

Single Cell Profiling Distinguishes Leukemia-Selective Chemotypes

Hannah L. Thirman^{1,2,3,4}, Madeline J. Hayes^{1,2,3}, Lauren E. Brown⁵, John A. Porco, Jr⁵, Jonathan M. Irish^{1,2,3**}

¹ Department of Cell and Developmental Biology, Vanderbilt University, Nashville, TN, USA.

² Department of Pathology, Microbiology and Immunology, Vanderbilt University Medical Center, Nashville, TN, USA.

³ Vanderbilt Center for Immunobiology, Vanderbilt University Medical Center, Nashville, TN, USA.

⁴ Chemical & Physical Biology Program, Vanderbilt University, Nashville, TN, USA.

⁵ Department of Chemistry and Center for Molecular Discovery (BU-CMD), Boston University, Boston, MA, USA.

**Corresponding author

Email: jonathan.irish@vanderbilt.edu (J.M.I.)

Running Title: Cytometric profiling identifies rocaglate pyrimidinones with exceptional bioactivity

Keywords: Rocaglates, chemical biology, phospho-flow, cytometry, cell signaling, cancer

Abstract

A central challenge in chemical biology is to distinguish molecular families in which small structural changes trigger large changes in cell biology. Such families might be ideal scaffolds for developing cell-selective chemical effectors – for example, molecules that activate DNA damage responses in malignant cells while sparing healthy cells. Across closely related structural variants, subtle structural changes have the potential to result in contrasting bioactivity patterns across different cell types. Here, we tested a 600-compound Diversity Set of screening molecules from the Boston University Center for Molecular Discovery (BU-CMD) in a novel phospho-flow assay that tracked fundamental cell biological processes, including DNA damage response, apoptosis, M-phase cell cycle, and protein synthesis in MV411 leukemia cells. Among the chemotypes screened, synthetic congeners of the rocaglate family were especially bioactive. In follow-up studies, 37 rocaglates were selected and deeply characterized using 12 million additional cellular measurements across MV411 leukemia cells and healthy peripheral blood mononuclear cells. Of the selected rocaglates, 92% displayed significant bioactivity in human cells, and 65% selectively induced DNA damage responses in leukemia and not healthy human blood cells. Furthermore, the signaling and cell-type selectivity were connected to structural features of rocaglate subfamilies. In particular, three rocaglates from the rocaglate pyrimidinone (RP) structural subclass were the only molecules that activated exceptional DNA damage responses in leukemia cells without activating a detectable DNA damage response in healthy cells. These results indicate that the RP subset should be extensively characterized for anticancer therapeutic potential as it relates to the DNA damage response. This single cell profiling approach advances a chemical biology platform to dissect how systematic variations in chemical structure can profoundly and differentially impact basic functions of healthy and diseased cells.

Introduction

Improving the quality of per-cell biological measurements has been a focus of cytometry over the last 20 years and has led to significant improvements in understanding cellular processes governing life, death, and specification of cell identity¹⁻⁴. However, in preclinical drug discovery and chemical biology it is common to use either well-based readouts, or alternatively, to deploy imaging or label free approaches that are neither single cell nor multiplexed, or lack in resolution or throughput. Suspension cytometry assays could potentially address these issues by balancing these factors while quantifying multiple features per cell^{5,6}.

Phospho-specific flow cytometry (phospho-flow) was developed to quantify a range of cellular functions at the single cell level both within and across different populations of cells in parallel⁷⁻⁹. The technique involves quantifying one or more features of individual cells that have been stained with a panel of fluorescently tagged antibodies directed against specific epitopes, such as surface proteins, transcription factors, and phospho-proteins. This approach can therefore assess multiple signaling pathways simultaneously in mixtures of cells without the need to physically isolate different cell subsets present in human tissue samples. Thus, phospho-flow offers a platform for chemical biology that combines clinical relevance and multi-parameter single cell data. Throughput, consistency, and costs have been improved by a technique called fluorescent cell barcoding (FCB) that labels cells from a given well or sample with a unique signature of fluorescent dye levels - a barcode - so that they can be mixed, stained, and analyzed as a single sample¹⁰⁻¹². Phospho-flow and FCB protocols have evolved over the last two decades to closely align with drug discovery and translational research workflows¹³⁻¹⁵. Phospho-flow-based single cell profiling has been used to screen sets of small molecules and metabolite extracts and has identified pathway- and cell-selective inhibitors in cell lines and primary samples^{3,6,16-19}. These prior studies focused on sets of diverse molecules; thus, there is an opportunity to explore multiplexed phospho-flow as a platform for dissecting structure-activity relationships (SAR).

Here, we aimed to develop bench and computational methods to apply single cell profiling to SAR studies. This approach aims to draw on quantitative structure-activity relationship (QSAR) approaches²⁰ and bring in single cell bench techniques and modern data science tools that draw on machine learning. Computational methods like t-distributed stochastic neighbor embedding (t-SNE)²¹⁻²³, UMAP^{24,25}, and PHATE²⁶ are used to analyze cytometry data²⁴ and are now growing in use in computational chemical biology^{6,16,26}. Recently developed algorithms, such as Tracking Responders EXpanding (T-REX) and Marker Enrichment Modeling (MEM), might be used to augment analysis and facilitate identification of rare cells and quantification of contextual protein enrichment^{27,28}. Tools like these that can resolve cellular heterogeneity and consider many readouts simultaneously may help to reveal connections between chemical structure and cellular bioactivity and accelerate development of molecules that selectively modulate chosen cell subsets and improve patient outcomes.

Here, we describe a multiplexed, single cell phospho-flow-based test of a chemical diversity library (**Diversity Set**) that identified a natural product chemotype that was next tested extensively for structure-activity relationships using single cell profiles of fundamental cellular processes (**Figure 1**). Here, the term ‘chemotype’ will be used to describe a set of molecules that share a molecular scaffold²⁹⁻³¹. The chemical family whose members most often displayed exceptional anti-leukemia bioactivity in initial tests with the **Diversity Set** was the rocaglates (flavaglines), a family of plant-derived translation inhibitors whose founding member, rocaglamide A (RocA), was identified for its anticancer activity³². The Boston University Center for Molecular Discovery (BU-CMD) houses a series of synthetic rocaglate structural variants³³ which can be informally categorized into three subclasses: 1) regular rocaglates (RR), exemplified by natural products RocA (**Supplementary Figure 1A**), silvestrol, and eliptifoline, containing a cyclopenta[*b*]-benzofuran scaffold and no other ring fusions, 2) rocaglate pyrimidinones (RPs), exemplified by natural products aglaroxin C (**Supplementary Figure 1B**) and aglaiastatin, containing a mono- or bicyclic pyrimidinone ring system fused to the cyclopenta[*b*]-benzofuran core,^{34,35} and 3) amidino-rocaglates (ADRs), a novel rocaglate subclass not found in nature, which contains an amidine ring fusion to the cyclopenta[*b*]benzofuran core and has been shown to exhibit the most potent translation inhibition for any rocaglate to-date^{36,37}. Following selection of rocaglates as a class for in-depth study, a representative **Rocaglate Set** of members, both natural and synthetic, of these rocaglate subclasses was assembled and deeply characterized for structure-activity relationships using a high dimensional single cell profiling panel.

Ultimately, this work reveals subclasses and individual molecules with desirable and reproducible bioactivity in leukemia cells and not in healthy human blood cells. Additionally, this study shows the value of using single cell profiling to characterize a large network of fundamental cell signaling processes, as striking differences in bioactivity and cell type selectivity were observed between different rocaglate classes. In particular, rocaglate pyrimidinones (RPs) stand out as a subgroup that exclusively targeted leukemia cells and not healthy cells, with individual members displaying exceptional ability to activate intrinsic DNA damage responses that lead to the selective destruction of malignant cells. These exceptional RPs should now be prioritized in SAR and pre-clinical translational studies.

Results

Rocaglates are highly bioactive in leukemia cells.

