non-human mammal), a vertebrate, a laboratory animal, a domesticated animal, an agricultural animal, or a companion animal. In embodiments, the subject is a human (e.g., a human patient). In embodiments, the subject is a rodent, a mouse, a rat, a hamster, a rabbit, a dog, a cat, a cow, a goat, a sheep, or a pig.

[00169] As used herein, “a companion diagnostic” is an *in vitro* device, which provides information that is essential for the safe and effective use of a corresponding drug or biological product. In some embodiments, the method is used in conjunction as a companion diagnostic.

[00170] As used in this Specification and the appended claims, the singular forms “a,” “an” and “the” include plural referents unless the context clearly dictates otherwise.

[00171] Unless specifically stated or obvious from context, as used herein, the term “or” is understood to be inclusive and covers both “or” and “and”.

[00172] 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, for example within 2 standard deviations of the mean. About is understood to be within 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0.5%, 0.1%, 0.05%, or 0.01% of the stated value. Unless otherwise clear from the context, all numerical values provided herein are modified by the term “about.”

[00173] A stated range is understood to be any value between and at the limits of the stated range. As examples, a range between 1 and 5 includes 1, 2, 3, 4, and 5; a range between 1 and 10 includes 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10; and a range between 1 and 100 includes 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, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100.

Diagnostic Kits and Probe Sets

[00174] The disclosure further provides a kit or probe array containing nucleic acid primers

and/or probes for determining the level of expression in a patient tumor specimen or cell culture of a plurality of genes listed in Tables 1. The probe array may contain 3000 probes or less, 2000 probes or less, 1000 probes or less, 500 probes or less, so as to embody a custom set for preparing gene expression profiles described herein. In some embodiments, the kit may consist essentially of primers and/or probes related to evaluating drug-sensitivity/resistant in a sample, and primers and/or probes related to necessary or meaningful assay controls (such as expression level controls and normalization controls, as described herein under "Gene Expression Assay Formats")

[00175] The kit for evaluating drug-sensitivity/resistance may comprise nucleic acid probes and/or primers designed to detect the expression level of ten or more genes associated with drug sensitivity/resistance, such as the genes listed in Tables 1-2. The kit may include a set of probes and/or primers designed to detect or quantify the expression levels of at least 5, 7, 10, or 20 genes listed in one of Tables 1-2. The primers and/or probes may be designed to detect gene expression levels in accordance with any assay format, including those described herein under the heading "Assay Format." Illustrative assay formats include polymerase-based assays, such as RT-PCR, TaqMan™, hybridization-based assays, for example using DNA microarrays or other solid support, nucleic acid sequence based amplification (NASBA), flap endonuclease- based assays. The kit need not employ a DNA microarray or other high density detection format.

[00176] In accordance with this aspect, the probes and primers may comprise antisense nucleic acids or oligonucleotides that are wholly or partially complementary to the diagnostic targets described herein (*e.g.*, Tables 1-2). The probes and primers will be designed to detect the particular diagnostic target *via* an available nucleic acid detection assay format, which are well known in the art. The kits of the invention may comprise probes and/or primers designed to detect the diagnostic targets *via* detection methods that include amplification, endonuclease cleavage, and hybridization.

EXAMPLES

[00177] In order that the invention disclosed herein may be more efficiently understood, examples are provided below. It should be understood that these examples are for illustrative purposes only and are not to be construed as limiting the invention in any manner.

Example 1: Pan-Cancer Cell Line Analysis of lncRNAs Affecting Drug Response

[00178] To address the critical problem of drug resistance in cancer, a comprehensive genome wide approach based on large scale analysis of pharmacogenetics from cancer cell lines was developed followed by forward genetic screening utilizing CRISPRa technologies targeting both protein coding and non-coding genes. This approach identified putative protein coding and noncoding RNA biomarkers that correlate with Ara-C response and determined if biological processes predictive of drug response were enriched amongst the candidate genes identified. By subsequently integrating these correlative biomarkers with forward genetic CRISPRa screening approaches, results were able to identify high confidence clinically actionable modulators of chemotherapy response. The new platform was termed “DICaS” for Dual protein-coding and non-coding Integrated CRISPRa Screening.

[00179] In order to comprehensively define resistance mechanisms to chemotherapy, cellular responses to Ara-C was examined, a mainstay in the treatment of hematological malignancies. A computational strategy was developed to identify genes that correlate with sensitivity or resistance to Ara-C by correlating the pharmacological profiles from the Cancer Target Discovery and Development (CTD2) database, which harbors Ara-C sensitivity data for almost 1,000 human cancer cell lines, with the transcriptome profiles of 760 corresponding cancer cell lines from the Cancer Cell Line Encyclopedia (CCLE). In order to maximize the chance of identifying high confidence gene targets it was imperative to integrate analysis of as many cell lines as possible, however, drug sensitivities across the panel of cell lines formed a skewed distribution, likely arising from characteristics conferred by tissue of origin and histological subtype. Indeed, cancer cell type annotations were able to explain a substantial amount of the variation in drug sensitivities (adjusted $R^2 = 0.5123$, ANOVA $p < 2.2e-16$). For example, differentially annotated cell types are associated with differential drug resistance, but this difference was corrected after removing cell type annotations. Thus, using a linear regression model to remove these effects on the drug sensitivities established a normalized distribution of Ara-C sensitivity for the 760 cell lines analyzed (Figure 1A).

[00180] A correlation analysis between the adjusted drug sensitivities and gene expression levels across the 760 cell lines was subsequently performed (Figure 1B). To determine appropriate Z-score thresholds for correlated genes, a receiver operating characteristic (ROC)

analysis, which determines the Z-score that simultaneously maximizes the “true positive rate” (TPR) and minimizes the “false negative rate” (FPR) from the correlation profiles was performed. A set of genes across the literature that correspond to differential Ara-C sensitivity was compiled, which was denoted as positive controls. Negative controls were any gene in the analysis not in the list of positive controls. Thus, from the ROC analysis identified 24.4% of genes to be highly correlated with Ara-C sensitivity or resistance, corresponding to a Z-score threshold of 1.16.

[00181] Consequently, results identified a high enrichment for a number of genes known to be involved in the metabolism of Ara-C, illustrating the applicability of such an approach to identifying chemotherapy resistance mechanisms. This analysis identified low expression of deoxycytidine kinase (DCK) and equilibrative nucleoside transporter 1 (ENT1, SLC29A1) to be correlated with increased resistance to Ara-C ($Z = -2.51$ and -1.61 , respectively), whereas high expression of cytidine deaminase (CDA) and SAM Domain and HD Domain 1 (SAMHD1) correlated with increased resistance ($Z = 2.54$ and 2.03 , respectively) (Figure 1B). ENT1/SLC29A1 and DCK are enzymes involved in the transport and phosphorylation of Ara-C, respectively, thereby localizing and priming Ara-C for its cytotoxic activity. Decreased expression levels or inactivating mutations have already been demonstrated to correlate with increased resistance to Ara-C and poorer AML patient outcomes. On the other hand, CDA and SAMHD1 have been shown to decrease the effective amount of activated Ara-C in the treated cell, so their elevated expression effectively neutralizes its cytotoxic effect. Interestingly, data showed a number of cell-cycle and DNA damage regulators which have been previously implicated in modulation of AraC sensitivity including p21/CDKN1A, CHEK1, CDC6, and BRCA1 (Figure 1B).

[00182] To obtain a more comprehensive picture of the biological pathways predictive of Ara-C resistance, gene set enrichment analysis was performed (GSEA) on the drug sensitivity-gene expression correlations using the pathway annotations from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Figure 1C). Results identified positive enrichment of a number of signaling pathways known to promote cell survival, including the Jak-STAT (NES = 1.385, $p = 0.013$), PI3K-Akt (NES = 1.232, $p = 0.025$), and MAPK (NES = 1.222, $p = 0.042$) signaling pathways and negative enrichment of the pyrimidine metabolic pathway (NES = -2.456, $p =$

0.00016), mechanisms related to DNA damage (*e.g.*, p53 signaling pathway: NES = -2.293, $p = 0.00016$) and RNA regulatory mechanisms (*e.g.*, RNA degradation: NES = -2.613, $p = 1.6 \times 10^{-4}$) (Figure 1C-1D). Surprisingly, a significant negative enrichment of pathways involved in oxidative phosphorylation was also observed (NES = -1.305, $p = 0.059$), which have been shown to be active in leukemia resistant to Ara-C therapy. To further confirm the relevance of these pathways in human AML, pre-treatment AML transcriptome profiles with corresponding disease-free survival data corresponding to 121 patients treated with Ara-C from The Cancer Genome Atlas (TCGA) (Cancer Genome Atlas Research Network, 2013) was correlated, which identified a large number of enriched pathways shared with our cell line predictions, including oxidative phosphorylation (NES = -1.994, $p = 1.1 \times 10^{-4}$) and RNA regulatory mechanisms (*e.g.*, RNA degradation: NES = -1.702, $p = 0.0011$) (Figure 8A-8B).

[00183] As many non-coding genes are thought to regulate their loci in a cis-manner, it was relevant to evaluate coding and non-coding cognate gene pairs for correlation with either resistance or sensitivity to Ara-C within our dataset. Thus, a genome-wide set of 997 coding/non-coding sense/antisense gene pairs and compared their drug sensitivity-gene expression correlations was compiled. In agreement with the concept of cis regulation by noncoding RNAs, observations showed a significant positive correlation between sense-antisense gene expression levels across the panel of 760 cell lines (Pearson correlation, median $R = 0.5312$; Wilcoxon rank-sum test, $p < 2.2 \times 10^{-16}$) (Figure 1E). Furthermore, cognate gene pairs also demonstrate significant positive correlation in drug sensitivity (Pearson correlation, $R = 0.5636$, $p < 2.2 \times 10^{-16}$). Importantly, analysis of these same cognate gene pairs amongst the TCGA AML patient cohort, identified a similarly significant positive correlation of gene expression). Of the 997 sense-antisense gene pairs analyzed, 101 pairs are both highly correlated with Ara-C sensitivity or resistance and are highly correlated with each other, as compared to 10 pairs out of 5,000 random gene pairs (Chi squared test, $p < 2.2 \times 10^{-16}$). Interestingly, cognate sense genes were found to be positively enriched in PI3K-Akt (NES = 1.426, $p = 0.0764$) and MAPK signaling pathways (NES = 1.787, $p = 0.0040$), implicating these sense-antisense gene pairs in a number of the previously identified enriched pathways.

Example 2: A CRISPRa approach to study AML resistance to Ara-C

[00184] To determine if this predictive analysis could be functionally validated in a high-

throughput manner, an unbiased CRISPR-based approach to identify novel genes contributing to Ara-C resistance was applied. Forward genetic screening has proven to be a useful tool for the identification of protein coding genes affecting phenotypes *in vitro* and *in vivo*, but to date technological barriers have limited the ability to study non-coding genes in a similar manner. CRISPRa, which takes advantage of the CRISPR gene targeting strategy for gene activation, is a novel and promising approach that overcomes many of the prior obstacles hindering the study of non-coding genes, enabling the modulation of expression of both coding and lncRNA genes from their endogenous loci. A CRISPRa-based system in AML cell lines was established to provide a comprehensive and integrative genome-wide study of both the coding and non-coding genes that contribute to Ara-C resistance.

[00185] A first step included identifying the most appropriate cell line model to facilitate the screening process, and the MOLM14 AML cell line was selected. This cell line was selected because its IC₅₀ (~0.13 μM) ranks it among the most sensitive AML cell lines, as suggested by published IC₅₀s (Figure 2A). To ensure that MOLM14's sensitivity to Ara-C could be readily manipulated, both overexpression of the anti-apoptotic B-cell lymphoma 2 (BCL2) gene and shRNA-mediated knockdown of DCK were evaluated (Figure 2B). By testing the sensitivity of MOLM14 cell lines to Ara-C over a range of concentrations, it was found that the overexpression of BCL2 resulted in an approximately four-fold increase in the IC₅₀ for Ara-C (from ~0.13 μM to 0.5 μM), while knockdown of DCK resulted in a highly significant protective effect with an IC₅₀ for Ara-C increased almost 300-fold to approximately 3.5 μM (Figure 2B). Thus, these data indicated that MOLM14 represented an Ara-C sensitive cell line that could be readily manipulated for the purposes of identifying new genes that promote resistance to chemotherapy.

[00186] To identify the most efficient CRISPRa approach for use in MOLM14 for the screening purposes, CRISPRa was used. CRISPRa takes advantage of an enzymatically dead Cas9 (termed dCas9) component that recognizes and binds to sgRNAs. The dCas9 has been modified to promote gene expression by facilitating recruitment of transcriptional machinery to the gene transcriptional start site (TSS). Currently, several CRISPRa systems exist for the purposes of transcriptional activation, including the VP64-p65-Rta (VPR), SunTag and synergistic activation mediator (SAM) systems. It was demonstrated that the application of SAM

CRISPRa constructs achieves robust activation of both coding mRNA and non-coding lncRNA genes using a HEK293 cell line model. Thus, SAM-mediated CRISPRa in MOLM14 cells as compared with K562 and HL60 were tested, two additional leukemia cell lines of varying sensitivity to Ara-C. Using a panel of previously validated sgRNAs targeting the promoters of both coding (TTN, RHOF, ASCL1, HBG1) and noncoding (MIAT, TUNA) genes, confirmation of the ability of these guides to activate transcription in HEK293T cells as previously published and established that the majority of sgRNAs gave the highest activation in MOLM14 cells as compared with the other two hematopoietic cell lines (Figure 9). These results suggest that differences in CRISPRa mediated gene activation across cell lines for a given gene may be driven by the locus properties, such as chromatin accessibility as previously reported, highlighting the potential ramifications of cell line choice for CRISPRa studies.

Genome wide CRISPRa Screening of Protein-Coding Genes in AML

[00187] The CRISPRa platform was then used to screen for protein-coding genes whose expression modulates resistance to Ara-C. Screening was carried out using a previously published sgRNA library designed to enhance the transcription of ~23,000 RefSeq-annotated protein coding transcripts encoded in the human genome. MOLM14 cells were engineered to express the dCas9-VP64 SAM and the p65 and HSF1 transcription activators stably as described above (Figure 2C). In two independent experimental replicates, cells were transduced at a low multiplicity of infection (MOI) with the sgRNA library. Cells were selected for 5 days and then subsequently treated with 0.25 μ M Ara-C for 14 days with cell viability monitored over the treatment period (Figure S3B). 500 cells/sgRNA were collected both immediately after selection (T0) and again after 14 days of Ara-C treatment (T14). DNA was isolated from cellular populations, sgRNA sequences were amplified by PCR, and libraries were sequenced to identify enriched and/or depleted sgRNAs in each sample (Figure 2C). After quantifying guide RNA abundances, library preparations were analyzed for potential technical bias by examining guide abundance distributions, correlation between technical replicates, and sample clustering by principal component analysis (PCA). PCA analysis uncovered an experimental batch effect which correlated with the first principal component, explaining 47.3% of the variance across the samples. Transcript-level representation between T0 and T14 identified a host of genes enriched and depleted in Ara-C treated cells after adjusting for experimental batch (Figure 2D).

[00188] To determine an appropriate FDR cutoff for the screen, d an ROC analysis was performed using the list of annotated Ara-C-related genes, which identified an FDR < 33.9% as an ideal threshold. Interestingly, both correlation analysis and our forward genetic screen revealed DCK to be the most significantly depleted gene, thereby indicating that strong transcriptional activation of DCK by CRISPRa confers high sensitivity to Ara-C in MOLM14 cells (Figure 1B and 2D). Indeed, this was confirmed by overexpressing the top-scoring sgRNA targeting the DCK promoter in MOLM14 cells and treating these cells with Ara-C. Treatment with increasing concentrations of Ara-C for 48 hours confirmed that DCK overexpression sensitizes to the cytotoxic effects of Ara-C. Furthermore, multiple genes suspected to modulate sensitivity to Ara-C, such as RRM1/2, STAT3, SMAD5, and AKT1 were identified by this unbiased high-throughput approach.

[00189] To identify potential biological pathways that may contribute to differential gene representation upon treatment with Ara-C, gene set enrichment analysis using the KEGG gene set annotations were performed (Figure 2E) which identified a number of pathways congruent with the cell line analysis. These pathways again included the Jak-STAT (NES = 1.281, $p = 0.0565$), PI3K-Akt (NES = 1.291, $p = 0.0216$), and MAPK signaling pathways (NES = 1.535, $p = 3.1e-4$) (Figure 2E). Importantly, data showed that a large overlap of 2,411 genes significantly enriched/depleted in both the cell line and protein-coding CRISPRa screening.

[00190] A subset of these overlapping genes, including ZBP1, MUL1 and P14K2A, whose expression was associated with poor prognosis and decreased disease free survival were subsequently chosen for validation (Figure 2F). To further characterize the ability of induced coding gene expression for these genes, and a number of others, to resist Ara-C cytotoxicity, MOLM14 infected with relevant sgRNAs with a range of Ara-C doses up to 0.5 μM for 48 hours were treated (Figure 2G). Overexpression of each gene in the MOLM14 cells resulted in an increased resistance to Ara-C and an increased IC₅₀, correlating with the protective effect observed in our screening. This protective effect was also associated with decreased apoptosis of sgRNA infected MOLM14 cells as measured by Annexin-V positivity upon treatment with 0.25 μM Ara-C for 72 hours (Figure 2H). Importantly, in the absence of Ara-C, growth curve analysis failed to reveal any difference in proliferative capacity of cells overexpressing top scoring sgRNAs (Figure 2I). These data further support the hypothesis that these genes are

enriched in the screening on the basis of their protective effects against Ara-C cytotoxicity.