An initial test of 600 diverse molecules from the BU-CMD compound collection, with representation from 120 distinct chemotypes (**Diversity Set**), was conducted at 10 μ M in MV411 leukemia cells using phospho-flow cytometry in combination with FCB for multiplexing. The goal of this experiment was to identify chemotypes with exceptional leukemia cell bioactivity. Exceptional activity following treatment with a compound or signaling input can be measured by phospho-flow in different ways, including exceptional potency (i.e., activity at nM or lower concentration)³⁸, bioactivity (i.e., orders of magnitude greater responses in a measured proteins; e.g., high per-cell γ H2AX indicating exceptionally strong DNA damage response)^{16,18,38,39}, pathway selectivity or signature profile (i.e., a specific combination of activities per cell, such as activation of γ H2AX without inhibition of p-4EBP1, indicating a DNA damage response without a halt to protein synthesis)^{38,39}, and/or exceptional selectivity (i.e., targeting a specific subpopulation of cells, such as targeting cancer cells without targeting healthy cells)^{16,38,40-42}. Here, the order of evaluation was first on bioactivity and selectivity, then on profile, and finally on evaluation of potency.

The panel selected for this initial test was designed to measure cell death, division, and translation and included post-translational modifications of proteins representing four fundamental cell processes: cleaved Caspase3 (c-CAS3, representing apoptosis), the gamma phosphorylation of H2AX (γ H2AX, representing a DNA damage response), phosphorylated Histone H3 (p-HH3, representing M phase and cell cycle activity), and S235/S236 phosphorylated S6 (p-S6, representing active protein synthesis and cell growth) (**Supplementary Table 1**). The arcsinh fold change in median fluorescence intensity vs. vehicle was calculated for each compound across each of the four functional readouts (**Figure 1E**). For more details on standard arcsinh scaling using the inverse hyperbolic sine in cytometry analysis, see original use in the Supplement of ⁴¹ and recent implementations ^{16,27,43,44}. While few compounds decreased the fold change in MFI vs. vehicle for any readout, 65 molecules were found to increase p-S6, c-CAS3, and γ H2AX significantly compared with vehicle. Of the 65 bioactive molecules selected, 13.8% (9/65) were rocaglates, a chemotype which comprised only 3.2% of the initial set of 600, and predominant in the bioactive set (**Figure 1F**). Nearly half (9/19) of the rocaglates tested in the **Diversity Set** were selected as bioactive. Of the bioactive rocaglates, driving activities included c-CAS3 and γ H2AX. Based on their high frequency of triggering c-CAS3 and/or γ H2AX in MV411 leukemia cells, rocaglates were prioritized for further investigation.

All 19 rocaglates in the initial 600 compound **Diversity Set** were RRs. To investigate whether rocaglates from different structural subgroups were bioactive in MV411 leukemia cells and whether this bioactivity extended to healthy cells, we tested an expanded library of 37 rocaglates, intentionally weighted toward known bioactive molecules, with representation from each of the three subclasses at 10 μ M on MV411 and healthy peripheral blood mononuclear cells (PBMCs) (**Figure 2A**). The primary staining panel was

selected to test a range of cell functions including cell growth, translation, and death and has been validated in previous work: Ki67 (cell proliferation ^{45,46}), phospho-S6 ribosomal protein at serine 240/244 (p-S6 S240/244) (growth, AKT/mTORC1 specific ^{47,48}), phosphorylation of gamma H2AX at serine 139 (γ H2AX) (DNA damage response ⁴⁹), phosphorylation of STAT3 at serine 727 (p-STAT3) (amplified transcriptional activity, downstream mTOR ⁵⁰), phosphorylation of STAT5 at tyrosine 694 (p-STAT5) (cell survival, highly phosphorylated in cancer cells ^{51,52}), phospho-S6 ribosomal protein at serine 235/236 (p-S6 S235/236) (growth, activated by ERK/RSK and AKT/mTORC1 ^{47,48}), phosphorylation of ERK1/ERK2 at threonine 202, tyrosine 204 (p-ERK1/2) (proliferation, activated by MAP kinase cascade ^{53,54}), phosphorylated histone H3 at serine 28 (p-HH3) (growth, cell cycle M phase ⁵⁵), phosphorylation of 4EBP1 at threonine 37/46 (p-4EBP1) (growth, activated by mTORC1 ⁵⁶), phosphorylation of lymphocyte-specific protein tyrosine kinase at tyrosine 505 (p-LCK) (SRC-family kinase involved in T-cell receptor signaling ⁵⁷), and phosphorylation of Akt at serine 473 (p-Akt) (Upstream of mTOR ^{41,58}) (**Supplementary Table 1**) ^{59,44,60}.

Next, the t-SNE algorithm was used to reduce the dimensionality of the data from 11-dimensional space to 2-dimensional space for visualization and analysis. The t-SNE plot was made using MV411 data from all 45 conditions tested (37 rocaglates, 3 vehicles, and 5 controls which are not shown) and then subdivided based on the subclass of origin (and vehicle) for comparison of high dimensional signaling profiles (**Figure 2B**). Striking differences were apparent in t-SNE plots for each of the three rocaglate subclasses when compared to the pooled vehicle data: vehicle cell density was localized to the upper left side of the map, whereas cell density for the rocaglate subclasses was nearly absent from this region. These differences were quantified by T-REX. Specifically, the t-SNE from each of the three respective rocaglate subclasses was compared to the pooled set of the 3 vehicle wells (**Figure 2C**). The T-REX plots revealed the cell populations enriched in vehicle-treated cells (blue) and rocaglate subclass-treated cells (red). For each T-REX plot, the T-REX degree of difference [(number of red cells + number of blue cells)/total number of cells] was computed to quantify how different subclass-treated cells were from vehicle-treated cells. All rocaglate subclass-treated cells had a T-REX degree of difference that was greater than 60%, indicating substantial change upon 16 hours of rocaglate treatment, relative to vehicle. In comparison, when T-REX was performed on vehicle 3 vs. the pooled set of vehicles 1 and 2, the T-REX degree of difference was <1%, indicating the vehicles were highly similar to each other. The contrast in cell density apparent in t-SNE and quantified by T-REX arises from phenotypic differences in MV411 cells treated with molecules from the three subclasses, as compared to vehicle, and confirms the original observation (**Figure 1**) that the three rocaglate subclasses display outstanding activity against leukemia cells.

To investigate the variation in bioactivity within each subclass, T-REX was applied by comparing the t-SNE from each compound to the pooled set of the 3 vehicle wells. As a representative of this analysis, the T-REX plot for **CMLD013342** vs. vehicle is shown in **Figure 2D** (data for remaining compounds are shown

in **Supplementary Figure 2**). Then, a T-REX degree of difference calculation was performed for each of the 40 comparisons (37 rocaglates + 3 vehicles). All 37 rocaglates had a T-REX degree of difference above a threshold of three times the interquartile range of the vehicle data over the vehicle median ($+3 \times$ IQR). Taken together, these results indicated that all tested rocaglates were bioactive against MV411 leukemia cells.

Rocaglates have distinct, subclass-specific bioactivity patterns in leukemia and healthy blood.

To further explore signaling patterns within and across the rocaglate subclasses, the field standard log-like arcsinh ratio of the median phospho-protein signal in treated cells as compared to the signal in the first vehicle (*i.e.*, $\text{arcsinh} [\text{MFI}_{\text{treated}}/\text{MFI}_{\text{vehicle1}}]$) was performed for each marker and compound. This resulted in values that range from 0 (no change) to $+1$ arcsinh unit (an increase of X -fold, yellow) or -1.5 (a decrease of $1.5X$ -fold, blue). In **Figure 3A**, compounds were arranged according to rocaglate subclass to allow comparison of phenotypic differences within and between subclasses. In **Figure 3B**, the same data were arranged by hierarchical clustering of the signaling profile (below). When organized by subclass, RPs displayed a more consistent signature than other subclasses (median IQR across all markers for RP = 0.07, RR = 0.2, and ADR = 0.2). The RP signature, which was also observed in some RRs, included exceptionally strong γ H2AX and inhibition of p-ERK, a signature indicating activation of DNA damage response and inhibition of MAPK proliferative signaling. In contrast to RPs, most ADRs inhibited γ H2AX and p-4EBP1 and activated p-AKT and p-STAT5, a signature suggesting activation of cell survival signaling that would be undesirable in an anti-cancer compound. RRs were heterogeneous in the pattern of phosphorylation seen following rocaglate treatment and had members that displayed patterns that were similar to either RPs or ADRs. Grouping by unsupervised clustering of bioactivity led to the successful separation of RPs into clusters 2 and 3 and ADRs into clusters 4 and 5 (except for **CMLD012600**); RRs were interspersed throughout clusters 2 - 5. One readout that strongly contrasted between clusters was γ H2AX; while cluster 3 had the highest γ H2AX compared to vehicle, cluster 2 had moderately high γ H2AX compared to vehicle, cluster 5 had similar γ H2AX compared to vehicle, and cluster 4 had low γ H2AX compared to vehicle. Structures for the 37 rocaglates grouped according to rocaglate subclass and hierarchical clustering can be found in **Supplementary Figure 3** and **Supplementary Figure 4**, respectively.

To investigate how γ H2AX induction patterns by rocaglates compared between MV411 leukemia cells and healthy blood, the percentage of γ H2AX positive (% γ H2AX+) cells for each compound (and vehicle) was quantified for each cell type (**Figure 4A**). A threshold of 15% γ H2AX+ cells was selected to distinguish compounds that trigger significant DNA damage in MV411 and PBMC. Only the vehicles and three RRs did not trigger significant DNA damage in either cell type. Nearly all ADRs and one RR (**SDS-1-021**⁶¹⁻⁶⁵, represented twice in the set as both racemic (**CMLD011880**) and enantioenriched (**CMLD010508**) stocks) triggered significant DNA damage in PBMC alone. Distinctly, two ADRs induced significant DNA damage in both cell types. Lastly, all RPs and some RRs led to leukemia-specific induction of γ H2AX. Quantification

of γ H2AX induction in both cell types further illuminated rocaglate structural subclass bioactivity signatures and outlier compounds within each subclass.

To identify compounds with desirable anti-leukemia activity, the compounds with the greatest leukemia-specific induction of DNA damage were next identified. The log₂ fold ratio of the percentage of γ H2AX⁺ cells was compared between MV411 and healthy PBMC cells (i.e., log₂ [% γ H2AX⁺_{MV411} / γ H2AX⁺_{PBMC}], **Figure 4B**). The ADRs that had PBMC-specific induction of DNA damage in **Figure 4A** were in the lower left quadrant, as they had the lowest leukemia-specificity of DNA damage induction and low induction of DNA damage in MV411. All RPs and the select RRs that were demonstrated to have leukemia-specific induction of γ H2AX in **Figure 4A** were found to be in the top right quadrant, as they had maximal DNA damage induction in MV411 and a specificity of this induction to leukemia cells. Three RPs, **CMLD012390**, **CMLD013342**, and **CMLD013348** had the highest log₂fold ratio of % γ H2AX⁺ in MV411 to % γ H2AX⁺ PBMC. The representative contour plot of **CMLD013342** depicted in **Figure 4C** demonstrated a distinct population of γ H2AX hyperactivated cells that was not seen in the contour plot for **CMLD012600**, a rocaglate with a similar % γ H2AX⁺ cells. This γ H2AX hyperactivated population of cells by **CMLD013342** was also not present in the PBMC contour plot. The uniform induction of a leukemia-specific DNA damage response associated with cell death and γ H2AX hyperactivated population of cells, helped establish RPs as an interesting subclass of rocaglates for further investigation as potential leukemia therapeutics.

The rocaglate RP subclass selectively induces a γ H2AX⁺ p-4EBP1⁺ bioactivity pattern seen only in leukemia cells.

To visualize the differences in high-dimensional signaling profiles across rocaglate subclasses, a T-REX analysis was performed by comparing the t-SNE from each of the three respective rocaglate subclasses to the pooled set of the other rocaglate structural subclasses (including vehicle) (**Figure 5A**). These T-REX plots identified the cell populations that were enriched in (red) and absent from (blue) a given subclass. Notably, T-REX illuminated a subpopulation of cells specifically induced by RPs (RP island). Coloring the grouped t-SNE for all 45 compounds based on measurements for each functional readout suggested that the RP island was a region of cells with high median fluorescence intensity of p-4EBP1 and γ H2AX (**Supplementary Figure 5**).

To determine the true specificity of this island to RPs, the percentage of cells in the RP island for each rocaglate was calculated (**Figure 5B**). A Wilcoxon rank sum test of the percentage of cells in the RP island for RPs vs. the remaining subclasses was conducted as a test of significance and generated a p-value of 0.015. The percentage of cells falling into the RP island was also correlated with transformed 98th percentile fluorescence intensity for γ H2AX and p-4EBP1 (**Supplementary Figure 6**). A Marker Enrichment Modeling (MEM)²⁸ analysis was performed on the RP island to characterize contextual protein enrichment (absolute MEM labels range from 0 = no enrichment to +10 = highest enrichment). The resulting MEM label was: γ H2AX+10, p-4EBP1+7, Ki67+5, p-ERK+5, and p-S6 - S240/244 +5 (**Figure**

5C). The enrichment for γ H2AX and p-4EBP1 indicated that RP island cells had exceptionally high bioactivity, seen in the activation of the DNA damage response (γ H2AX), and a distinct signature, seen in retention of some mTOR pathway activity (p-4EBP1).

To assess the leukemia-specificity of the RP-island, as activation of the DNA damage response in healthy cells would be detrimental, a t-SNE analysis was performed on the MV411 and PBMC together to generate a new embedding with corresponding x and y axis scales (**Figure 5D**). After dimensionality reduction with t-SNE, cells were separated based on origin (MV411 or PBMC) (**Figure 5D**). The MV411 t-SNE analysis revealed a distinct island of cells from the remaining cell density on the lower right side of the plot. When quantified, this cell population was >99% specific to MV411. This leukemia-specific island was confirmed as the RP island after a MEM analysis generated the same protein enrichment label as displayed in **Figure 5C** (**Figure 5E**). Together, these results established that RPs activate a strong and leukemia-specific DNA damage response through the RP island.

RP structure correlates with exceptional DNA damage responses in leukemia cells.

While the Wilcoxon rank sum test provided a subclass-level validation of the RP specificity of the RP island, further investigation of bioactivity and structure at the compound level was necessary for understanding which compounds shifted cells to the RP island and why. To address this, the t-SNE map of pooled RPs was further subdivided based on compound of origin, and the percentage of total cells falling into the RP island was quantified for each RP (**Figure 6**). This compound-level analysis highlighted three RPs, **CMLD012390**, **CMLD013342**, and **CMLD013348**, as the only compounds with greater than 5% of cells in the RP island gate. Markedly, these are the same compounds with the highest log₂fold ratio of % γ H2AX+MV411/% γ H2AX+ PBMC as reported in **Figure 4B**. These three compounds commonly possess three unique structural features that are found in isolation in the other six RPs, specifically a non-aryl pyrimidinone ring substituent (termed the RP “R-group”), a 4'-methoxy substituent on the rocaglate “B”-ring, and a monocyclic, *N*-unsubstituted pyrimidinone ring. This suggested that a combination of multiple structural features might drive SAR for the RPs with high cell density in the RP island. Thus, the RP island was associated with an exceptional, leukemia-specific induction of the DNA damage response and prominent structural features were identified for future SAR studies.

To further investigate the potency and timing of γ H2AX activation by these special RPs, dose-response and time course experiments were performed in MV411 cells. For the dose-response, six doses (10, 5, 1, 0.1, 0.05, 0.01, and 0 μ M) were selected to encompass a 1000-fold range and the top two RPs, **CMLD012390** and **CMLD013342** were chosen for testing. A γ H2AX + gate was drawn on biaxial plots of Side Scatter Area (SSC-A) vs. γ H2AX based on γ H2AX levels in the vehicle-treated cells. The percentage of cells within the γ H2AX + gate was quantified for each concentration and used to calculate the half-maximal activating concentration (EC₅₀) (**Figure 7A**). **CMLD012390**, the RP with the greatest percentage of cells in the RP island in **Figure 6**, had an EC₅₀ of 213 nM, and **CMLD013342**, the RP with the second

greatest percentage of cells in the RP island had a slightly higher EC₅₀ of 492 nM. **CMLD012390** and **CMLD013342** were also used to challenge MV411 cells at 1, 4, and 16 hours to investigate the timing of γ H2AX induction. Within 4 hours, mild γ H2AX induction was seen for both compounds; this γ H2AX induction increased in potency to match levels seen in both the **Rocaglate Set** and dose-response within 16 hours (**Figure 7B** and **Supplementary Figure 7**). Thus, these two RPs shared multiple structural commonalities and demonstrated a strong and reproducible induction of DNA damage response within 4 hours. Going forward, future SAR studies should involve direct, pairwise comparisons to discern the individual impact of the unique structural features in these RPs on specificity in selectively inducing DNA damage responses in leukemia cells and not in healthy blood.