Example 3: Functional Genome Wide Screening of lncRNAs in AML

[00191] To study the functional roles of lncRNA genes in Ara-C resistance, an sgRNA library was designed, using a comprehensive set of 14,701 lncRNA genes, covering all major classifications of lncRNAs obtained by curating a merge of the human Gencode V22 and Broad Human lincRNA transcriptome annotations (Figure 3A). This approach was termed CRISPRa SAM-mediated approach “CRISPR activation of lncRNA” (CaLR) screening. One of the major advantages of CRISPRa is its ability to activate transcription at canonical or alternative TSSs in their endogenous genomic contexts. This is especially important in the study and analysis of lncRNAs, which are poorly annotated and for which many alternative TSSs have been proposed. Therefore, at least 4 sgRNAs per lncRNA covering a total of 22,253 TSSs using an algorithm that selected two proximal guides at least 50 bp upstream of individual TSSs, separated by a minimum of 50 bp, on both positive and negative strands was designed. Guides with degenerate nucleotides or inappropriate restriction sites were excluded from the library, resulting in 88,444 targeting guides (Figure 3A). To test the efficiency of activation within the library, the sgRNAs predicted to target the TUNA lncRNA gene (n=4 sgRNAs) was tested. A validated TSS of the MIAT lncRNA gene (MIAT-01, n=5 sgRNAs) and a predicted TSS of MIAT (MIAT-06, n=4 sgRNAs) (Figure 10B and 10C). It was found that each of these sets of sgRNAs contained at least two guides that promoted increased expression of the relevant lncRNA (Figure 10B), and at least one of the tested sgRNAs for each of TUNA and MIAT-01 was able to achieve a similar level of transcriptional activation as the corresponding validated sgRNAs. Furthermore, results were able to confirm the existence of the predicted TSS associated with MIAT-06, increasing the confidence in both the accuracy of the TSS annotations used and the ability of the sgRNA library to activate gene expression at these annotated TSSs.

[00192] An additional cohort of single sgRNAs was tested from the library for their ability to drive transcriptional activation of their respective gene products in both MOLM14 and HEK293T cells. While the majority of sgRNAs exhibited activity in at least one of the MOLM14 or HEK293T cell lines, data showed a significant degree of cell-specific transcriptional activation, highlighting the tissue-specific factors governing lncRNA expression. In addition, it was noted that one of these sgRNAs appeared to promote unexpected silencing of its targeted

promoter. Next the stable MOLM14 CRISPRa cell line was transduced with sgRNAs of the CaLR library at low MOI in two independent experimental replicates. Cells were selected for 5 days, and aliquots were collected for sequencing analysis to establish library representation at time zero (T0) as described above. Cells were subsequently treated with 0.25 μ M Ara-C for 14 days (T14), the remaining cells were harvested, and DNA was isolated from both T0 and T14 samples as outlined above (Figure 2C). Subsequently, sgRNA sequences were amplified by PCR, and libraries were sequenced to identify enriched and/or depleted sgRNAs in each sample.

[00193] After quantifying guide RNA abundances, library preparations were analyzed as above for potential technical bias. PCA analysis uncovered an experimental batch effect which correlated with the first principal component, explaining 46.2% of the variance across the samples. In order to estimate the false positive rate within the non-coding RNA screening, 99 non-targeting sgRNAs were included that should not be selected for in response to Ara-C treatment (Figure 3A). These 99 non-targeting sgRNAs behaved as expected across both experimental batches and were utilized to determine an appropriate FDR cut-off to control for our false positive rate, corresponding to 4.76%. This analysis identified 1,158 highly depleted and 2,000 highly enriched lncRNA transcripts (Figure 3B). Interestingly, several cancers associated lncRNA genes were identified within enriched sgRNAs, including sgRNAs targeting Taurine Up-Regulated 1 (TUG1), HOXA Transcript Antisense RNA, Myeloid-Specific 1 (HOTAIRM1) and Plasmacytoma variant translocation 1 (PVT1) (Figure 3B). Expression analysis of lncRNAs and coding genes from AML patient cohorts within the TCGA revealed that at least 50% of lncRNAs genes (both enriched and depleted) can be detected across these leukemias, compared to about 85% of protein coding genes (Figure 3C). As expected, transcript levels for individual protein coding mRNAs are significantly higher than the transcript levels for lncRNAs (Figure 3D; $p < 2.2e-16$). Interestingly, a significantly higher proportion of lncRNA genes found to be enriched following Ara-C treatment was detected (Chi-squared test, $p = 6.92e-3$) and the expression levels of these lncRNAs were significantly higher in AML patients (Wilcoxon rank-sum test, $p = 5.4e-7$), as compared with depleted lncRNA genes following Ara-C treatment (Figure 3C-3D), suggesting a functional role for lncRNAs in mediating resistance to Ara-C.

[00194] A guilt-by-association analysis was used on the lncRNAs significantly enriched in the

transcript-based analysis of the CaLR screen to infer putative regulatory networks contributing to modulation of Ara-C response. By carrying out co-expression analysis using the funcpred tool, putative functional annotations for these lncRNAs were generated. Coexpression was computed from RNAseq data for human tissues collected by the GTEx consortium, and functional enrichment was determined with respect to the MSigDB Hallmark Gene Sets. Using this approach identified two distinct gene set networks, each composed of a number of distinct cellular pathways. The two networks consist of (1) oxidative phosphorylation and fatty acid metabolism pathways and (2) leukemia development and progression. Enrichment of these pathways in the first network is reflective of the role of the mitochondria in regulating nucleotide metabolism, while specific pathways enriched in the latter network include leukemia associated pro-survival pathways (*e.g.*, Interferon response, IL6/JAK/STAT3 signaling, TNF α /NF κ B signaling), pathways associated with cell cycle regulation and proliferation and epithelial mesenchymal transition (EMT) related pathways (*e.g.*, TGF β signaling, EMT). Interestingly, JAK/STAT and TNF α /NF κ B are known to play important roles in the maintenance of normal hematopoiesis and are frequently deregulated in leukemia, while recent data has also highlighted an important role for EMT-related genes in the pathogenesis of AML, demonstrating the expression of multiple EMT-related genes to be associated with poor outcomes in human AML. Thus, the analysis of functional associations identified for significantly enriched lncRNAs from our screen point to deregulation of key pathways regulating cell proliferation and hematopoietic cell function. Although results were able to categorize broad groups of lncRNA function in this manner, there is still appreciable overlap between the various functional capabilities of these genes and the cellular programs that they associate with.

[00195] Because the understanding of the functional role of lncRNAs is limited to handful of well-studied examples, a short list of novel annotated lncRNAs was compiled to characterize further. These lncRNAs were significantly enriched in both our functional screening and our cell line analysis (Figure 3B). Co-expression analysis to associate individual lncRNA transcript levels with their most highly correlated protein coding genes from our analysis of the 760 CCLE cell line panel, again identified many of the pathways uncovered in our global analysis, suggesting that these lncRNAs play roles in survival pathways known to affect leukemia and drug resistance, including JAK/STAT, PI3K-AKT, RAS, EGFR and Hippo pathway as above (Figure

3E).

Example 4: Validation of Top lncRNA Candidates

[00196] To validate the findings from the screening experimentally, eleven genes significantly enriched were chosen, and two genes significantly depleted in our screening for further characterization (Figure 3B and 3E). Of these 13 genes selected from our screening, 10 were also found to be candidate genes predicted to influence Ara-C response in our cell line analysis. Individual sgRNAs targeting these lncRNA genes were overexpressed in the stable CRISPRa SAM MOLM14 cell line and treated with 0.25 μM Ara-C for 48 hours to evaluate the ability of these genes to protect cells from Ara-C-mediated cytotoxicity. The enriched sgRNAs resulted in a significant protection over control cells, with a 2.5-5.0-fold increase in cell viability post-Ara-C treatment (Figure 4A), while the two depleted genes resulted in decreased viability of approximately 2 fold in response to Ara-C (Figure 4A). To confirm the specificity of these sgRNAs for their target lncRNAs, the fold induction in expression of the lncRNA induced by these sgRNAs post-infection was examined (Figure 4B). Indeed, data confirmed that each of the guides promoted a strong expression of their individual lncRNA targets, with expression increased across the different sgRNAs examined (Figure 4B). To further characterize the ability of induced lncRNA expression to resist Ara-C cytotoxicity, MOLM14 infected with the relevant sgRNAs were treated with a range of Ara-C doses up to 0.5 μM for 48 hours (Figure 4C). Overexpression of each enriched lncRNA in the MOLM14 cells resulted in a decreased sensitivity to Ara-C and an increased IC₅₀ (Figure 4C), correlating with the protective effect observed in Figure 4A, while the depleted lncRNA genes also behaved as expected.

[00197] To address how these lncRNAs may be promoting cell viability in the presence of Ara-C, the ability of each of the targeted lncRNAs to promote either increased proliferation of the MOLM14 cells or increased survival upon treatment was examined. Growth curve analysis was carried out over a period of 4 days in the absence of Ara-C treatment, and proliferation was assessed by MTS-viability assay for each of the sgRNAs. Out of the candidate lncRNAs, three appeared to promote proliferation (AL353148.1, LINC02426; AL157688.1) (Figure 4D), which was confirmed by BrdU (Bromodeoxyuridine) incorporation. These data indicated that enrichment of the corresponding sgRNAs might be facilitated by the ability of their corresponding lncRNA to drive increased proliferation. On the other hand, to evaluate the

ability of candidate lncRNAs to promote survival, MOLM14 cells stably infected with sgRNAs were treated with 0.25 μM Ara-C for 72 hours, and cells were then stained with Annexin-V and propidium iodide (PI). Control cells demonstrated approximately 40% of Annexin-V/PI double positive staining at this time, and while all sgRNAs were able to promote increased survival to some extent (Figure 4E), several targeted lncRNAs demonstrated a significant ability to protect from apoptosis (AC012150.1; GAS6-AS2) (Figure 4E, right panel). This marked survival advantage for these sgRNAs provides compelling evidence for an anti-apoptotic role for their targeted lncRNAs. These results were further confirmed in an independent HL-60 hematopoietic cell line for sgRNAs targeting the AC106897.1, AC008073.2, AC091982.2, and GAS6-AS2 lncRNA genes, demonstrating that these phenotypes are not specific to the MOLM14 cell line.

[00198] Given that Ara-C promotes extensive genotoxic stress and the analysis of the significantly enriched lncRNAs from the CaLR screen are suggested to regulate cell cycle regulation and DNA repair, it was hypothesized that some of these lncRNAs may impact these pathways. Thus, MOLM14 cells overexpressing targeted lncRNAs were examined for DNA damage as induced by Ara-C. Cells were again treated with 0.25 μM Ara-C for 24 hours and subsequently stained by immunofluorescence for the presence of phosphorylated $\gamma\text{H2A.X}$, a marker of DNA damage response (DDR) (Figure 4F, left panels). The number of phospho- $\gamma\text{H2A.X}$ positive foci were counted, and average numbers of foci/cell were used to evaluate the overall DDR. One of the candidate lncRNA genes triggered a strong reduction in phospho- $\gamma\text{H2A.X}$ staining (AL353148.1) (Figure 4F), suggesting that the protective effect afforded by this lncRNA gene may be mediated by (1) an ability of the cells to detoxify the DNA damage-inducing Ara-C nucleotides more strongly, (2) a deficiency in the ability of the cells to recognize and respond appropriately to damage, or (3) an increased ability of these cells to repair the damage induced by Ara-C. For two of our candidate lncRNAs (GAS6-AS2 and AC008073.2), not only were they identified as candidates in both our cell line analysis and functional CaLR screening, but higher expression levels of these lncRNA genes were also associated with poor prognosis and decreased disease-free survival in AML patients treated with Ara-C (Figure 4G). Taken together these data reify our screening process as a platform to identify clinically relevant lncRNAs that may modulate Ara-C cytotoxicity through targeting a number of cellular processes.

Example 5: lncRNA GAS6-AS2 Regulates the GAS/AXL Signaling Axis

[00199] The computational pharmacogenomic cell line based analysis was integrated with both the coding and non-coding functional screens. Statistical analysis of cognate gene pairs across the integrated analysis demonstrated significant enrichment of 7 sense-antisense gene pairs, relative to a randomized pool of 5000 coding-noncoding gene pairs (Chi-Squared Test, $p < 2.2 \times 10^{-16}$). These seven sense-antisense cognate gene pairs were predicted to regulate resistance to Ara-C (Figure 5A). Interestingly, the activity of these sense-antisense cognate gene pairs was found to be both correlative and anti-correlative relative to each other, representing potentially diverse regulatory modules between coding and non-coding genes whereby noncoding may either positively or negatively regulate expression of their cognate coding gene. Of the 7 cognate pairs identified, GAS6/GAS6-AS2 appeared to be one of the best candidates for further analysis based on the enrichment of both the coding and noncoding genes. In addition, while GAS6 is already known to play an important role in drug resistance in cancer, including AML, the role and function of GAS6-AS2 remains unknown. Therefore, its role and function was characterized further.

[00200] To confirm the on-target effect of our CRISPRa sgRNA, the 8 different sgRNAs within the library that were designed to target the GAS6-AS2 promoter were overexpressed (Table 3). Each of these sgRNAs were initially examined for their ability to promote cell viability upon treatment with 0.25 μM Ara-C for 48 hours. As expected, the majority of these sgRNAs led to a significant increase in cell survival (Figure 5B). Importantly, a strong correlation between the levels of GAS6-AS2 activation and the resistance to Ara-C was observed (Figure 5B and 5C). While sgRNAs #2 and #6 failed to achieve a statistically significant increase in fold viability (Figure 5B), these sgRNAs led only to a moderately activated GAS6-AS2, indicating a dose dependent effect and specificity of GAS6-AS2 as directly responsible for the increased tolerance of MOLM14 cells against Ara-C. Additionally, similar to sgRNA #4, which was validated, as outlined in Figure 4, overexpression of sgRNAs #1 and #3 clearly promoted decreased sensitivity to Ara-C, as shown by increased IC₅₀ (Figure 5C and 5D). Moreover, both sgRNAs #1 and #3 also promoted a similarly potent ability to reduce apoptosis of MOLM14 infected cells when treated with 0.25 μM Ara-C for 72 hours (Figure 5E). Thus, the GAS6-AS2 lncRNA appears to be a bona fide promoter of resistance to the effects of the chemotherapy drug

Ara-C and represents a novel candidate gene to promote resistance to therapy in AML.

[00201] AML is known to develop as a multi-clonal disease, and resistant clones are frequently observed to exist even in early stages of the disease. The selective pressure of treatment leads to rapid clonal evolution and the emergence of resistant clones. Indeed, following Ara-C treatment of a mixed population of two MOLM14 cells, one expressing high GAS6-AS2 (labeled with red fluorescent protein) and the other expressing a non-targeting sgRNA (labeled with blue fluorescent protein), the GAS6-AS2 expressing clone emerged as dominant and was significantly enriched post-treatment (Figure 5F). These results were also confirmed *in vivo*. To this end, equal numbers of both GAS6-AS2 expressing and control cells were transplanted into immunocompromised NSG mice as outlined in Figure 5G. Engrafted mice were then treated with Ara-C (30 mg/kg/day) for 5 days and allowed to recover for an additional 2 days (Figure 5G). Bone marrow was harvested and analyzed by flow cytometry. Results showed significant enrichment ($p=0.002$) of the GAS6-AS2-Red over the Control-Blue-labeled cells (Red/Blue cells=10.9±6.4), consistent with the *in vitro* observations (Figure 5H-5I). Importantly, within a non-treated cohort, both populations of cells were observed to be present in an equal ratio, demonstrated that GAS6-AS2-overexpressing cells did not exhibit a proliferative advantage over control cells (Figure 5H and 5I). Furthermore, mice transplanted with GAS6-AS2 overexpressing cells alone had a greater tumor burden post-Ara-C treatment as compared with control cells, again demonstrating the capacity of GAS6-AS2 to promote resistance to Ara-C treatment.

[00202] Several lncRNAs have been shown to exert their functional role by cis-regulation of neighboring genes and this is further supported by our genome wide analysis of sense-antisense cognate gene pairs (Figure 1E-1F and Figure 5A). As GAS6-AS2 lies in an antisense head-to-head manner with the coding gene GAS6, it was suggested that the GAS6/GAS6-AS2 cognate gene pair may function in this manner. Importantly, the GAS6-AS2 lncRNA displayed nuclear (and cytoplasmic) localization suggesting that it may have the potential to regulate transcription of the GAS6 gene. GAS6 is an important ligand for the Tyro3-Axl-Mer (TAM) receptor tyrosine kinase signaling axis, controlling known pro-survival signals which are upregulated in AML, as well as additional cancer types. Indeed, upregulation of GAS6/TAM signaling strongly correlates with resistance to chemotherapy and is a predictor of poor survival. In line with this suggestion, GAS6 expression levels were found to be strongly correlated with GAS6-AS2 expression in

MOLM14 cells after GAS6-AS2 CRISPRa modulation (Figure 6A). In addition, observation in a striking correlation between these cognate gene pairs across the diverse 760 CCLE cell line panel (Pearson's $R = 0.8762$, $p < 2.2e-16$) (Figure 6B), as well as for a diverse set of primary human cancer types including AML (Figure 6C).

[00203] Activation of the GAS6/TAM pathway has been reported to promote MAPK, JAK/STAT and NFkB signaling (Schoumacher and Burbridge, 2017), with both MEK-ERK and S6K-RPS6 signaling axes as known downstream targets of TAM signaling. Thus, the transcriptional activation of GAS6-AS2 was next examined to see if it also had the capacity to promote downstream signaling from GAS6/TAM. To this end, MOLM14 cells were transduced with one of either 3 individual sgRNAs targeting GAS6-AS2 or 3 non-targeting control sgRNAs. Western blot analysis of lysates from these cells confirmed activation of the pathway (Figure 6D). Importantly, results showed both pERK and pRPS6 to be strongly phosphorylated in response to GAS6-AS2 activation (Figure 6D). Surprisingly in a variety of cancer subtypes, including AML, GAS6-AS2 expression levels share strong correlations not only with GAS6 but also to its target receptor AXL (Figure 6E). This correlation with AXL was also mirrored in our 760 CCLE cell line panel (Pearson's $R = 0.6064$, $p < 2.2e-16$) (Figure 6F), and by overexpression of GAS6-AS2 in both MOLM14 and HEK293 cells *in vitro*. These data suggested that GAS6-AS2 may be able to regulate the TAM receptor signaling axis at a number of levels.