Discussion

An advantage of single cell chemical biology approaches is the potential to test a greater space of cell biology readouts and to do so with primary cells that include diverse healthy cell types. Thus, phospho-flow and single cell chemical biology approaches can be used to include a range of cells that will be present in the human tissues in which compounds will need to function *in vivo*. Here we report a test of diverse compounds facilitating structure-activity relationship studies based on multiparameter, single cell phospho-flow assays. This revealed rocaglates as a highly bioactive chemotype and identified all rocaglate pyrimidinones tested and most amidino-rocaglates tested as chemotypes that respectively selected for leukemia or healthy blood cells. The comparison between primary human blood and leukemia cells acts as a proof of concept for potentially employing human tissues in early phase SAR studies.

Since phospho-flow was first combined with FCB for chemical biology screening⁶, this approach has been used for screening and characterization of promising compounds based on cell type- and pathway-specific effects. The screen of small molecules described here was similarly performed in combination with FCB for multiplexing and used a focused panel of four functional states of proteins to efficiently profile the modulation of relevant pathways involved in leukemia cell translation and death. The screening set of 600 structurally diverse compounds is a relatively small library compared to what might be possible if this approach were scaled; part of the success of the study with this focused library is likely due to the selection of key functional signaling events to measure as bioactivity readouts. Here, multiplexed cytometry measured 2,400 cell-based biological readouts at the single cell level in one experiment. Future studies might go beyond this work by leveraging additional advantages of phospho-flow signaling profiles, especially the ability to resolve impacts on different subpopulations of primary cells in blood, tumors, or other tissue. Sixty-five promising activators were selected for further investigation based on activity at p-S6, here interpreted as measuring growth and translation, c-CAS3, an indicator of apoptosis, and/or γ H2AX, a readout of DNA damage. While previous studies have focused on selection of individual hits for follow-up testing, this approach was focused instead on identifying promising chemotypes. Of the 65-compound bioactive set, 9 of the compounds were rocaglates, a natural product family that has received a great deal of attention for promising anti-cancer and anti-infective activities.

While rocaglate bioactivity has been of interest since rocaglamide A (RocA), the prototypic rocaglate, was first isolated in 1982, this study was the first phospho-flow-based test of rocaglates and one of only a few that have characterized the bioactivities of various structural subclasses within the chemotype³². Larger screens of >200 structurally diverse rocaglates have mainly monitored clamping of the three eIF4A homologs (eIF4A1-3) to RNA, *in vitro* translation, and cytotoxicity^{36,37,66}. This test set of 37 rocaglates from three structural subclasses profiled bioactivity using 11 functional states of proteins across MV411 and PBMC cells, enabling 22 cell-based biological readouts to be tested for each compound simultaneously at the single cell level.

Cytometric profiling of the 37 rocaglates in combination with t-SNE and T-REX enabled us to establish that these molecules were all bioactive against leukemia cells, considering all 11 protein functional readouts measured at the single cell level. Additionally, investigation of fluorescence intensity for each molecule and readout enabled rocaglate bioactivity within and across subclass to be compared. There were noticeable similarities in fluorescence intensity for readouts within each subclass and differences across subclass: RPs were distinguished by high γ H2AX, p-S6 S240/244, and p-HH3 and ADRs demonstrated low γ H2AX and p-4EBP1 and high p-AKT and p-STAT5. However, a dendrogram of transformed fluorescence intensity across all 11 functional readouts and 37 compounds only separated ADRs from RPs; this suggested that the previously identified structural subclasses may not fully capture bioactivity differences.

Analyzing γ H2AX, a readout that contrasted across the clusters of rocaglates formed in **Figure 3**, in both PBMC and MV411 shed even further light on bioactivity differences between structural subclasses. While some RRs and all RPs activated γ H2AX in MV411 and not PBMC, ADRs only activated γ H2AX in PBMCs. For RRs these findings are not surprising; rocaglamide A, an RR, has demonstrated leukemia-specific bioactivity in previous studies⁶⁷. However, ADRs have been found to be the most potent *in vitro* RNA clampers of eIF4A1 and eIF4A2^{35,37}. Interestingly, RPs, the subfamily demonstrating uniform leukemia-specific induction of γ H2AX, have primarily been reported as agents for hepatitis C virus (HCV) and to show a possible bias for the inhibition of viral entry over translation inhibition³⁴. The natural product aglaroxin C (**CMLD011866**) has been cited as a potent cytotoxic agent against multiple cancer cell lines, though not yet as selective for cancer cells as opposed to healthy cells⁶⁸⁻⁷¹. Therefore, these findings reveal leukemia-specific bioactivity, or a lack thereof, for RPs and ADRs, respectively.

By leveraging t-SNE, T-REX, and MEM, we were able to identify a small subpopulation of cells with inflated γ H2AX that was only present in RP-stimulated MV411 cells. Given that this population was <3% of total MV411 cells, it is likely that this would have been overlooked by approaches only accounting for bulk cell responses. Further investigation allowed us to uncover **CMLD012390**, **CMLD013342**, and **CMLD013348** as the primary contributors to this island of cells. This response was not seen by aglaroxin C (**CMLD011866**) though its potent cytotoxic activity is warranted based on prior literature^{69,70}. γ H2AX is produced in response to histone H2AX phosphorylation on a serine four residues from the carboxyl

terminus and acts as a sensitive marker for DNA double-stranded breaks (DSBs). Many existing cancer therapies such as etoposide and mitoxantrone act by introducing sufficient DSBs to activate cell death pathways.⁷² Therefore, the ability of these top three RPs to potently activate γ H2AX in only leukemia cells should now be studied for potential clinical use. Rocaglates have been shown to have two key classes of DEAD-box RNA helicase targets: the three eIF4A homologs, and more recently DDX3^{73,74}. While eIF4A is implicated in translation initiation, DDX3 is involved in a range of functions related to RNA metabolism⁷⁵.

The compounds **CMLD012390**, **CMLD013342**, and **CMLD013348** have not been previously reported as distinct members of the RP subfamily. Notably, these compounds possess some structural similarities; they commonly have a methoxy group on their B-ring (the southeastern ring), a non-aryl R-group (the northern projection from the pyrimidinone), and a monocyclic, N-unsubstituted pyrimidinone with unique steric properties that, in contrast to *N*-substituted RPs such as **CMLD011866**, is capable of equilibrating to an aromatic hydroxypyrimidine tautomeric form. Further structure-activity relationship and target identification studies may be used to confirm which structural commonalities possessed by the top three RPs drive their exceptional γ H2AX activation ability. While the RP-specific γ H2AX signature (the RP island) for **CMLD012390**, **CMLD013342**, and **CMLD013348** may imply an as-yet unknown target, it is also possible that differences in potency or timing of engagement with eIF4A1 and/or DDX3 drive this unique activity signature.

In this study, we investigated structure-activity relationships for 37 rocaglates across 11 key functional proteins representing hallmarks of cell biology in MV411 leukemia cells and PBMCs. The functional states of proteins chosen, while more expansive than a single target- or readout-based approach, did not capture all potentially impacted pathways. For confirming the leukemia-specificity of the bioactive molecules identified, additional testing on other leukemia cell lines and primary leukemia samples would help to show the range of impact in the disease. While only the top three RPs shifted cells to the γ H2AX high, RP island, there were many other compounds worth investigating further. From a clinical standpoint, the RRs **CMLD011890** and **CMLD011891** also had among the highest leukemia-specific induction of γ H2AX. Additionally, the RR **CMLD011880/CMLD010508 (SDS-1-021)** and the ADRs **CMLD012565** and **CMLD012600** had distinct bioactivities within their subclass; the structural determinants of these are yet unknown and could also potentially be explained by engagement with specific targets.

The subset of rocaglate pyrimidinones (RPs) with exceptional, leukemia-specific γ H2AX activation identified through successive testing of 600 diversity compounds followed by 37 select rocaglates demonstrate the value of an expanded biological readout space for deconvolving the impact of small chemical changes. We anticipate this combined set of experimental and computational approaches to be adaptable for testing other promising families of structurally related molecules across a range of functional readouts, cell types, and disease states.

Methods

Screening Compounds:

Initial screens were performed on a **Diversity Set** of 600 compounds selected from the BU-CMD's in-house small molecule screening collection that was comprised of 4,445 compounds at the time of **Diversity Set** assembly. The **Diversity Set** was curated to contain a roughly proportional count of members from each of the 120 distinct structural chemotypes present in the parent screening collection. **Diversity Set** representatives for each chemotype (range: 1-23 compounds per chemotype; average: 5.4 members per chemotype) were randomly chosen from a preselected field of available compounds meeting a specific stock volume threshold, in order to ensure availability of follow-up material for validation and secondary assays.