[00204] To further investigate the role of GAS6-AS2 in regulating GAS6 and AXL, the K562 leukemia cell line, which was found to express high levels of GAS6-AS2, GAS6, and AXL relative to MOLM14 was used (Figure 6G), and which was demonstrated to be highly resistant to Ara-C treatment (Figure 6G). Thus, GAS6-AS2 activity was attenuated in K562 cells by knockdown of the lncRNA using two specific locked nucleic acid (LNA)-enhanced anti-sense oligonucleotides (ASO), which leads to RNase H-dependent decay of RNA transcripts in the nucleus. Indeed, GAS6-AS2 knockdown by ASO leads to a significant decrease in both GAS6 and AXL levels (Figure 7A), as well as sensitizing K562 cells to the activity of Ara-C (Figure 6B).

[00205] Previous studies found that AXL transcription is regulated by methylation of CpGs upstream of its TSS. Direct methylation analysis using a bisulfite assay identified 6 highly

methyated sites in the AXL promoter. Correspondingly, GAS6-AS2-overexpressing cells show significant decreases in the methylation levels at these highly methylated CpG sites (Figure 7C), suggesting that GAS6-AS2 has the potential to regulate both GAS6 in a cis-acting manner and AXL in a trans-acting manner. To characterize the global function of GAS6-AS2 in cancer, an unbiased Kmeans clustering based on the coding and non-coding gene expression across 53 AML patients was performed. A large number of genes known to be regulated by promoter methylation were clustered together with GAS6-AS2 (Figure 7D), supporting our hypothesis that GAS6-AS2 regulates gene expression of its targets through CpG modification. Based on these results, it was suggested that GAS6-AS2 may function through a DNA methyltransferase, and in so doing act to regulate CpG methylation of the AXL promoter. Thus, an unbiased screening of the 760 CCLE cell line panel was used for candidate DNA methyltransferases that would correlate with Ara-C sensitivity. Interestingly, results showed decreased expression of DNMT1 and DNMT3A as strongly associated with increased resistance amongst our cell line panel (Figure 7E). Importantly, GAS6-AS2 was observed to be significantly enriched in RNAs bound to DNMT1 relative to IgG negative control in a recently published DNMT1 RNA-IP sequencing (RIP-Seq) dataset (Figure 7F) suggesting that GAS6-AS2 mediates transregulation of AXL by coordinating activity of DNMT proteins at the AXL promoter. Thus, this data supports a model whereby increased transcription and expression of GAS6-AS2 promotes upregulation of both the GAS6 ligand and its TAM receptors to promote cellular survival and resistance to Ara-C treatment in AML (Figure 7G).

[00206] Accordingly, the present disclosure provides a superior computational approach which is more clinically relevant by removing the effect of cancer/tissue subtype prior to predictive modeling, thereby focusing the analysis on genetic features not associated with differences between cancer subtypes in the analyzed data sets.

OTHER EMBODIMENTS

[00207] It is to be understood that while the disclosure has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the disclosure, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

INCORPORATION BY REFERENCE

[00208] All patents and publications referenced herein are hereby incorporated by reference in their entireties.

[00209] The publications discussed herein are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that the present invention is not entitled to antedate such publication by virtue of prior invention.

[00210] As used herein, all headings are simply for organization and are not intended to limit the disclosure in anyway.

References

- Aken, B.L., Ayling, S., Barrell, D., Clarke, L., Curwen, V., Fairley, S., Fernandez Banet, J., Billis, K., García Girón, C., Hourlier, T., *et al.* (2016). The Ensembl gene annotation system. Database (Oxford). 2016.
- Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with highthroughput sequencing data. *Bioinformatics* 31, 166–169.
- Bachas, C., Schuurhuis, G.J., Zwaan, C.M., van den Heuvel-Eibrink, M.M., den Boer, M.L., de Bont, E.S.J.M., Kwidama, Z.J., Reinhardt, D., Creutzig, U., de Haas, V., *et al.* (2015). Gene Expression Profiles Associated with Pediatric Relapsed AML. *PLoS One* 10, e0121730.
- Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A. a., Kim, S., Wilson, C.J., Lehár, J., Kryukov, G. V., Sonkin, D., *et al.* (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 483, 603–307.
- Basu, A., Bodycombe, N.E., Cheah, J.H., Price, E. V., Liu, K., Schaefer, G.I., Ebright, R.Y., Stewart, M.L., Ito, D., Wang, S., *et al.* (2013). An interactive resource to identify cancer genetic and lineage dependencies targeted by small molecules. *Cell* 154, 1151–1161.
- Benekli, M. (2002). Constitutive activity of signal transducer and activator of transcription 3 protein in acute myeloid leukemia blasts is associated with short disease-free survival. *Blood* 99, 252–257.
- Beroukhi, R., Mermel, C.H., Porter, D., Wei, G., Raychaudhuri, S., Donovan, J., Barretina, J., Boehm, J.S., Dobson, J., Urashima, M., *et al.* (2010). The landscape of somatic copy-number alteration across human cancers. *Nature* 463, 899–905.
- Cabili, M., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., and Rinn, J.L. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* 25, 1915–1927. Cancer Cell Line Encyclopedia Consortium, and Genomics of Drug Sensitivity in Cancer
- Consortium (2015). Pharmacogenomic agreement between two cancer cell line data sets. *Nature* 528, 84–87. Cancer Genome Atlas Research Network (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N. Engl. J. Med.* 368, 2059–2074.
- Chavez, A., Scheiman, J., Vora, S., Pruitt, B.W., Tuttle, M., P R Iyer, E., Lin, S., Kiani, S., Guzman, C.D., Wiegand, D.J., *et al.* (2015). Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods* 1–5.
- Chavez, A., Tuttle, M., Pruitt, B.W., Ben, E.-C., Chari, R., Dmitry, T.-O., Haque, S.J., Cecchi, R.J., Kowal, E.J.K., Buchthal, J., *et al.* (2016). Comparison of Cas9 activators in multiple species. *Nat Methods* 13, 563–567.

- Derrien, T., Johnson, R., Bussotti, G., Tanzer, a, Djebali, S., Tilgner, H., Guernec, G., Merkel, a, Gonzalez, D., Lagarde, J., *et al.* (2012). The GENCODE v7 catalogue of human long noncoding RNAs: Analysis of their structure, evolution and expression. 1775–1789.
- Dimitrova, N., Zamudio, J.R., Jong, R.M., Soukup, D., Resnick, R., Sarma, K., Ward, A.J., Raj, A., Lee, J.T., Sharp, P. a., *et al.* (2014). LincRNA-p21 Activates p21 In cis to Promote Polycomb Target Gene Expression and to Enforce the G1/S Checkpoint. *Mol. Cell* 54, 777–790.
- Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., *et al.* (2012). Landscape of transcription in human cells. *Nature* 489, 101–108.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21.
- Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., and Huber, W. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* 21, 3439–3440.
- Engreitz, J.M., Haines, J.E., Perez, E.M., Munson, G., Chen, J., Kane, M., McDonel, P.E., Guttman, M., and Lander, E.S. (2016). Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature* 539, 452–455.
- Farge, T., Saland, E., de Toni, F., Aroua, N., Hosseini, M., Perry, R., Bosc, C., Sugita, M., Stuani, L., Fraisse, M., *et al.* (2017). Chemotherapy-resistant human acute myeloid leukemia cells are not enriched for leukemic stem cells but require oxidative metabolism. *Cancer Discov.* 7, 716–735.
- Galmarini, C.M., Thomas, X., Graham, K., El Jafaari, A., Cros, E., Jordheim, L., Mackey, J.R., and Dumontet, C. (2003). Deoxycytidine kinase and cN-II nucleotidase expression in blast cells predict survival in acute myeloid leukaemia patients treated with cytarabine. *Br. J. Haematol.* 122, 53–60.
- Garnett, M.J., Edelman, E.J., Heidorn, S.J., Greenman, C.D., Dastur, A., Lau, K.W., Greninger, P., Thompson, I.R., Luo, X., Soares, J., *et al.* (2012). Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature* 483, 570–575.
- Garzon, R., Volinia, S., Papaioannou, D., Nicolet, D., Kohlschmidt, J., Yan, P.S., Powell, B.L., Kolitz, J.E., Moore, J.O., Eisfeld, A., *et al.* (2014). Expression and prognostic impact of lncRNAs in acute myeloid leukemia. *Proc. Natl. Acad. Sci. U. S. A.* 1, 1–6.

- Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., and Ploegh, H.L. (2014). Resource Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* 159, 647–661.
- Herold, N., Rudd, S.G., Ljungblad, L., Sanjiv, K., Myrberg, I.H., Paulin, C.B.J., Heshmati, Y., Hagenkort, A., Kutzner, J., Page, B.D.G., *et al.* (2017). Targeting SAMHD1 with the Vpx protein to improve cytarabine therapy for hematological malignancies. *Nat. Med.* 23, 256–263.
- Hon, C., Ramilowski, J., Harshbarger, J., Bertin, N., Rackham, O., Gough, J., Denisenko, E., Schmeier, S., Poulsen, T., Severin, J., *et al.* (2017). An atlas of human long non-coding RNAs with accurate 5' ends. *Nature*.
- Hong, C.C., Lay, J.D., Huang, J.S., Cheng, A.L., Tang, J.L., Lin, M.T., Lai, G.M., and Chuang, S.E. (2008). Receptor tyrosine kinase AXL is induced by chemotherapy drugs and overexpression of AXL confers drug resistance in acute myeloid leukemia. *Cancer Lett.* 268, 314–324.
- Horlbeck, M.A., Gilbert, L.A., Villalta, J.E., Adamson, B., Pak, R.A., Chen, Y., Fields, A.P., Park, C.Y., Corn, J.E., Kampmann, M., *et al.* (2016). Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *Elife* 5, 1–20.
- Huarte, M. (2015). The emerging role of lncRNAs in cancer. *Nat. Med.* 21, 1253–1261.
- Imielinski, M., Berger, A.H., Hammerman, P.S., Hernandez, B., Pugh, T.J., Hodis, E., Cho, J., Suh, J., Capelletti, M., Sivachenko, A., *et al.* (2012). Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* 150, 1107–1120.
- Iyer, M.K., Niknafs, Y.S., Malik, R., Singhal, U., Sahu, A., Hosono, Y., Barrette, T.R., Prensner, J.R., Evans, J.R., Zhao, S., *et al.* (2015). The landscape of long noncoding RNAs in the human transcriptome. *Nat. Genet.* 47, 199–208.
- Jayavelu, K., and Srivastava, V.M. (2015). Pharmacogenomics RNA expression of genes involved in cytarabine metabolism and transport. *16*, 877–890.
- Johnson, W.E., Li, C., and Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8, 118–127.
- Joung, J., Konermann, S., Gootenberg, J.S., Abudayyeh, O.O., Platt, R.J., Brigham, M.D., Sanjana, N.E., and Zhang, F. (2017). Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nat. Protoc.* 12, 828–863.
- Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2014). Data,

information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res.* **42**, D199–D205.

Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., *et al.* (2014). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **2**–11.

Lamba, J.K. (2009). Genetic factors influencing cytarabine therapy. *Pharmacogenomics* **10**, 1657–1674.

Lee-Sherick, a B., Eisenman, K.M., Sather, S., McGranahan, A., Armistead, P.M., McGary, C.S., Hunsucker, S. a, Schlegel, J., Martinson, H., Cannon, C., *et al.* (2013). Aberrant Mer receptor tyrosine kinase expression contributes to leukemogenesis in acute myeloid leukemia. *Oncogene* **32**, 1–10.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079.

Li, W., Xu, H., Xiao, T., Cong, L., Love, M.I., Zhang, F., Irizarry, R.A., Liu, J.S., Brown, M., and

Liu, X.S. (2014). MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* **15**, 554.

Liu, S.J., Horlbeck, M.A., Cho, S.W., Birk, H.S., Malatesta, M., He, D., Attenello, F.J., Villalta, J.E., Cho, M.Y., Chen, Y., *et al.* (2017). CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. *Science* (80). **355**, eaah7111.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550.

Macanas-Pirard, P., Broekhuizen, R., González, A., Oyanadel, C., Ernst, D., García, P., Montecinos, V.P., Court, F., Ocqueteau, M., Ramirez, P., *et al.* (2017). Resistance of leukemia cells to cytarabine chemotherapy is mediated by bone marrow stroma, involves cell-surface equilibrative nucleoside transporter-1 removal and correlates with patient outcome. *Oncotarget* **8**, 23073–23086.

Malani, D., Murumägi, A., Yadav, B., Kontro, M., Eldfors, S., Kumar, A., Karjalainen, R., Majumder, M.M., Ojamies, P., Pemovska, T., *et al.* (2017). Enhanced sensitivity to glucocorticoids in cytarabine-resistant AML. *Leukemia* **31**, 1187–1195.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*; Vol 17, No 1 Next Gener. Seq. Data Anal.

- Mudduluru, G., and Allgayer, H. (2008). The human receptor tyrosine kinase Axl gene promoter characterization and regulation of constitutive expression by Sp1, Sp3, and CpG methylation. *Biosci. Rep.* 28, 161–176.
- Negoro, E., Yamauchi, T., Urasaki, Y., Nishi, R., Hori, H., and Ueda, T. (2011). Characterization of cytarabine-resistant leukemic cell lines established from five different blood cell lineages using gene expression and proteomic analyses. *Int. J. Oncol.* 38, 911–919.
- Perron, U., Provero, P., and Molineris, I. (2017). In silico prediction of lncRNA function using tissue specific and evolutionary conserved expression. *BMC Bioinformatics* 18, 144.
- Quinn, J.J., and Chang, H.Y. (2016). Unique features of long non-coding RNA biogenesis and function. *Nat. Rev. Genet.* 17, 47–62.
- Ramos, N., Mo, C., Karp, J., and Hourigan, C. (2015). Current Approaches in the Treatment of Relapsed and Refractory Acute Myeloid Leukemia. *J. Clin. Med.* 4, 665–695.
- Rathe, S.K., Moriarity, B.S., Stoltenberg, C.B., Kurata, M., Aumann, N.K., Rahrmann, E.P., Bailey, N.J., Melrose, E.G., Beckmann, D. a, Liska, C.R., *et al.* (2014). Using RNA-seq and targeted nucleases to identify mechanisms of drug resistance in acute myeloid leukemia. *Sci. Rep.* 4, 6048.
- Rees, M.G., Seashore-Ludlow, B., Cheah, J.H., Adams, D.J., Price, E. V, Gill, S., Javaid, S., Coletti, M.E., Jones, V.L., Bodycombe, N.E., *et al.* (2016). Correlating chemical sensitivity and basal gene expression reveals mechanism of action. *Nat. Chem. Biol.* 12, 109–116.
- Rinn, J.L., and Chang, H.Y. (2012). Genome Regulation by Long Noncoding RNAs. *Annu. Rev. Biochem.* 81, 145–166. Di Ruscio, A., Ebralidze, A.K., Benoukraf, T., Amabile, G., Goff, L.A., Terragni, J., Figueroa, M.E., De Figueiredo Pontes, L.L., Alberich-Jorda, M., Zhang, P., *et al.* (2013). DNMT1- interacting RNAs block gene-specific DNA methylation. *Nature* 503, 371–376.
- Sahu, A., Singhal, U., and Chinnaiyan, A.M. (2015). Long Noncoding RNAs in Cancer: From Function to Translation. *Trends in Cancer* 1, 93–109.
- Saland, E., Boutzen, H., Castellano, R., Pouyet, L., Griessinger, E., Larrue, C., de Toni, F., Scotland, S., David, M., Danet-Desnoyers, G., *et al.* (2015). A robust and rapid xenograft model to assess efficacy of chemotherapeutic agents for human acute myeloid leukemia. *Blood Cancer J.* 5, e297.
- Schmitt, A.M., and Chang, H.Y. (2016). Long Noncoding RNAs in Cancer Pathways. *Cancer Cell* 29, 452–463.
- Schneider, C., Oellerich, T., Baldauf, H.-M., Schwarz, S.-M., Thomas, D., Flick, R., Bohnenberger, H., Kaderali, L., Stegmann, L., Cremer, A., *et al.* (2016). SAMHD1 is a

biomarker for cytarabine response and a therapeutic target in acute myeloid leukemia. *Nat. Med.* **23**, 250–255.

Schoumacher, M., and Burbridge, M. (2017). Key Roles of AXL and MER Receptor Tyrosine Kinases in Resistance to Multiple Anticancer Therapies. *Curr. Oncol. Rep.* **19**, 19.

Seashore-Ludlow, B., Rees, M.G., Cheah, J.H., Cokol, M., Price, E. V., Coletti, M.E., Jones, V., Bodycombe, N.E., Soule, C.K., Gould, J., *et al.* (2015). Harnessing Connectivity in a Large-Scale Small-Molecule Sensitivity Dataset. *Cancer Discov.* **5**, 1210–1223.

Shechner, D.M., Hacisuleyman, E., Younger, S.T., and Rinn, J.L. (2015). Multiplexable, locus specific targeting of long RNAs with CRISPR-Display. *Nat. Methods* **12**, 664–670.

Smoot, M.E., Ono, K., Ruscheinski, J., Wang, P.-L., and Ideker, T. (2011). Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **27**, 431–432.

Stavropoulou, V., Kaspar, S., Brault, L., Sanders, M.A., Juge, S., Morettini, S., Tzankov, A., Iacovino, M., Lau, I.-J., Milne, T.A., *et al.* (2016). MLL-AF9 Expression in Hematopoietic Stem Cells Drives a Highly Invasive AML Expressing EMT-Related Genes Linked to Poor Outcome. *Cancer Cell* **30**, 43–58.

Steelman, L.S., Pohnert, S.C., Shelton, J.G., Franklin, R.A., Bertrand, F.E., and McCubrey, J.A. (2004). JAK/STAT, Raf/MEK/ERK, PI3K/Akt and BCR-ABL in cell cycle progression and leukemogenesis. *Leukemia* **18**, 189–218.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., *et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550.

Takagaki, K., Katsuma, S., Kaminishi, Y., Horio, T., Nakagawa, S., Tanaka, T., Ohgi, T., and Yano, J. (2004). Gene-Expression Profiling Reveals Down-Regulation of Equilibrative Nucleoside Transporter 1 (ENT1) in Ara-C-Resistant CCRF-CEM-Derived Cells. *J. Biochem.* **136**, 733–740.

Tibes, R., Bogenberger, J.M., Chaudhuri, L., Hagelstrom, R.T., Chow, D., Buechel, M.E., Gonzales, I.M., Demuth, T., Slack, J., Mesa, R. a, *et al.* (2012). RNAi screening of the kinome with cytarabine in leukemias. *Blood* **119**, 2863–2872.

Vis, D.J., Bombardelli, L., Lightfoot, H., Iorio, F., Garnett, M.J., and Wessels, L.F. (2016). Multilevel models improve precision and speed of IC50 estimates. *Pharmacogenomics* **17**, 691–700.

Wang, T., Wei, J.J., Sabatini, D.M., and Lander, E.S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80–84.

Whitman, S.P., Kohlschmidt, J., Maharry, K., Volinia, S., Mrózek, K., Nicolet, D., Schwind, S., Becker, H., Metzeler, K.H., Mendler, J.H., *et al.* (2014). GAS6 expression identifies high-risk adult AML patients: potential implications for therapy. *Leukemia* 28, 1252–1258.

Wu, G., Ma, Z., Hu, W., Wang, D., Gong, B., Fan, C., Jiang, S., Li, T., Gao, J., and Yang, Y. (2017). Molecular insights of Gas6/TAM in cancer development and therapy. *Cell Death Dis.* 8, e2700.

Xu, L., Chen, Y., Dutra-Clarke, M., Mayakonda, A., Hazawa, M., Savinoff, S.E., Doan, N., Said, J.W., Yong, W.H., Watkins, A., *et al.* (2017). BCL6 promotes glioma and serves as a therapeutic target. *Proc. Natl. Acad. Sci. U. S. A.* 114, 3981–3986.

Yang, W., Soares, J., Greninger, P., Edelman, E.J., Lightfoot, H., Forbes, S., Bindal, N., Beare, D., Smith, J. a., Thompson, I.R., *et al.* (2013). Genomics of Drug Sensitivity in Cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* 41, 955–961.

CLAIMS

What is claimed is:

1. A method for determining an RNA signature indicative of sensitivity or resistance to a candidate cancer therapy, the method comprising:

identifying one or more coding RNAs whose level of expression correlates with sensitivity or resistance to a candidate cancer therapy in cells of a cancer subtype, and
identifying one or more noncoding RNAs whose level of expression correlates with sensitivity or resistance to the candidate cancer therapy in cells of the cancer subtype,

thereby determining an RNA signature that is indicative of sensitivity or resistance to the candidate cancer therapy; and

determining the presence or absence of the RNA signature in a cell or tissue of the cancer subtype, and classifying the cell or tissue for sensitivity or resistance to the candidate cancer therapy.
2. The method of claim 1, wherein the cancer subtype is defined by tissue of origin and histological subtype.
3. The method of claim 1, wherein the cancer subtype is a solid tumor.
4. The method of claim 1, wherein the cancer subtype is a hematological malignancy.
5. The method of claim 1, wherein the cancer subtype is a sarcoma, which is optionally squamous cell carcinoma, fibrosarcoma, myosarcoma, osteogenic sarcoma, angiosarcoma, or endotheliosarcoma.
6. The method of claim 1, wherein the cancer subtype is a carcinoma, which is optionally adenocarcinoma.
7. The method of claim 1, wherein the cancer subtype is Small Cell Lung Cancer (CLC), Non-Small Cell Lung Cancer (NSCLC), or mesothelioma.
8. The method of claim 1, wherein the cancer subtype is a brain cancer or glioblastoma.
9. The method of claim 1, wherein the cancer subtype is a breast cancer, lymphoma, prostate cancer, pancreatic cancer, liver cancer, kidney cancer, colon or colorectal cancer,

ovarian cancer, endometrial cancer, cervical cancer, testicular cancer, or melanoma.

10. The method of claim 1, wherein the cancer is a leukemia, which is optionally acute myeloid leukemia (AML), chronic myelogenous leukemia (CML), or acute lymphoblastic leukemia (ALL).

11. The method of any one of claims 1 to 10, wherein the cells of a cancer subtype are cell lines.

12. The method of any one of claims 1 to 10, wherein the cells of a cancer subtype are primary cell cultures.

13. The method of claim 11 or claim 12, wherein the cells of a cancer subtype comprise at least about 50 or at least about 100 or at least about 200 cell lines or primary cultures.

14. The method of claim 13, wherein the sensitivity or resistance of the cell lines or primary cell cultures to the candidate cancer therapy is determined by a chemosensitivity assay.

15. The method of any one of claims 1 to 14, wherein one or more coding RNAs are involved in DNA replication, Cell cycle, Pyrimidine metabolism, Homologous recombination, p53 signaling pathway, Base excision repair, Nucleotide excision repair, Mismatch repair, One carbon pool by folate, Non-homologous end-joining, Citrate cycle (TCA cycle), Apoptosis, Cellular senescence, MAPK signaling pathway, PI3K-Akt signaling pathway, Jak-STAT signaling pathway, Hematopoietic cell lineage, Oxidative phosphorylation, Fatty acid degradation, Cytokine-cytokine receptor interaction, Ribosome, RNA transport, mRNA surveillance pathway, RNA degradation, Spliceosome, and/ or Purine metabolism.

16. The method of any one of claims 1 to 15, wherein the non-coding RNAs are long non-coding RNAs.

17. The method of any one of claims 1 to 16, wherein the RNA signature includes the expression levels of one or more coding and non-coding RNA gene pairs, and optionally at least 5 or at least 10 coding and non-coding gene pairs.

18. The method of any one of claims 1 to 17, wherein the RNA signature includes the expression levels for at least about 10 RNAs, or at least about 15 RNAs, or at least about 25 RNAs, or at least about 50 RNAs.

19. The method of any one of claims 1 to 18, wherein the candidate cancer therapy is a chemotherapy.
20. The method of claim 19, wherein the chemotherapy comprises cytarabine, paclitaxel, docetaxel, doxorubicin, platinum-based chemotherapeutics that is optionally cisplatin or carboplatin, mitomycin, mithramycin, methotrexate, 5-fluorouracil, vinorelbine, topotecan, irinotecan, bleomycin, actinomycin, topoisomerase I and II inhibitors, anthracycline, epirubicin, idarubicin, mitoxantrone, valrubicin, etoposide, teniposide, and rubitecan.
21. The method of claim 19 or 20, wherein the candidate cancer therapy is a chemotherapy combination.
22. The method of claim 21, wherein the chemotherapy combination is TFAC (paclitaxel, 5-fluorouracil, doxorubicin, and cyclophosphamide), EC (epirubicin and cyclophosphamide), or FEC (5-fluorouracil, epirubicin, and cyclophosphamide).
23. The method of any one of claims 1 to 22, wherein the level of expression of coding and non-coding RNAs are determined by a hybridization assay, RNA sequencing, and/or quantitative PCR.
24. The method of claim 23, wherein the levels of expression of the coding and non-coding RNAs is determined in a biopsy sample from a subject, and the presence of the RNA signature determined.
25. The method of claim 24, wherein the levels of expression of the coding and non-coding RNAs in the biopsy sample are determined using quantitative PCR, RNA sequencing, and/or hybridization assay.
26. The method of claim 25, wherein the levels of expression of the coding and non-coding RNAs are determined in the cell lines or primary cell cultures using a hybridization assay or RNA sequencing, and the levels of expression of the coding and non-coding RNAs are determined in the biopsy sample using quantitative PCR.
27. The method of any one of claims 1 to 26, wherein the cell or tissue is classified using one or more classification schemes selected from Correlation Analysis, Principal Components Analysis, Naïve Bayes, Support Vector Machines, Nearest Neighbors, Decision Trees, Logistic, Artificial Neural Networks, and Rule-based schemes.

28. A method for treating cancer in a subject in need thereof, comprising
determining an RNA signature indicative of sensitivity or resistance to a candidate cancer therapy according to the method of any one of claims 1 to 27; and
administering to the subject a cancer therapy that the cancer cell or tissue is classified as sensitive to.
29. The method of claim 28, wherein the cancer therapy is a chemotherapy.
30. The method of claim 29, wherein the chemotherapy comprises cytarabine, paclitaxel, docetaxel, doxorubicin, platinum-based chemotherapeutics that is optionally cisplatin or carboplatin, mitomycin, mithramycin, methotrexate, 5-fluorouracil, vinorelbine, topotecan, irinotecan, bleomycin, actinomycin, topoisomerase I and II inhibitors, anthracycline, epirubicin, idarubicin, mitoxantrone, valrubicin, etoposide, teniposide, and rubitecan.
31. The method of claim 29 or claim 30, wherein the candidate cancer therapy is a chemotherapy combination.
32. The method of claim 31, wherein the chemotherapy combination is TFAC (paclitaxel, 5-fluorouracil, doxorubicin, and cyclophosphamide), EC (epirubicin and cyclophosphamide), or FEC (5-fluorouracil, epirubicin, and cyclophosphamide).
33. A composition for use in the method of any one of claims 28 to 32.

FIGURE 1

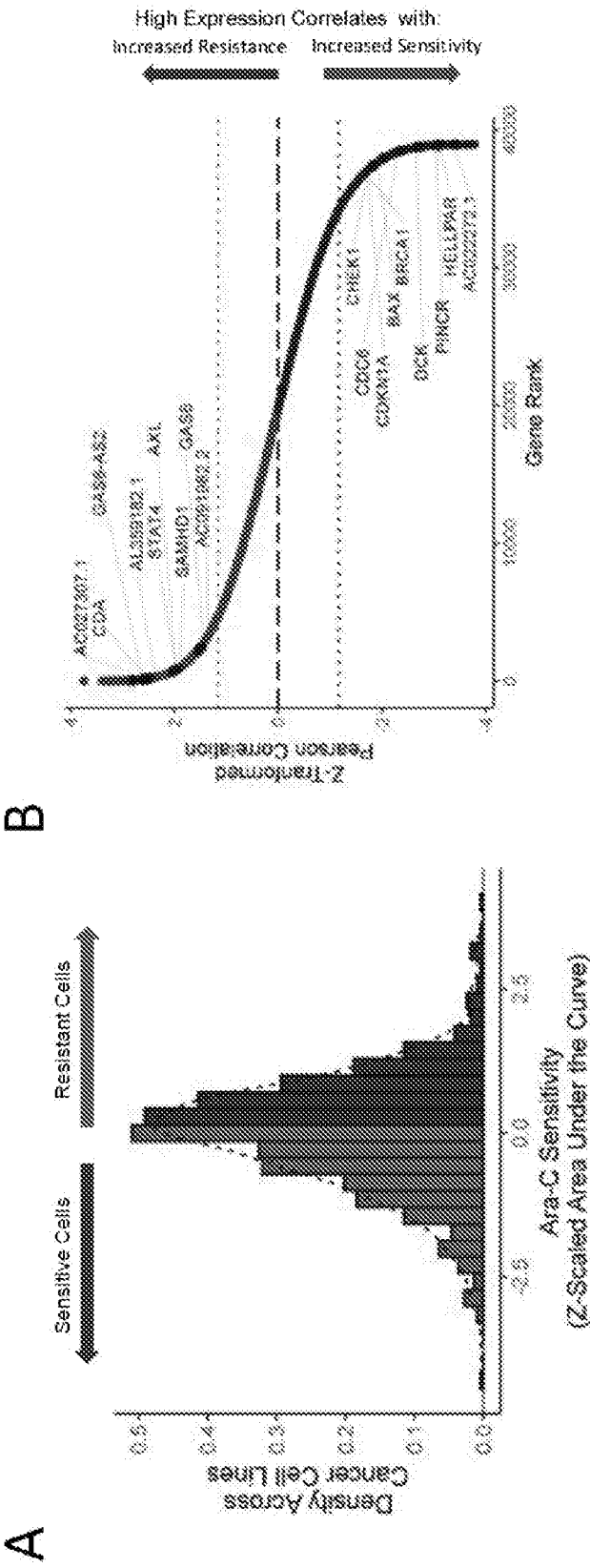


FIGURE 1 CONT.

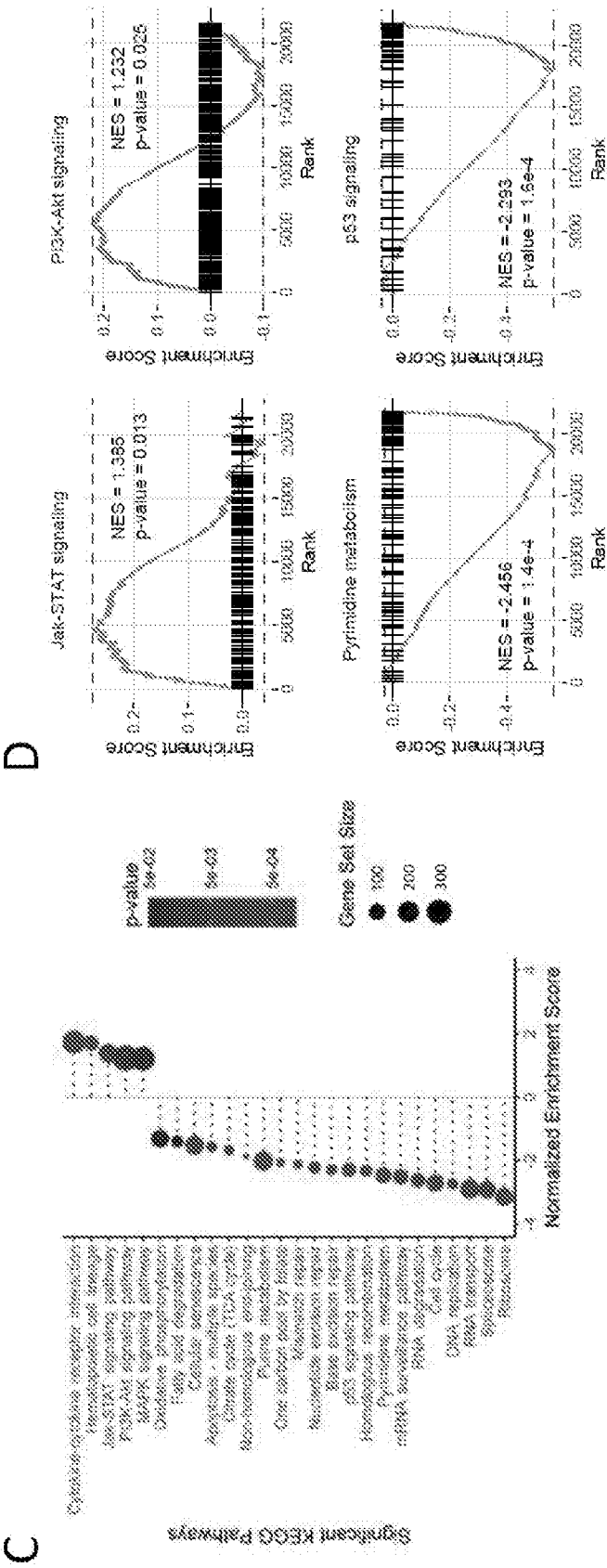


FIGURE 1 CONT.