Follow-up assays were performed on a hand-chosen cohort of 37 rocaglates (**Supplementary Figure 3**), selected to include a diverse representation of bioactive rocaglate subclasses, with additional consideration of known bioactivity, structural diversity, and available stock quantity.

PBMC Collection and Preservation:

Peripheral blood mononuclear cells (PBMCs) were obtained in accordance with the Declaration of Helsinki following protocols approved by Vanderbilt University Medical Center IRB. PBMCs were collected, isolated, and cryopreserved from approximately 20 mL of freshly drawn blood as previously described⁷⁶. Briefly, peripheral blood was drawn into sodium heparin anticoagulant, and PBMCs were isolated by centrifugation after layering on top of a Ficoll-Paque PLUS (GE Healthcare Bio-Sciences) gradient.

MV411 Cell Culture:

MV411 cells were obtained from ATCC (CRL-9591) and confirmed to be mycoplasma negative. MV411 was cultured in IMDM (Gibco 12440-053), supplemented to a final concentration of 10% fetal bovine serum. Cells were incubated in a water-jacketed 5% CO₂ incubator at 37 °C and maintained at densities between 100 thousand and 1 million cells per ml of culture media, fed every other day, and passaged every 4 days.

Fluorescent Cell Barcoding (FCB) Assays.:

The general protocol followed Schares et. al. Barcoding plates were prepared as described in Section 6.1 of Schares et. al.¹³.

Diversity Set Experiment:

For performing **Diversity Set** testing, the compound plate was prepared as follows: control compounds (3 replicates of DMSO, staurosporine, etoposide, aphidicolin, rapamycin, and nocodazole) were prepared to final concentrations listed in **Supplementary Table 2** and plated at 1 µL into columns 6 and 12 of 8 96

well tissue culture plates (Fisher). The **Diversity Set** compounds were acquired from the BU-CMD and plated at a volume of 1 μ L into the remaining 80 wells of the 8 96-well tissue culture plate for a final concentration of 10 μ M.

MV411 cells from suspension culture (0.5-1 million cells/mL) were acquired, pelleted, and resuspended in media to around 500,000 cells/mL at least 2 hours before start of the experiment as described in Schares et. al. Cell suspension was dispensed at 200 μ L to each of the 8 96 wells plates and pipetted to mix. The plate was incubated at 37°C, 5% CO₂ for 16 hr. The remaining steps were performed as described in Schares et. al. Briefly, cells were stained for viability with 0.04 μ g/mL Ax700, fixed with 1.6% paraformaldehyde, and permeabilized with 100% ice-cold methanol. Cells were stained using eight concentrations of pacific blue and six concentrations of pacific orange (per 48 wells) for generating a unique fluorescent barcode for each well, and one concentration of AlexaFluor 750 as an internal control. Cells were pooled into one flow cytometry tube per 48-well plate (16 tubes total) for staining with the following antibody panel: c-CAS3, γ H2AX, p-S6 S240/244, and p-HH3 (more information detailed in **Supplementary Table 1**). Compensation controls for each antibody and dye and bead controls were prepared and used for the set-up of the flow cytometer. Flow cytometry data was acquired on a 5-laser (355 nM, 405 nM, 488 nM, 561 nM, and 635 nM) BD LSR II Fortessa instrument.

Rocaglate Set Experiment:

For the **Rocaglate Set** experiment, the compound plate was prepared as follows. Control compounds (3x DMSO, staurosporine, etoposide, **CMLD010335**, rapamycin, and nocodazole) were prepared to final concentrations listed in **Supplementary Table 2** and plated at 1 μ L in columns 6 and 12 of a 96 well tissue culture plate. The **Rocaglate Set** (RR = 20, ADR = 8, RP = 9) acquired from the BU-CMD was plated at a volume of 1 μ L into the remaining 80 wells of the 96-well tissue culture plate at a final concentration of 10 μ M.

As described above, MV411 cells from suspension culture were acquired, pelleted, and resuspended in media at least 2 hours before start of the experiment. PBMCs were thawed from cryopreservation and resuspended at $\sim 2 \times 10^6$ cells/mL. The cell suspension was dispensed at 200 μ L to each of the 96 wells of 2 compound plates, one for each cell type, and pipetted to mix. The plate was incubated at 37°C, 5% CO₂ for 16 hr. The remaining steps were performed as described above, pooling cells into two total flow cytometry tubes for antibody staining: one for PBMC and one for MV411¹³. Antibody staining panel included p-STAT3, p-STAT5, p-ERK, p-HH3, p-4EBP1, p-S6 S240/244, p-S6 S235/236, Ki67, and γ H2AX (**Supplementary Table 1**). Flow cytometry controls were performed as above. Flow cytometry data was acquired on a four laser (405 nM, 488 nM, 561 nM, and 640 nM) Cytek Biosciences Aurora spectral flow cytometer following spectral unmixing with compensation controls.

Dose-Response:

CMLD012390 and **CMLD013342** were prepared at the following doses and plated at a volume of 1 μ L into the first two columns of rows A-H of a 96-well plate, respectively: 0.01, 0.05, 0.1, 0.5, 1, 5, 10, and 0 μ M.

As described above, MV411 cells from suspension culture were acquired, pelleted, and resuspended in media at least 2 hours before start of the experiment. The cell suspension was dispensed at 200 μ L to each of the two columns of the 96 well plate and pipetted to mix. The plate was incubated at 37°C, 5% CO₂ for 16 hr. Cells were stained for viability, fixed, and permeabilized. Cells for each compound were then stained using eight concentrations of pacific blue, one concentration of pacific orange, and one concentration of AlexaFluor 750. Cells for each compound were pooled into their own flow cytometry tube for staining with the same antibody panel as the **Rocaglate Set** experiment (**Supplementary Table 1**). Flow cytometry controls and data collection were performed as in the **Rocaglate Set** experiment.

Time Course:

CMLD012390 and **CMLD013342** were prepared to a final concentration of 10 μ M. As described above, MV411 cells from suspension culture were acquired, pelleted, and resuspended in media. Cell suspension was dispensed at 200 μ L to the first three columns of rows A-C of a 96-well plate. A 16-hour reverse time course was begun by pipetting 1 μ L of DMSO, **CMLD012390**, and **CMLD013342** into columns 1, 2, and 3, respectively of row A and pipetted to mix. After 12 h, 1 μ L of DMSO, **CMLD012390**, and **CMLD013342** was dispensed into row B for the 4-h time point. Lastly, after 15 h, 1 μ L of DMSO, **CMLD012390**, and **CMLD013342** was dispensed into row C for the 1-h time point. At the completion of the reverse time course, all three rows were stained for viability, fixed, permeabilized, and stained using three concentrations of pacific blue, three concentrations of pacific orange, and one concentration of AlexaFluor 750. Cells were pooled into one flow cytometry tube and stained with the following antibody panel: LC3, p-EIF2 α , p-MLKL, c-CAS3, and γ H2AX (*cf.* **Supplementary Table 1**). Flow cytometry controls were performed as above. Flow cytometry data was acquired on a three-laser (405 nM, 488 nM, and 640 nM) Cytek Biosciences Northern Lights spectral flow cytometer following spectral unmixing with compensation controls.

FCB Data Preprocessing and Analysis:

Data was uploaded and stored in Cytobank for scaling, quality control gating, compensation, and analysis of unmixed cytometry data (FCS file format). Raw median fluorescence intensity values were transformed to a hyperbolic arcsine scale. For the **Diversity Set** test, a cofactor of 150 was selected for all functional readouts. For the **Rocaglate Set** experiment, the default cofactor of 6000 was selected for all markers except for Ki67 (25000), p-S6 S240/244 (12000), p-LCK (12000), p-STAT3 (12000), p-STAT5 (25000), p-S6 S235/236 (12000), p-HH3 (12000), and p-4EBP1 (12000). For the dose-response and time course experiments, a cofactor of 6000 was selected for all functional readouts. Quality control gating (QC) consisted of QC1: FSC-A vs. FSC-H for singlets, QC2: FSC-A vs. SSC-A for intact cell bodies, QC3: FSC-A vs. Ax750 for barcode uptake control, and QC4: Ax750 vs. Ax700 for viable cells. Scaled and gated

samples were compensated and then computationally deconvoluted using the DebarcodeR algorithm. The resulting FCS files for each well were uploaded to Cytobank for storage and further analysis.

T-REX Analyses:

T-REX takes a pair of dimensionally reduced maps of equal size as inputs, creates a plot that highlights hotspots of cells in phenotypic regions that are the most different between the two files, and provides a T-REX degree of difference value for each analysis performed. The dimensionality reduction tool used here was t-SNE which was performed as follows: after debarcoding, the FCS files corresponding to stained cellular events from all 48 wells for both cell types were uploaded to one Cytobank experiment and fed into a t-SNE-CUDA analysis through Cytobank (settings: channels = 11 functional readouts included in the panel, iterations = 10,000, perplexity = 60, automatic learning rate, early exaggeration = 12). FCS files containing protein measurements for all readouts with the t-SNE axes appended were exported into R for subsequent analysis. Each dimensionally reduced map was sampled such that each well and rocglate structural subclass, when applicable, was equally represented before applying T-REX. The T-REX algorithm was applied in R using a k value of 60. A modular data analysis workflow including UMAP as the dimensionality reduction tool, K-Nearest Neighbors (KNN), and Marker Enrichment Modeling (MEM) was developed in R and is available online (<https://github.com/cytolab/T-REX>).

MEM Protein Enrichment Analyses:

Marker Enrichment Modeling from the MEM package (<https://github.com/cytolab/mem>) was used to characterize feature enrichment in the RP specific island identified. MEM normally requires a comparison of a population against a reference control, such as a common reference sample, all other cells, or induced pluripotent stem cells^{28,77,78}. Here, a statistical reference point intended as a statistical null hypothesis was used as the MEM reference. For this statistical null MEM reference, the magnitude was zero and the IQR was the median IQR of all features chosen for the MEM analysis. Values were mapped from 0 enrichment to a maximum of +10 relative enrichment. The contribution of IQR was zeroed out for populations with a magnitude of 0.

Cell Gating Analyses:

Gating was performed in Cytobank to quantify the % γ H2AX positive cells in response to each compound in MV411 and PBMC. This analysis was performed by plotting Side Scatter Area vs. γ H2AX for each compound and cell type and using the polygon gate tool to draw an area that maximizes the difference in percentage % γ H2AX positive cells between the vehicle and control wells. Quantifying the percentage of cells in the RP island was also performed using the polygon gate in Cytobank.

EC₅₀ Calculations:

EC₅₀s were calculated in R using the drm function from the drc package. A four-parameter log-logistic function was selected with the following unfixed parameters: hill slope, minimum value, maximum value,

and EC₅₀. The dependent variable for the dose-response formula was the percentage of cells in the γ H2AX+ gate which was calculated on Cytobank and imported into R. The independent variable was the common logarithm of the dose for each concentration.

Acknowledgements

Research was supported by the following funding resources: NIH/NCI grants U01 TR002526 (HLT, MJH, LEB, JAP, Jr., JMI), R01 CA226833 (JMI, HLT, MJH), R35 GM118173 (JAP, Jr.), the Vanderbilt-Ingram Cancer Center (VICC, P30 CA68485), the Michael David Greene Brain Cancer Fund (JMI), the Southeastern Brain Tumor Foundation (JMI), and the Ben & Catherine Ivy Foundation (JMI).

Author Contributions

All authors contributed to developing the study and reviewing and editing the manuscript. LEB and JAP, Jr. developed the process to select molecules for the **Diversity Set** and **Rocaglate Set**, provided access to the **Diversity Set** and **Rocaglate Set**, and contributed to rocaglate subtype binning, R-group decomposition, and structure-activity analysis. MJH and HLT collected and processed data. HLT and JMI conducted flow cytometry data analysis and interpretation. HLT developed data analysis scripts. HLT and JMI drafted the manuscript. JMI and JAP, Jr. provided financial support.

Declaration of Interests

All authors declare no competing interests.