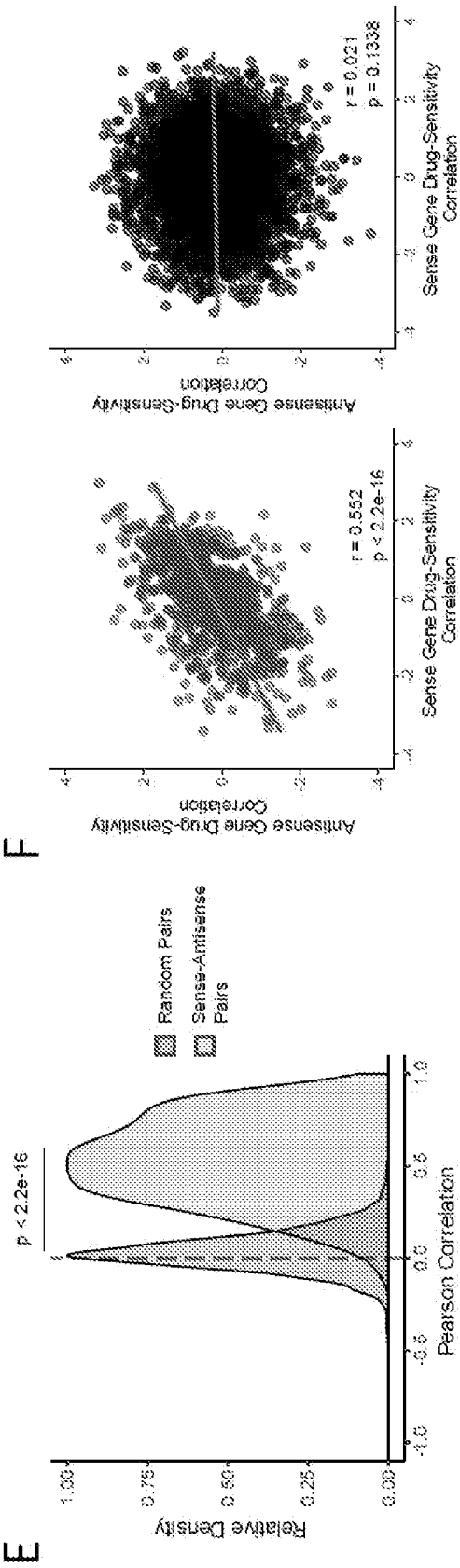


FIGURE 2

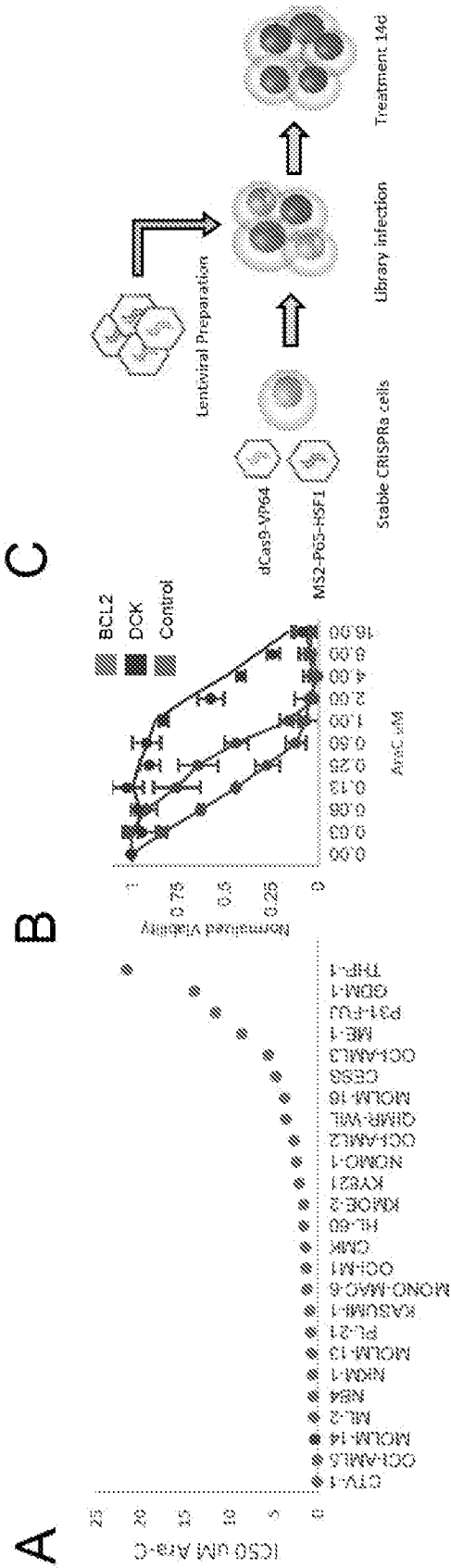
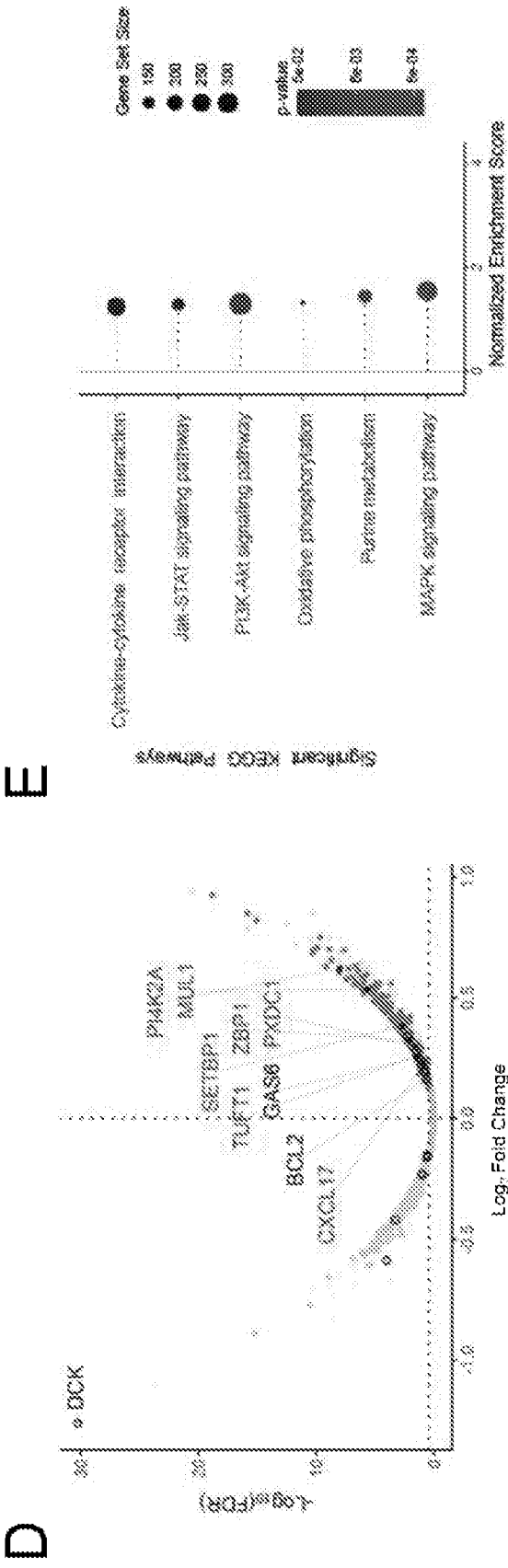


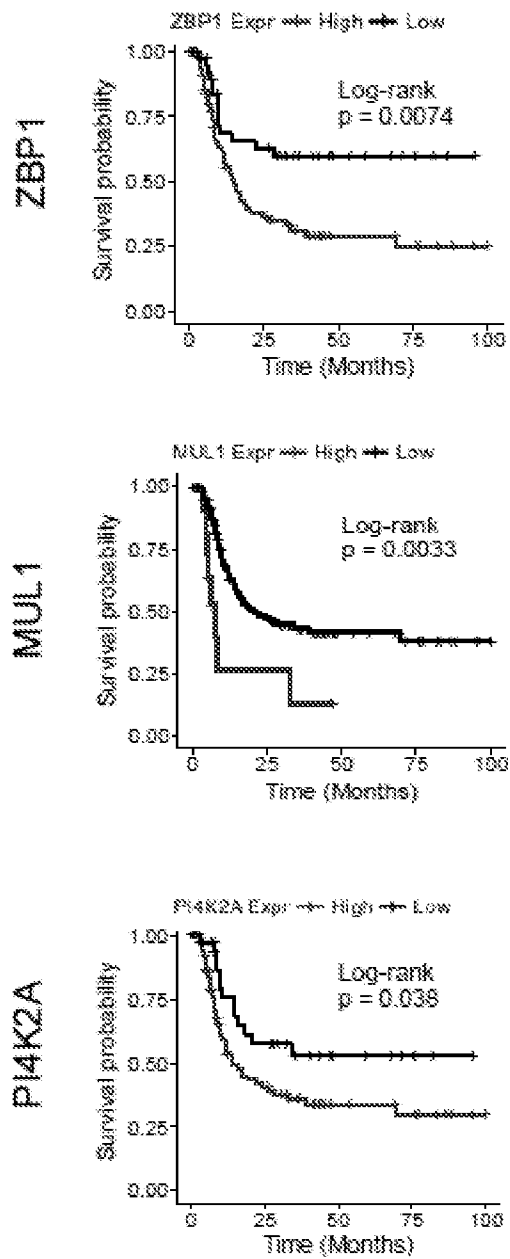
FIGURE 2 CONT.



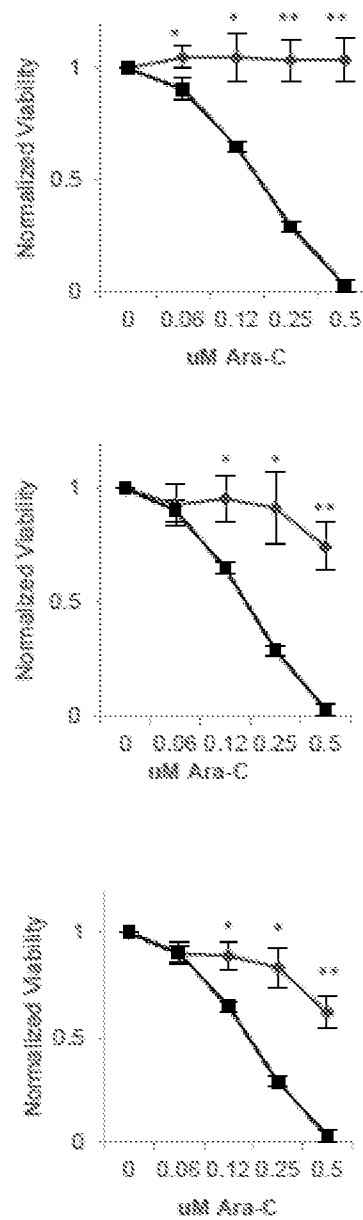
6/27

FIGURE 2 CONT.

F



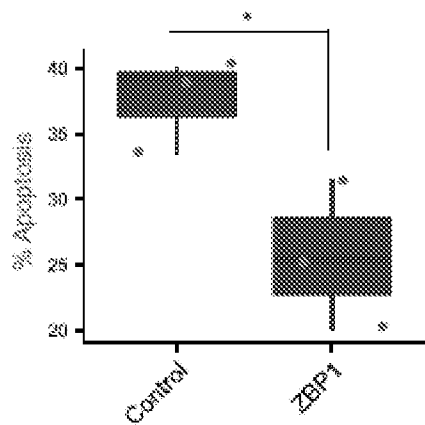
G



7/27

FIGURE 2 CONT.

H



I

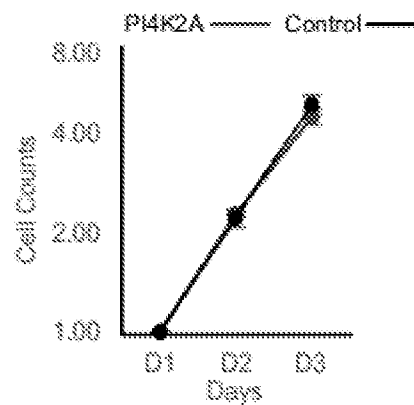
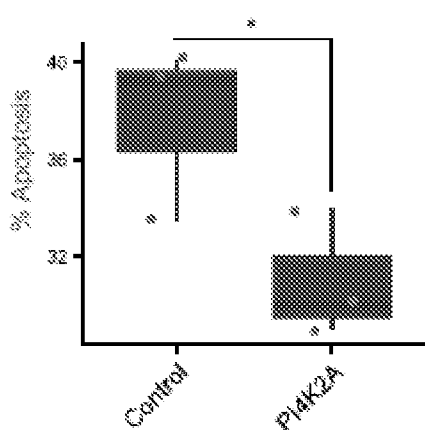
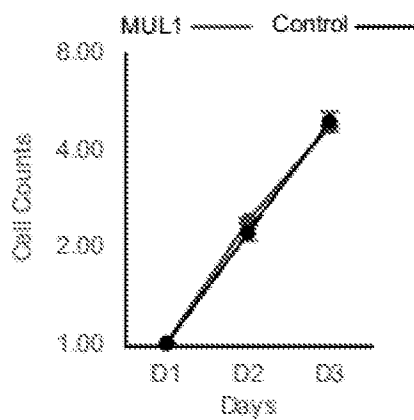
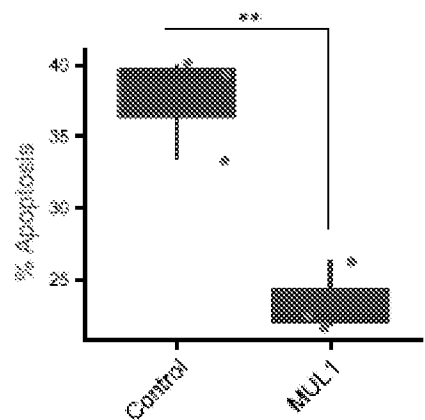
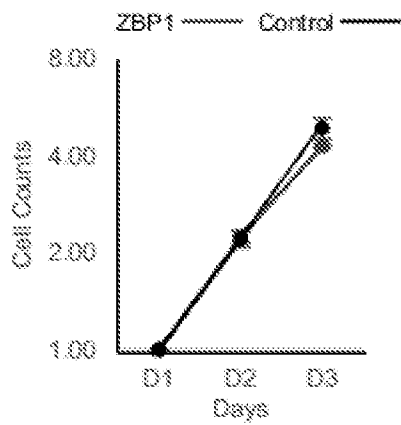
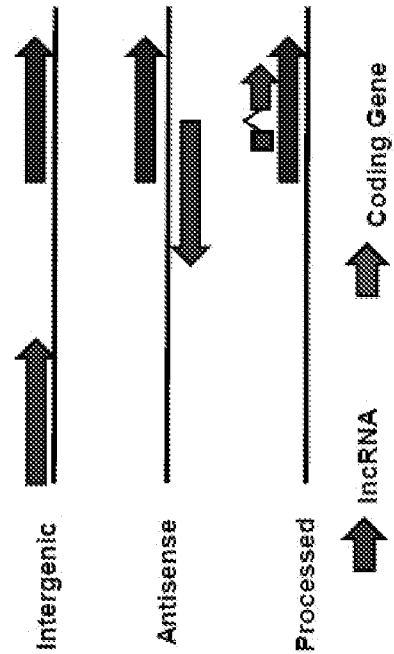


FIGURE 3

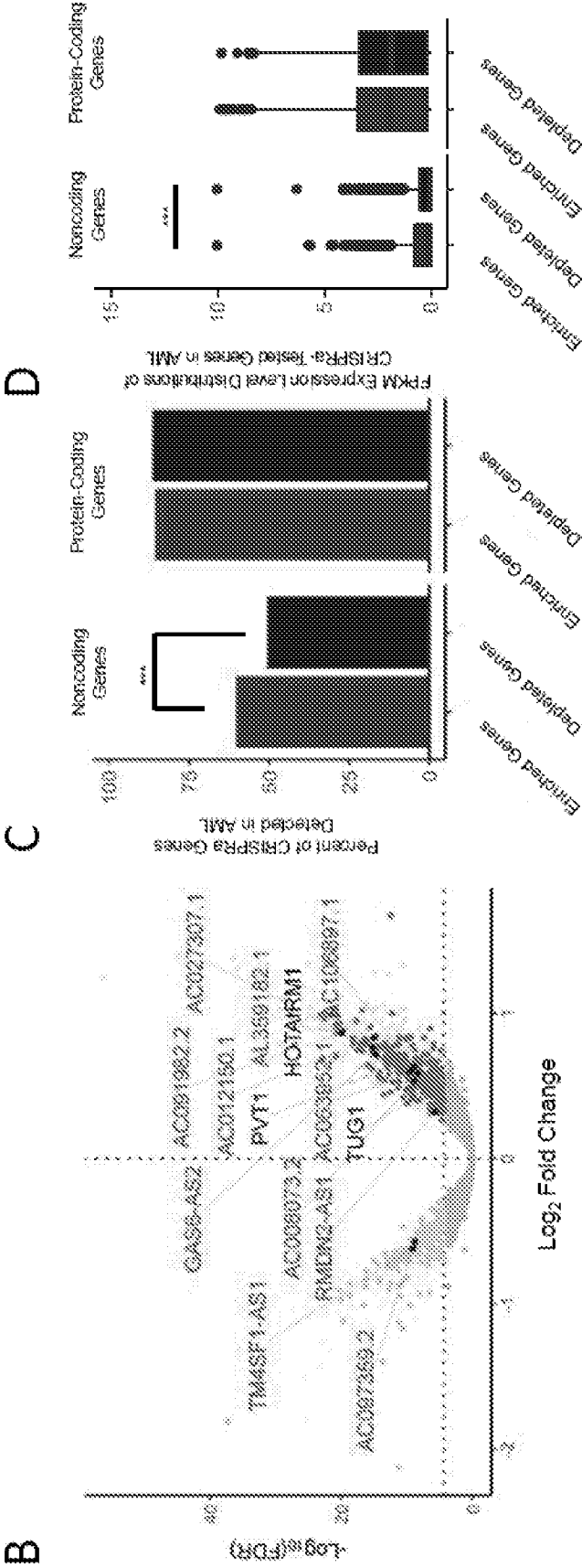
A

Sub-library	Genes	TSS	sgRNAs
Human Intergenic lincRNA	7,563	10,781	42,836
Human Antisense	5,532	8,545	33,965
Human Processed Transcript	1,606	2,927	11,643
Non-Targeting Control			99
Total	14,701	22,253	88,543



9/27

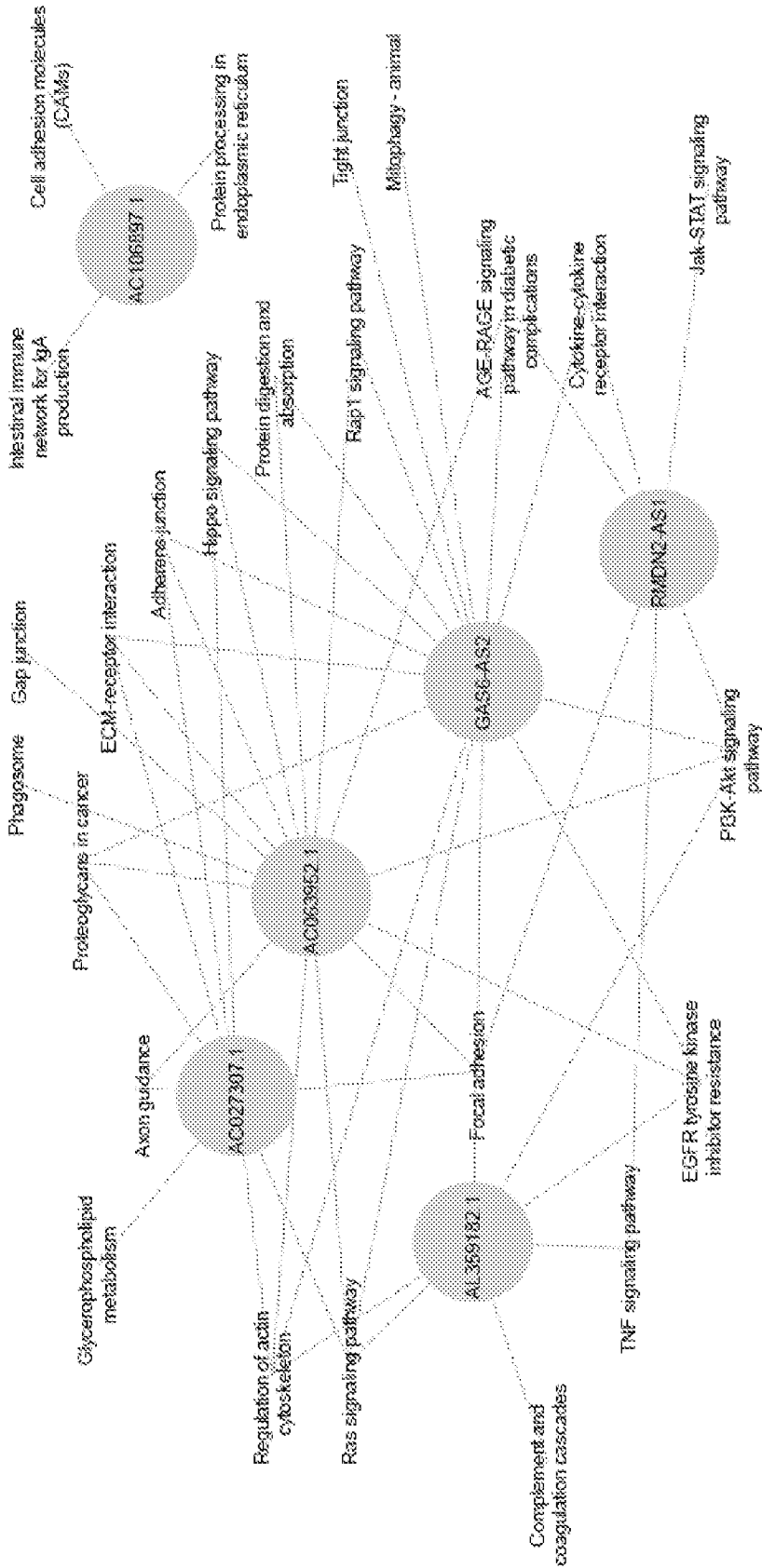
FIGURE 3 CONT.