Thirman et. al. – Figure 1

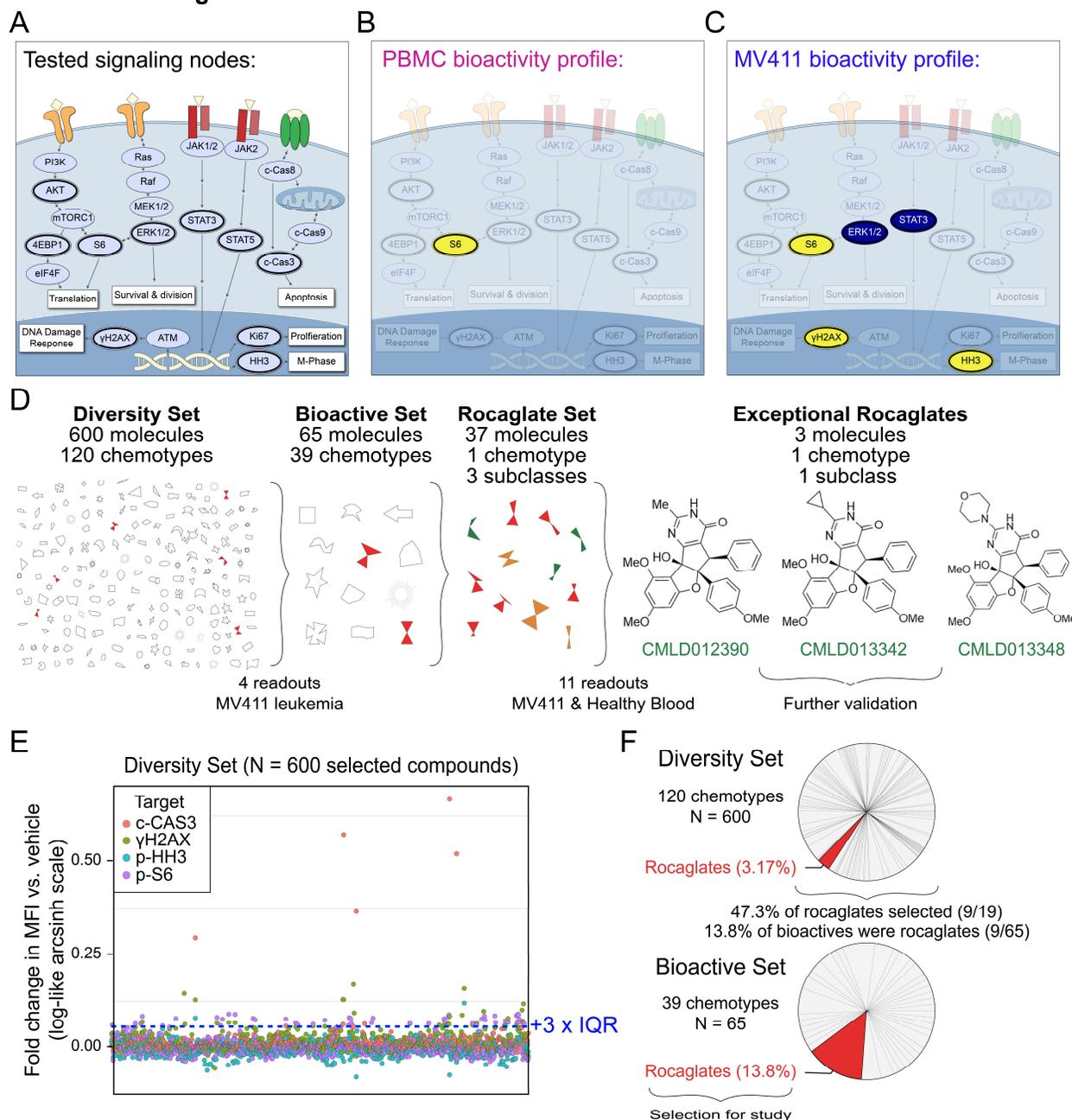


Figure 1 – Rocaglates displayed exceptional activity in a multiplexed single cell bioactivity assay. **A)** A model of a cell annotated with the key readouts measured (dark outlines). mTOR pathway activation is explored through testing p-AKT, p-ERK1/2, p-S6, and p-4EBP1, which regulate translation, among other cellular processes. Activation of the DNA damage response is explored through γH2AX. Ki67 and p-HH3 are indicative of proliferation and M-phase of the cell cycle, respectively. p-STAT3 and p-STAT5 are transcription factors that regulate the expression of cell cycle, survival, and pro-inflammatory genes. Activation of apoptosis is measured through the detection of c-CAS3. Key readouts activated (yellow) or inhibited (blue) by the exceptional three rocaglates (shown in D) are depicted for PBMC in **B)** and MV411 in **C)**. **D)** Depiction of the progression from 600 compound **Diversity Set** to three exceptional rocaglates. A symbolic representation of the 600-compound **Diversity Set** comprised of 120 chemotypes, the 65 molecules identified as bioactive, and the **Rocaglate Set** selected for structure-bioactivity relationship studies. The names and structures for the three exceptional rocaglates are shown. The number of total molecules, chemotypes, and subclasses, where relevant, are depicted above each respective phase of testing. The number of readouts and cells used for testing are depicted below each phase. **E)** The arcsinh scaled fold change in median fluorescence intensity vs. vehicle for the 600-compound **Diversity Set** of natural products on the four readouts tested. Each readout is represented by a different colored circle. The bioactivity threshold was drawn at the vehicle median + 3 x interquartile range (IQR) (0.0573). **F)** The percentage of the 600 compounds from each of the 120 chemotypes represented in the **Diversity Set**. The pie slice corresponding to rocaglates is highlighted in red. Bioactive molecules were selected based on the threshold shown in **E)**. The percentage of the 65 bioactive molecules from each of the 39 chemotypes represented is shown, with the pie slice corresponding to rocaglates highlighted in red.

Thirman et. al. – Figure 2

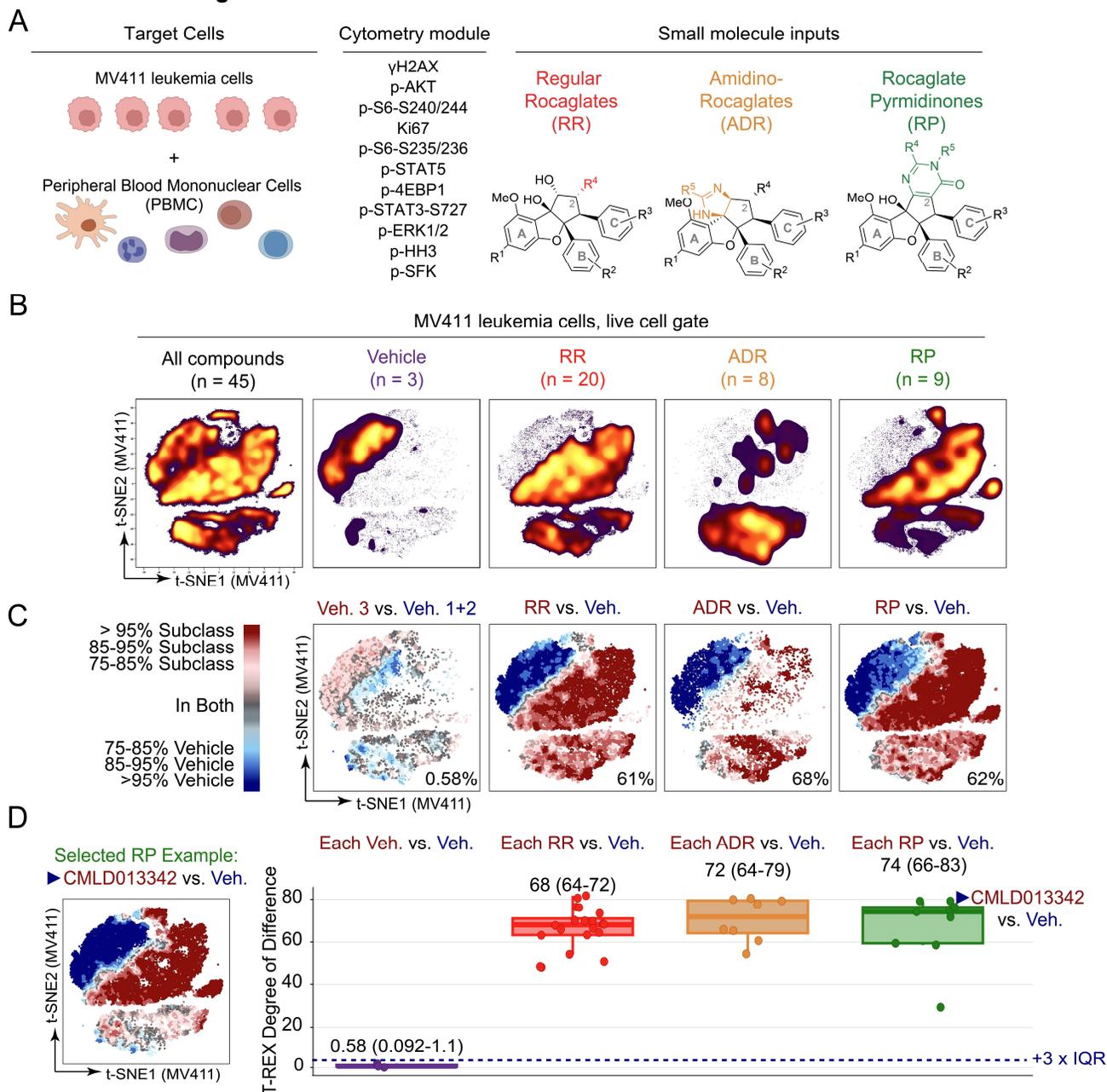
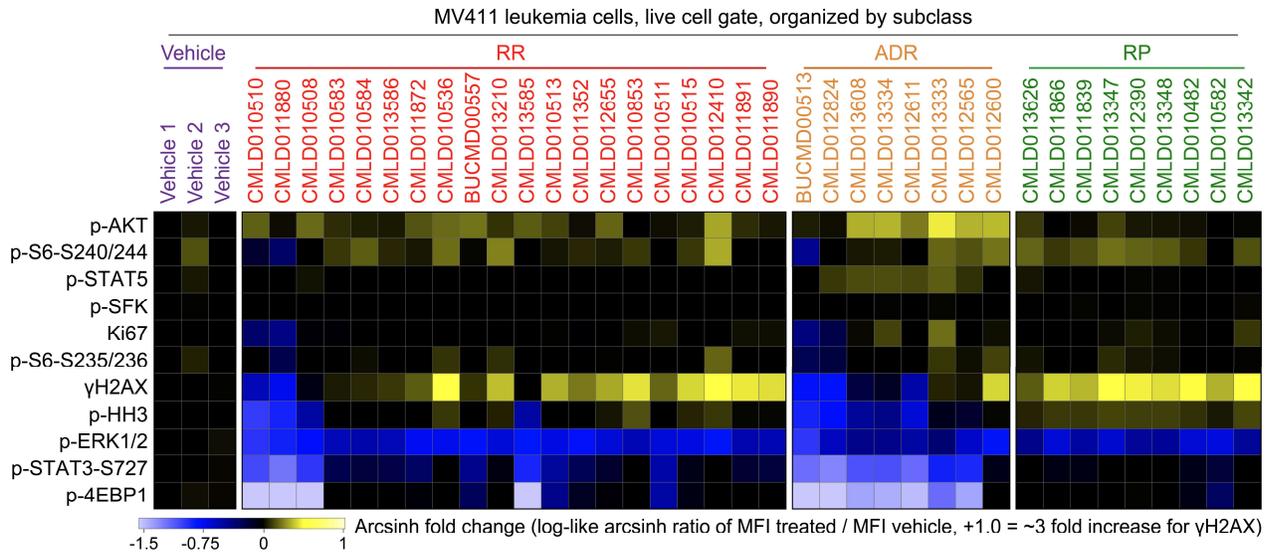


Figure 2 – Rocaglates are a chemical family with bioactivity against leukemia cells. **A)** Shown, are the three elements core to a single cell chemical biology experiment conducted using flow cytometry – target cells, a module of cytometry readouts, and small molecule inputs - and the inputs selected for these studies. The MV411 leukemia cell line and human peripheral blood mononuclear cells were selected as the set of target cells. Eleven readouts of core cell functions were selected for measuring per cell as part of the cytometry module. Lastly, rocaglates falling into the three depicted structural subclasses – regular rocaglates, amidino rocaglates, and rocaglate pyrimidinones – were chosen as the small molecule inputs. **B)** Plot depicting the result of performing a t-SNE analysis on the entire pre-processed MV411 dataset (All compounds). The t-SNE of all compounds is divided based on the category of input (Vehicle, RR, ADR, and RP) and colored based on cell density. **C)** T-REX plots depict regions of significant difference between the t-SNE of one rocaglate subclass vs. vehicle-treated cells in MV411. As a representative of variation between the three tested vehicle wells, a T-REX plot is shown for the analysis of vehicle 3 vs. the pooled set of vehicles 1 and 2 (leftmost plot). The T-REX degree of difference [(number of red cells + number of blue cells)/total number of cells] for each analysis is reported in the lower right corner of each plot. **D)** Box and whisker plot of T-REX degree of difference for each compound vs. vehicle-treated cells in MV411. Corresponding box and whiskers are grouped and colored according to rocaglate structural subclass. The median and IQR range of the T-REX degree of difference for each subclass is provided above the respective set of box and whiskers. The dashed navy line is shown at the vehicle median + 3 x IQR to indicate a significance threshold. Representative T-REX plot for **CMLD013342** vs. vehicle is shown to the left of the box and whisker plot for reference on the analysis conducted. T-REX plots for the remaining compounds are depicted in **Supplementary Figure 2B**.