10/27

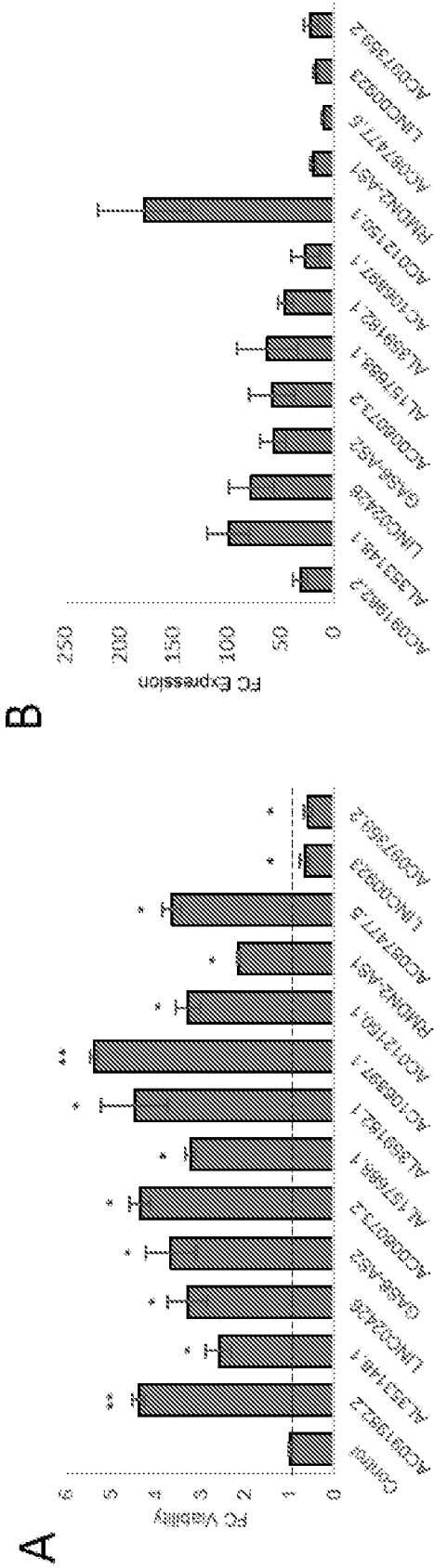
FIGURE 3 CONT.

E



11/27

FIGURE 4



12/27

FIGURE 4 CONT.

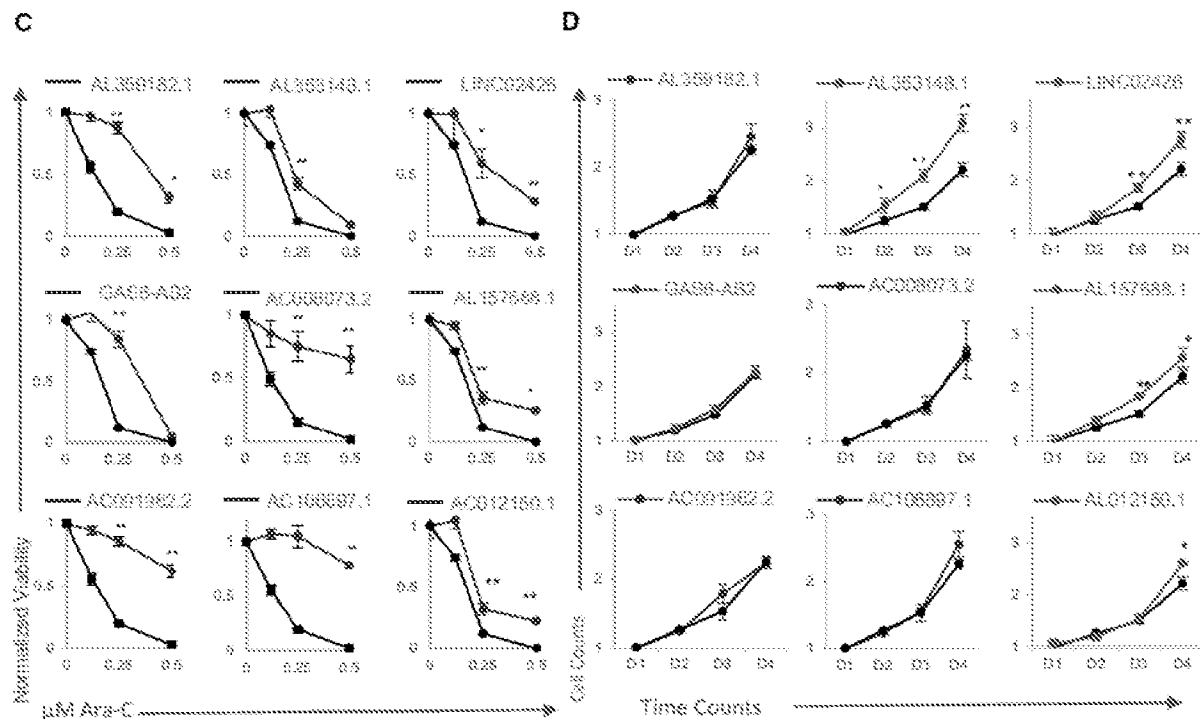


FIGURE 4 CONT.

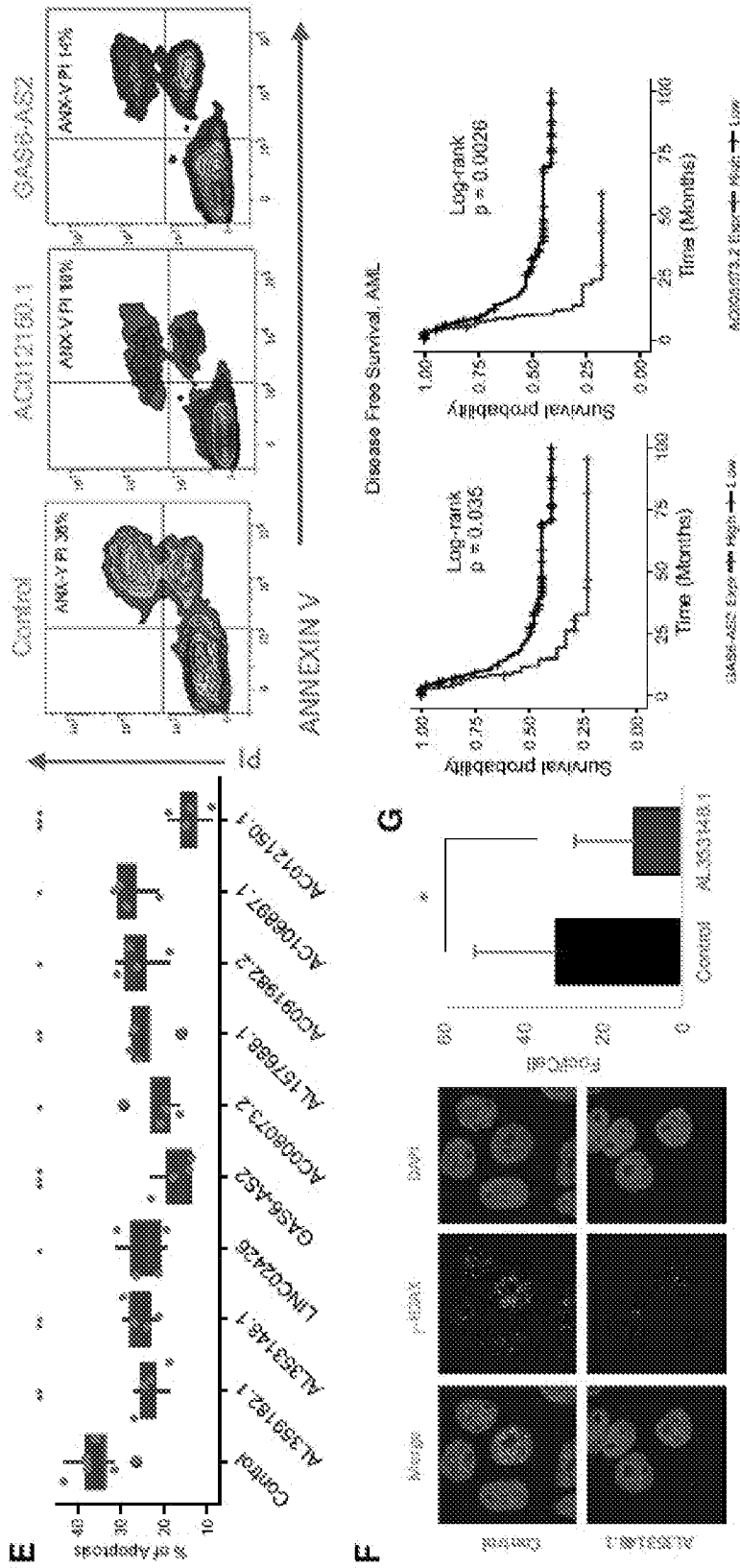


FIGURE 5

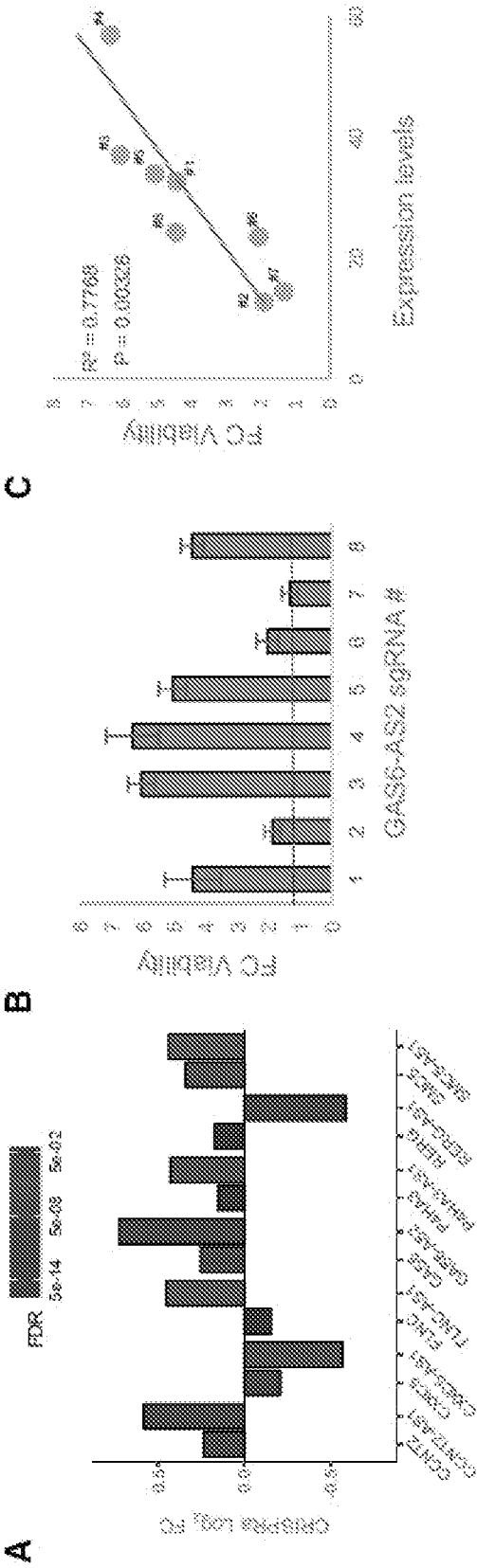


FIGURE 5 CONT.

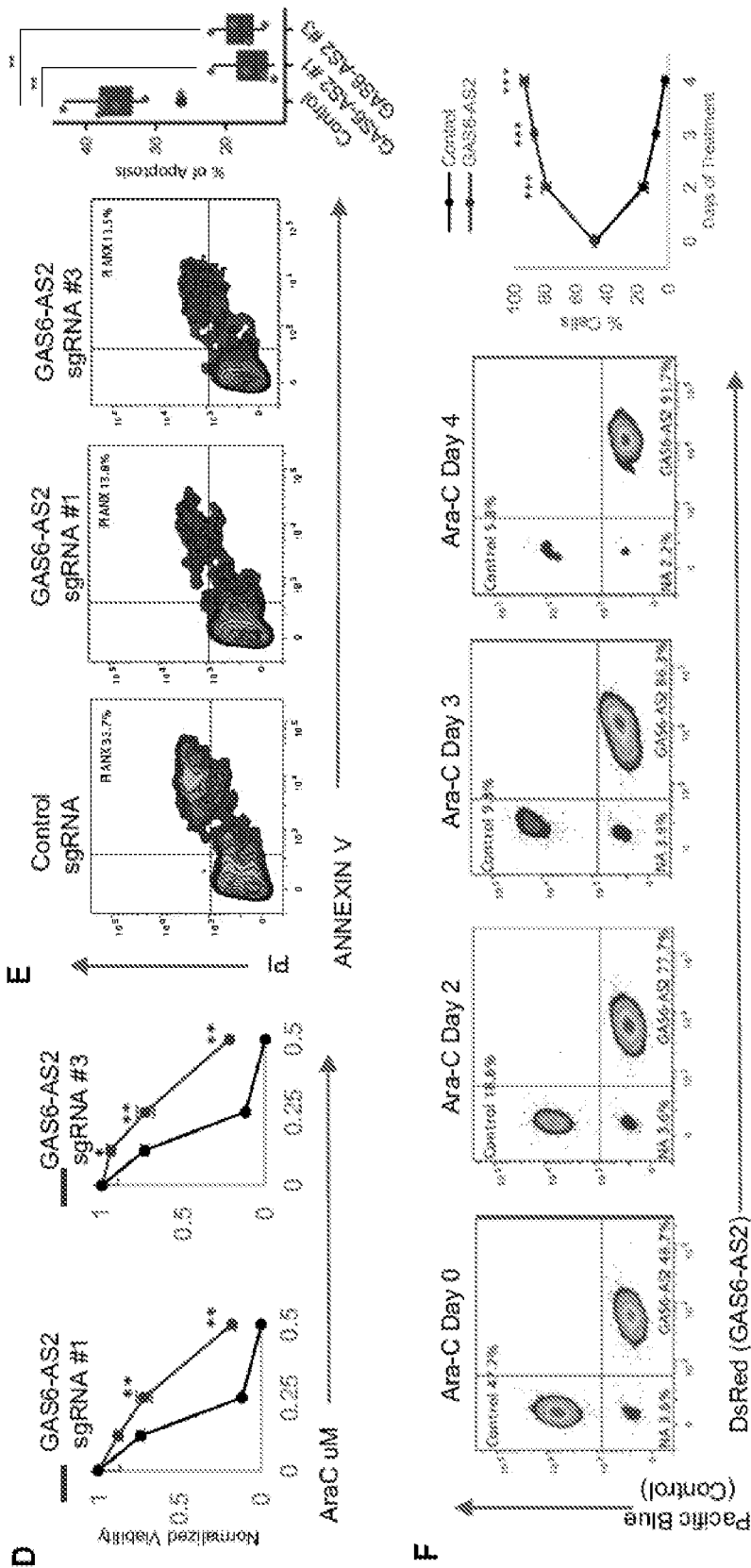


FIGURE 5 CONT.

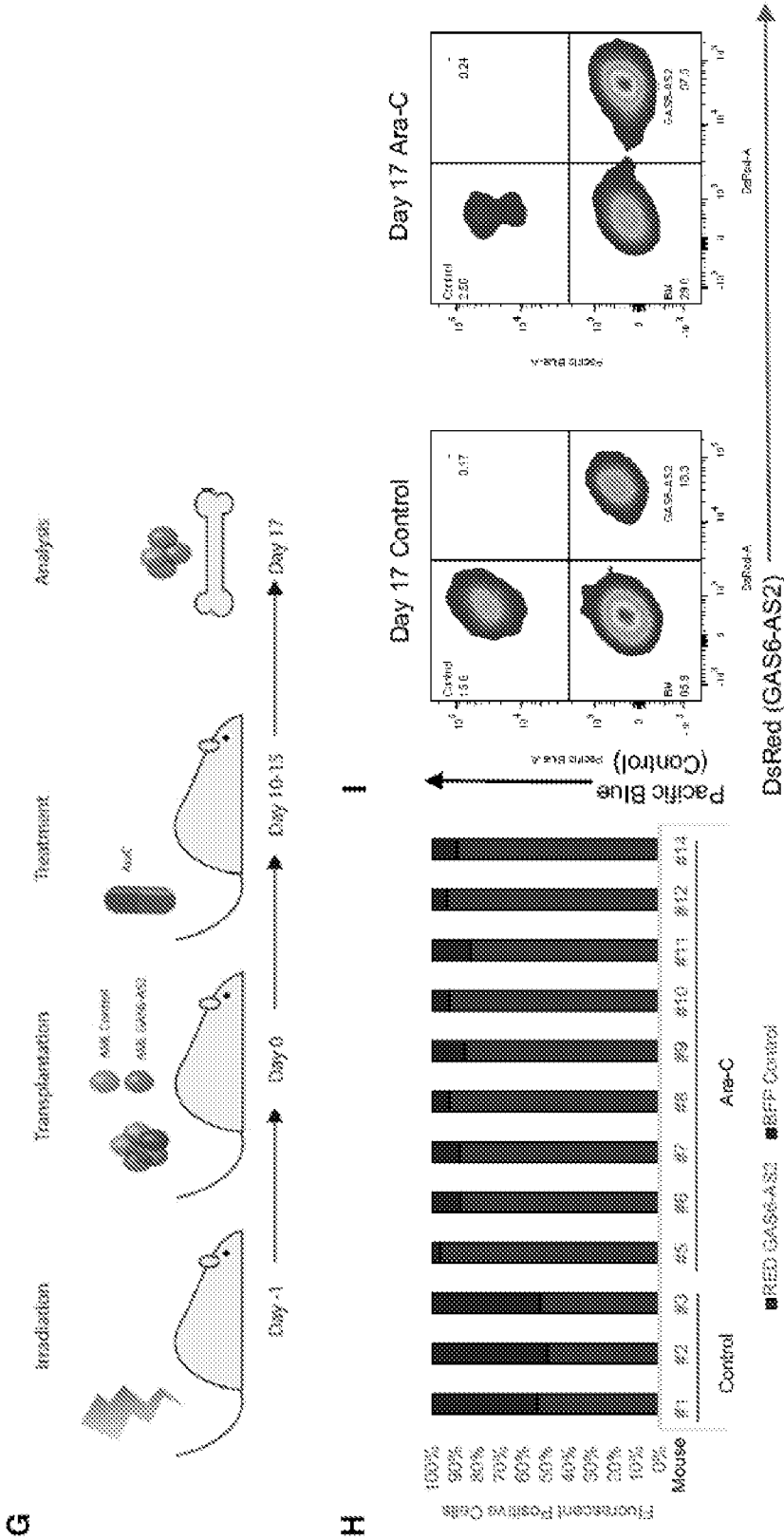
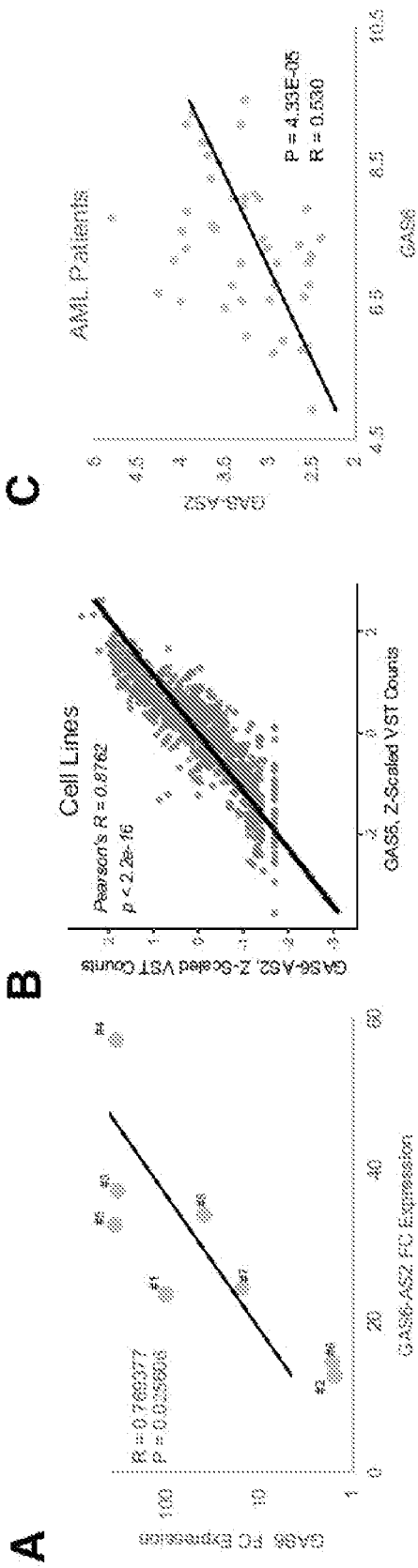


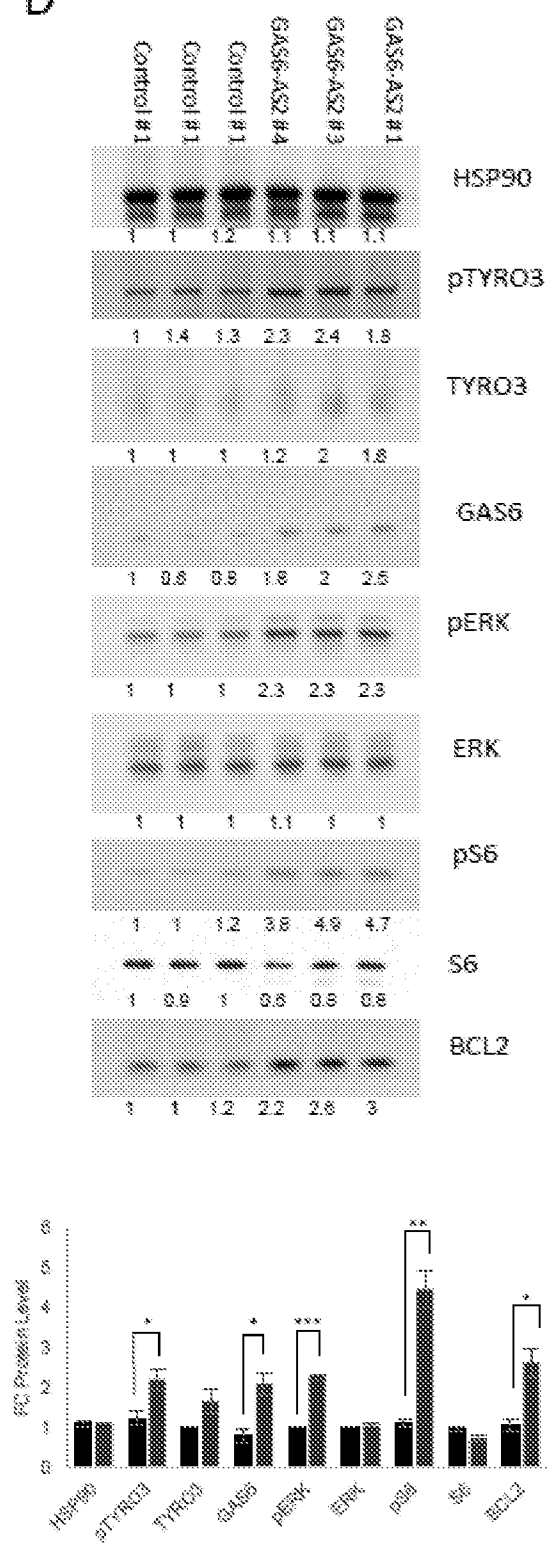
FIGURE 6



18/27

FIGURE 6 CONT.

D



19/27

FIGURE 6 CONT.

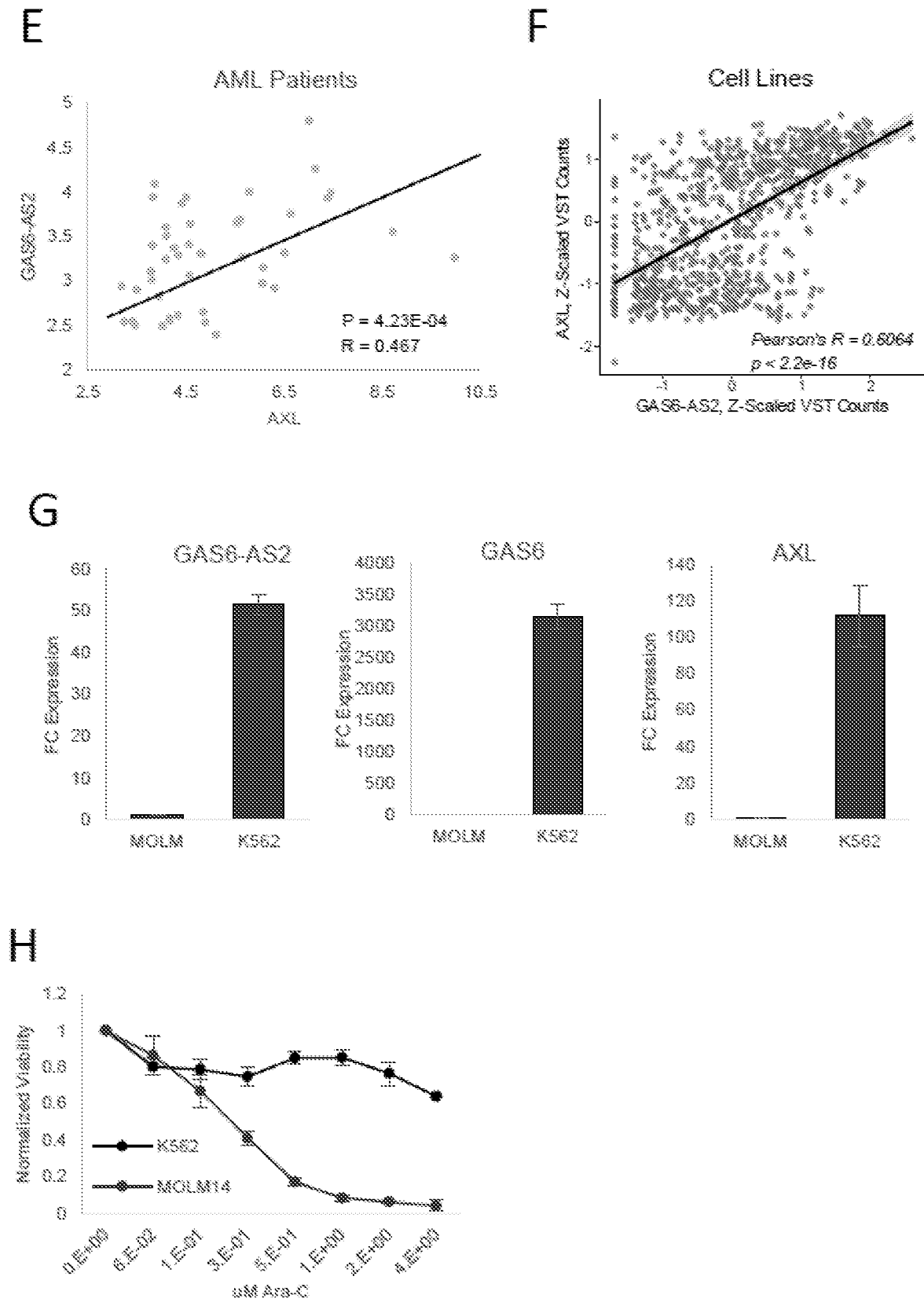


FIGURE 7

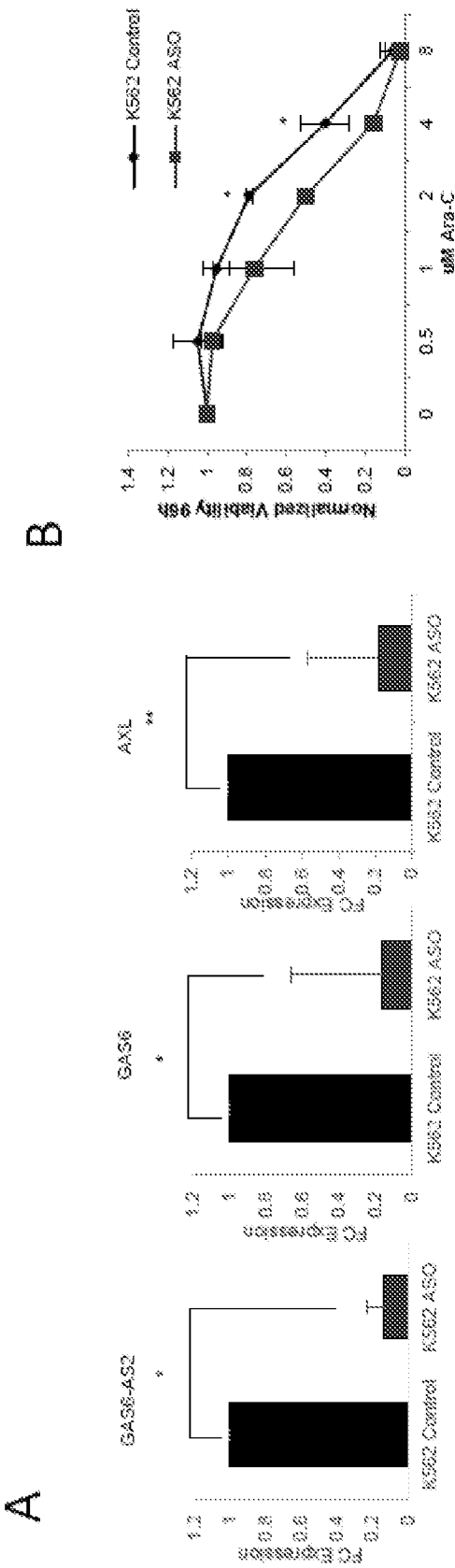
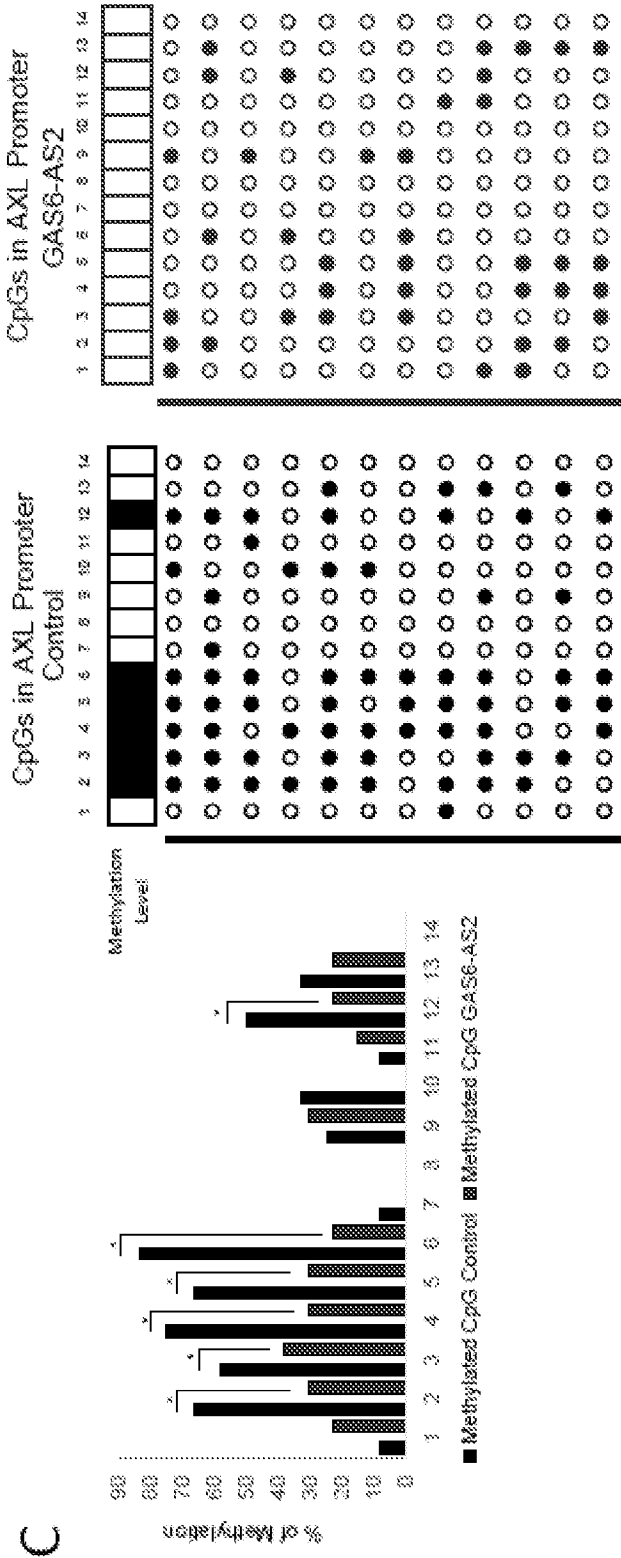


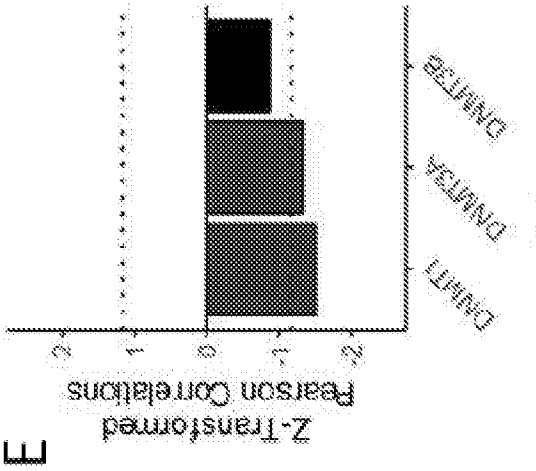
FIGURE 7 CONT.



D

Name	Source	P-Value	FDR
Human leukemia	GeneSigDB	2.43E-40	5.95E-37
Wnt signaling pathway	PantherDB	6.36E-06	3.66E-04
Solute carriers	genenames.org	1.19E-17	4.88E-15
Genes with high-CpG-density promoters (HCP)	MSigDB C2: CGP Curated Gene Sets (v5.1)	4.70E-44	4.61E-40

FIGURE 7 CONT.



F

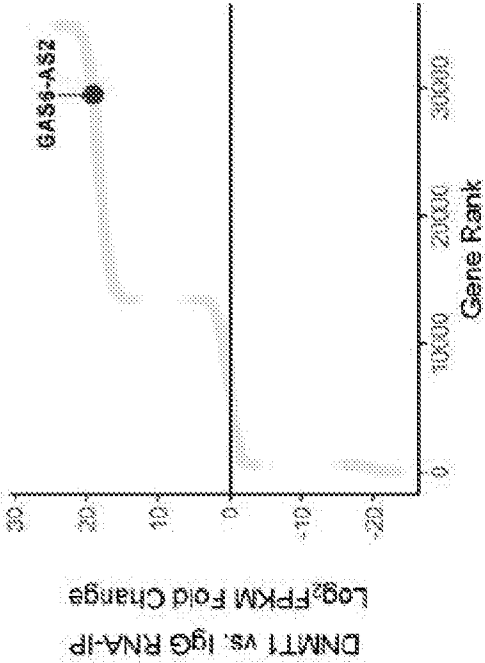
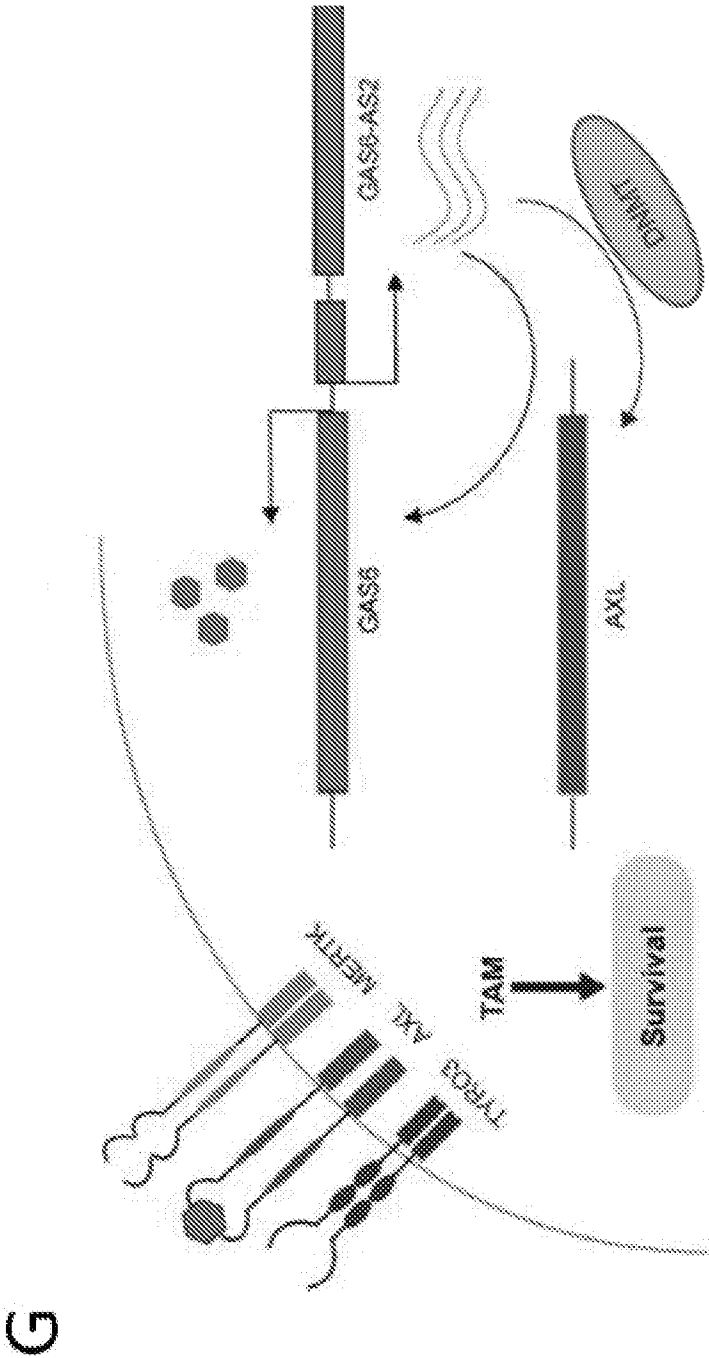


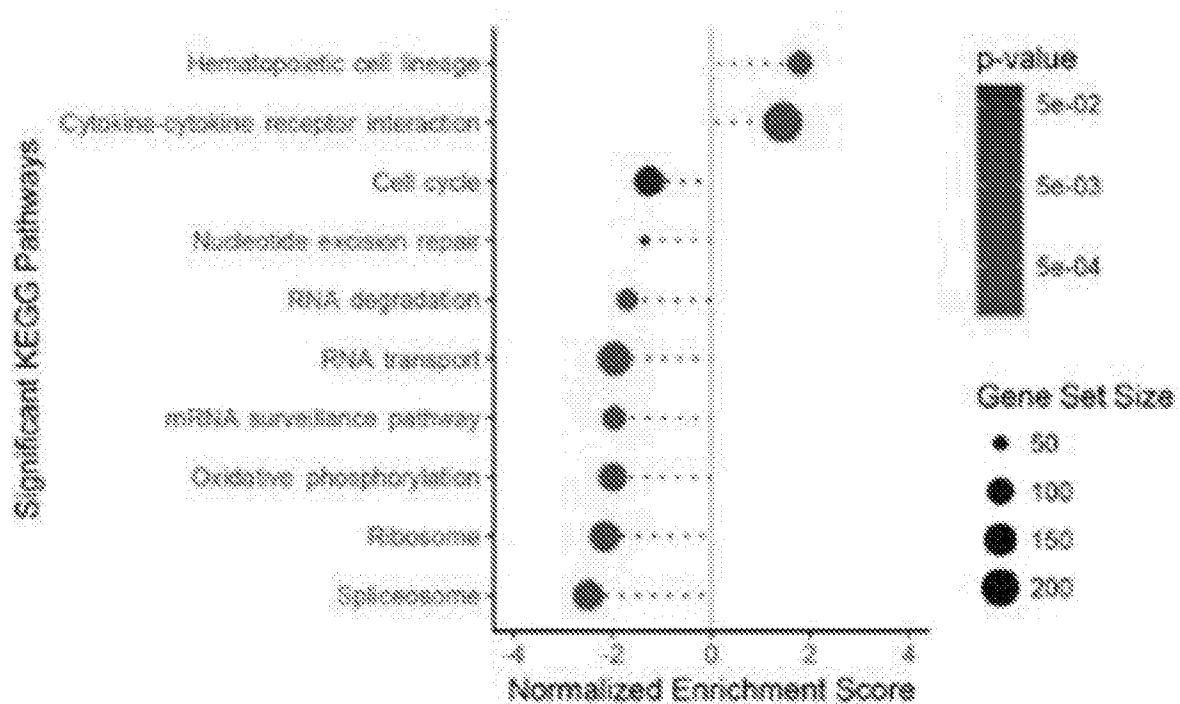
FIGURE 7 CONT.



24/27

FIGURE 8

8A



8B

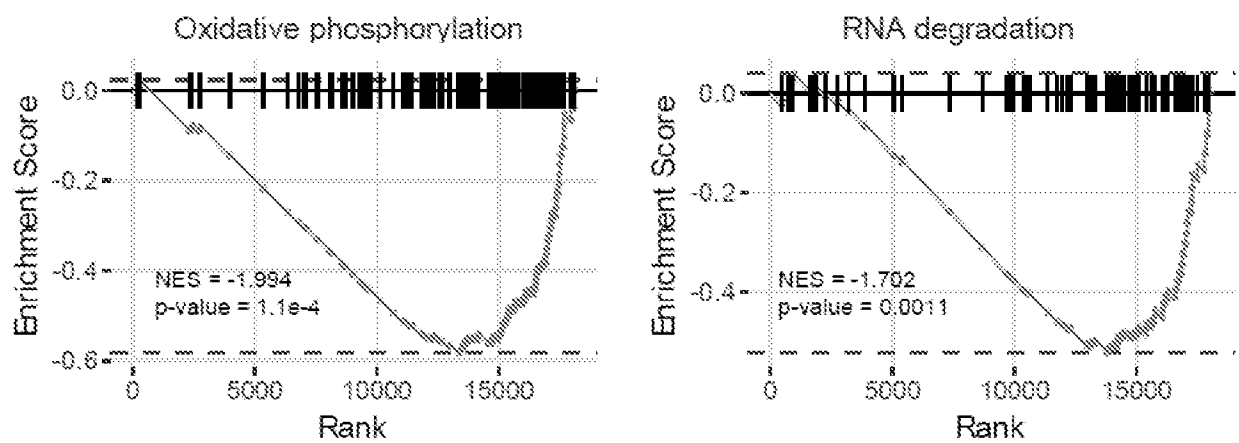


FIGURE 9

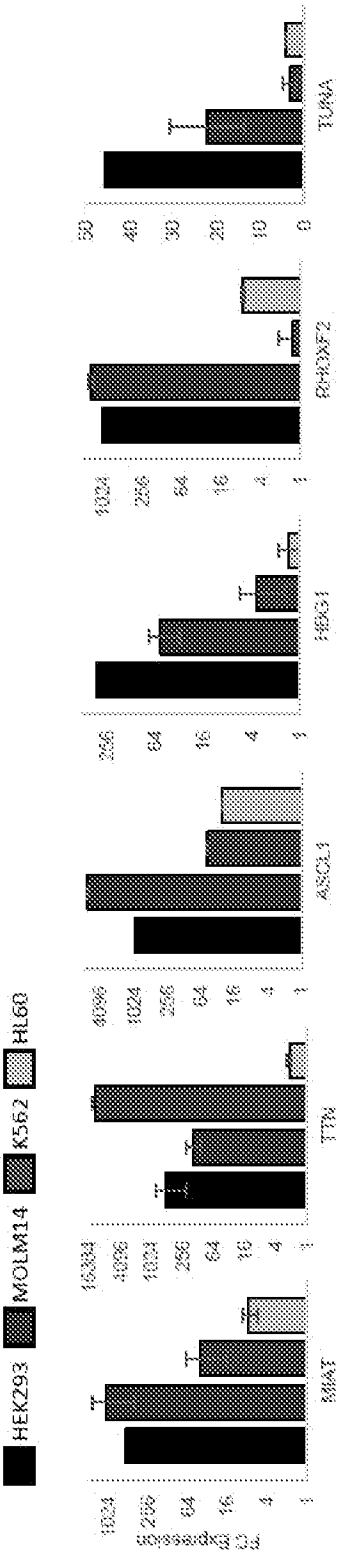
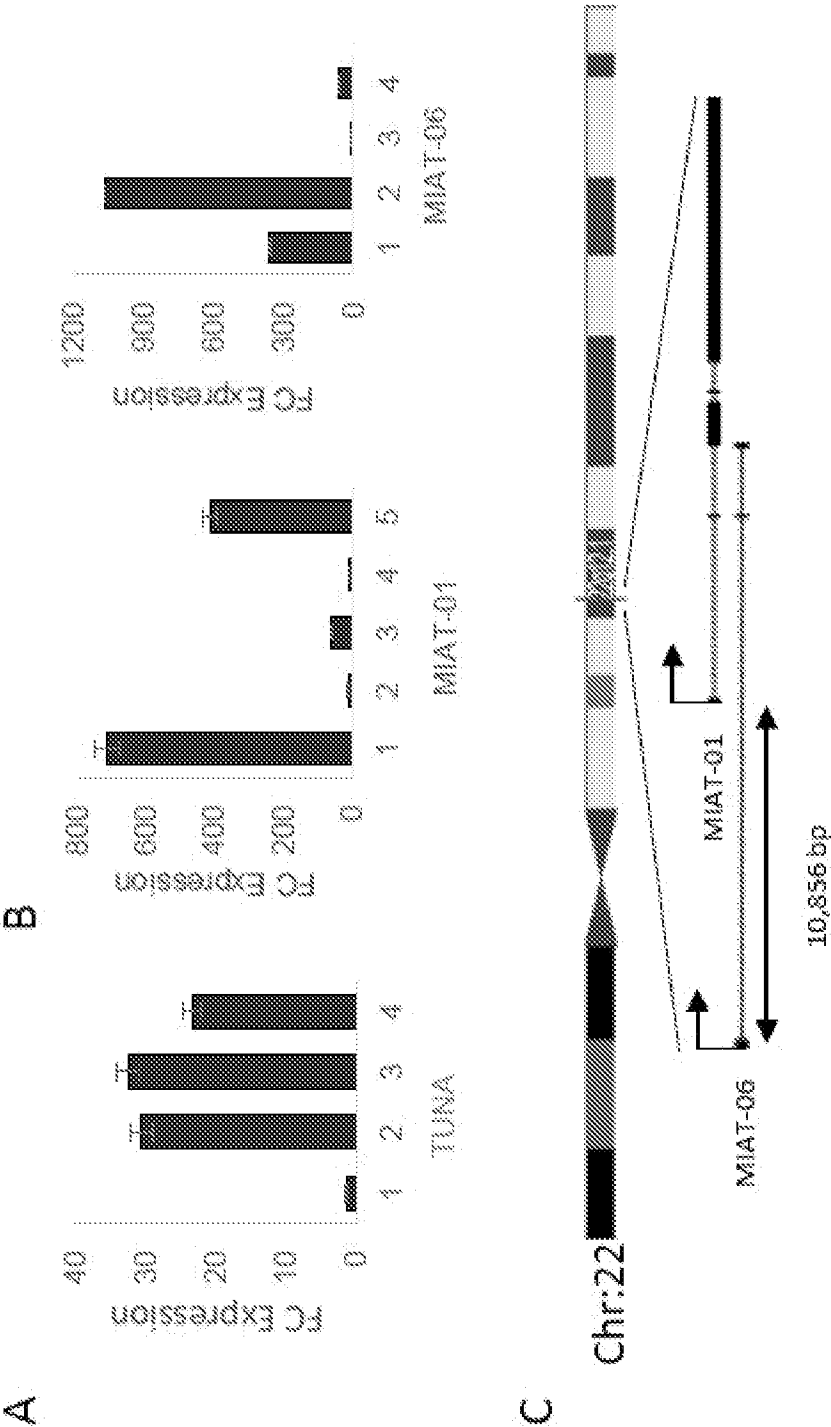
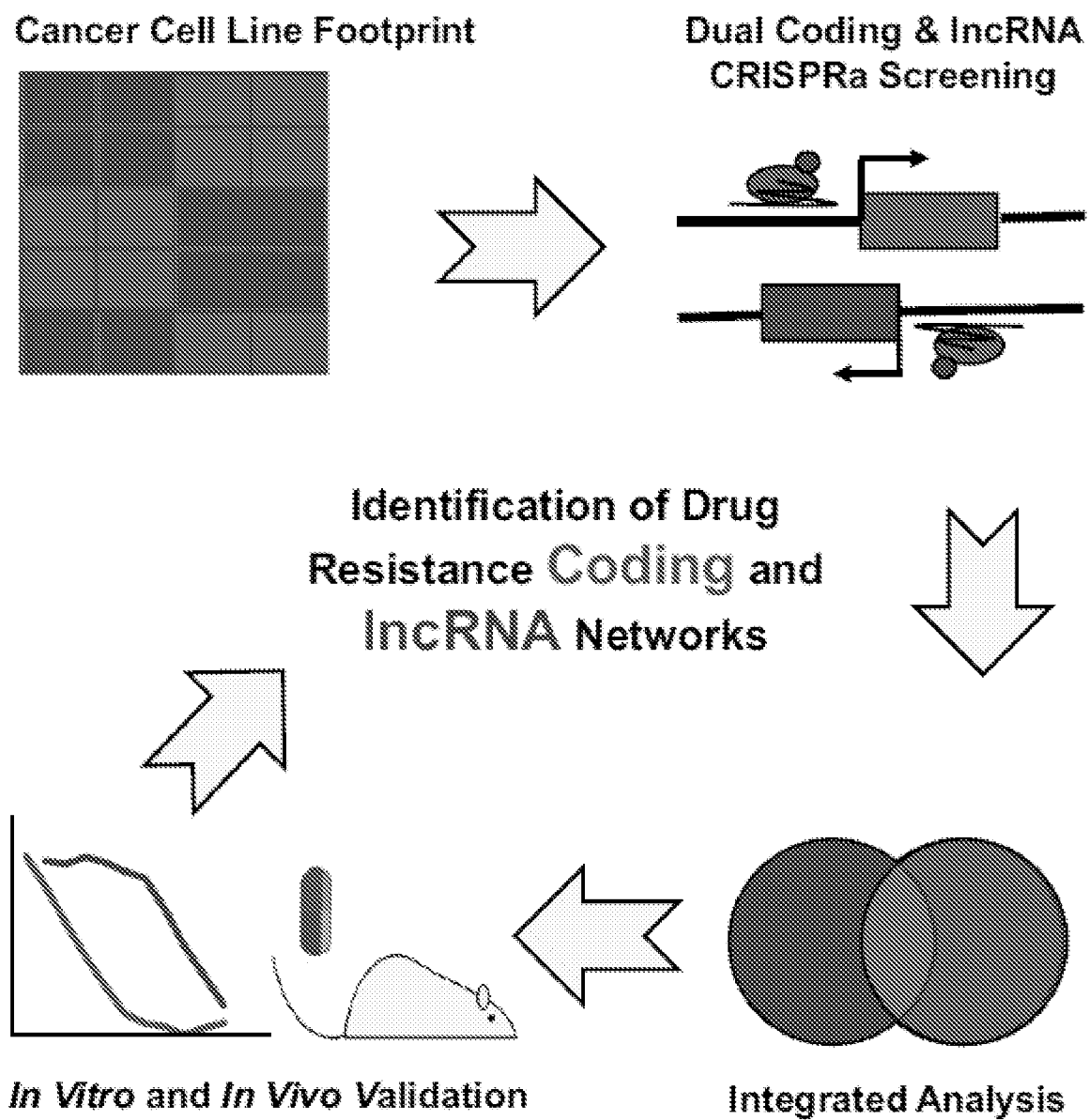


FIGURE 10



27/27

FIGURE 11



INTERNATIONAL SEARCH REPORT

International application No.

PCT/US19/19413

Box No. 1 Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

- a. ☒ forming part of the international application as filed: .
☒ in the form of an Annex C/ST.25 text file.
☐ on paper or in the form of an image file.
- b. ☐ furnished together with the international application under PCT Rule 13*ter*. 1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
- c. ☐ furnished subsequent to the international filing date for the purposes of international search only:
☐ in the form of an Annex C/ST.25 text file (Rule 13*ter*. 1(a)).
☐ on paper or in the form of an image file (Rule 13*ter*. 1(b) and Administrative Instructions, Section 713).

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US19/19413

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: 13-33
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US19/19413

A. CLASSIFICATION OF SUBJECT MATTER

IPC - C12Q 1/6886; G01N 33/574; C12P 19/34 (2019.01)

CPC - C12Q 1/6886; G01N 33/574; C12P 19/34

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History document

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

See Search History document

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

See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	WO 2012/167278 A1 (ALMAC DIAGNOSTICS LIMITED) 06 December, 2012; abstract; paragraphs [0020], [0043], [0045], [0052]-[0053], [0068], [0113]; claims 1, 3-8, 50-53	1-6, 8-10, 12/1-6, 12/8-10 ----- 7, 11/1-10, 12/7
Y	(KUNER, R) Lung Cancer Gene Signatures and Clinical Perspectives. Microarrays. 13 December, 2013; Vol. 2; pages 318-339; page 324, paragraph 2; page 326, paragraph 2; DOI: 10.3390/microarrays2040318	7, 11/1-10, 12/7
A	WO 2009/158143 A1 (THE UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL, et al.) 30 December, 2009; whole document	1-10, 11/1-10, 12/1-10

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search

09 April 2019 (09.04.2019)

Date of mailing of the international search report

23 APR 2019

Name and mailing address of the ISA/

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Shane Thomas

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774