Thirman et. al. – Figure 3

A



B

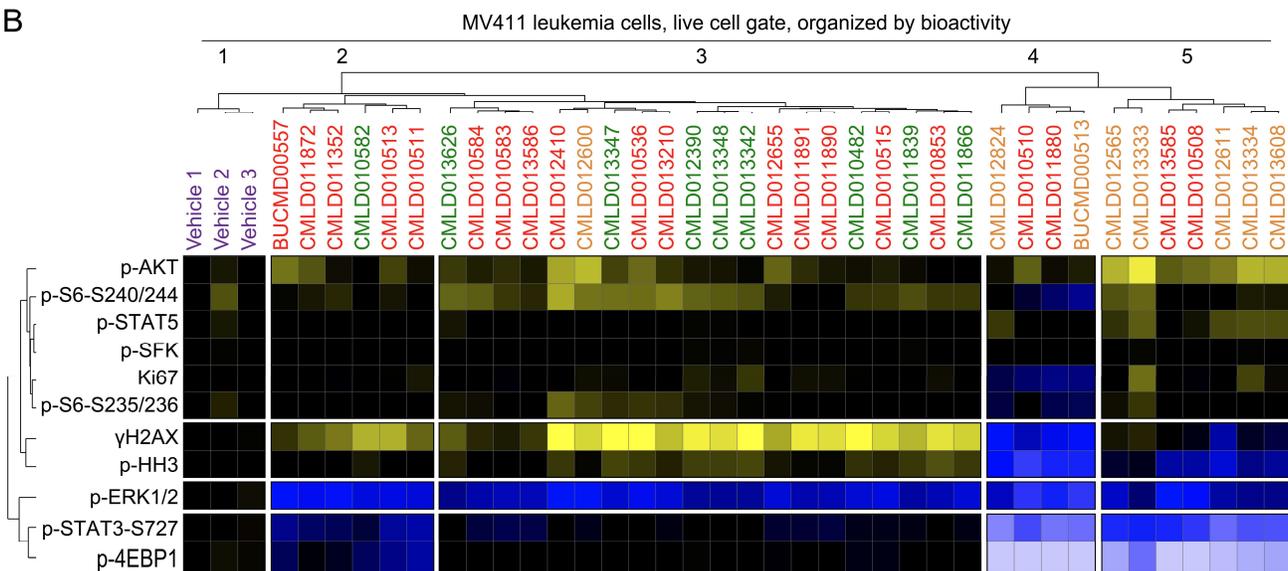


Figure 3 – Rocaglate subclasses had distinct patterns of bioactivity. A) Heatmap depicting the arcsinh ratio of the median fluorescence intensity for each compound (listed on top of heatmap) and readout (listed left of heatmap) by the median fluorescence intensity of Vehicle 1. Cells on the heatmap range from light blue for the lowest values to bright yellow for the highest values. Compounds are grouped and colored according to the rocaglate subclass listed on the top of the plot. **B)** Heatmap as in Figure 2B clustered according to a dendrogram of the transformed median fluorescence intensity for each compound and readout.

Thirman et. al. – Figure 4

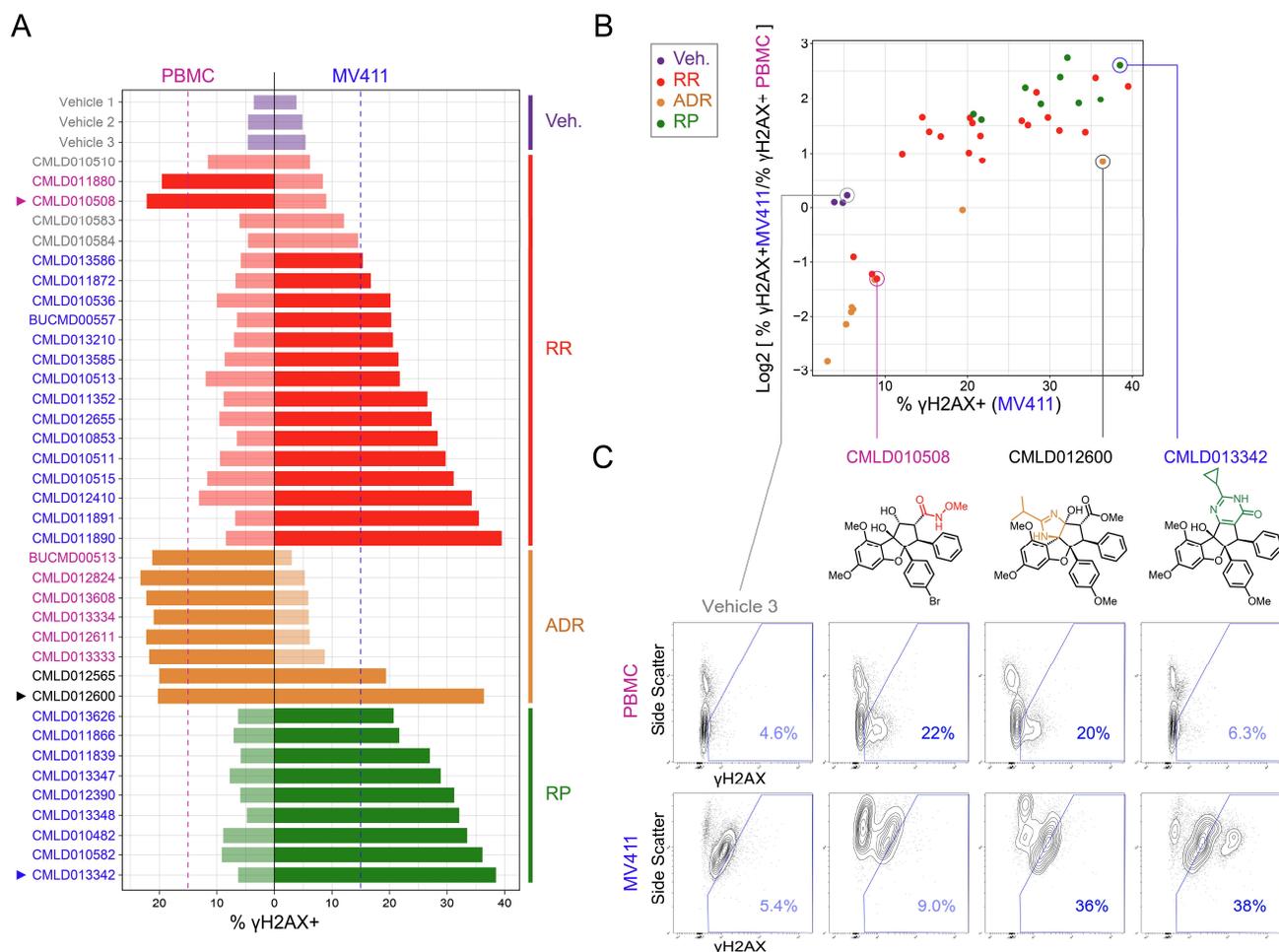


Figure 4 – MV411 specific activity was observed in the RP subfamily and some regular rocaglates. A) Bar plot of percentage of γ H2AX positive ($\% \gamma$ H2AX +) cells for each compound grouped and colored according to rocaglate subclass listed on the right side of the plot. Bars on the left side of the vertical black line correspond to the compound response in peripheral blood mononuclear cells (PBMC) and the right side corresponds to the compound response in MV411 cells. The dotted vertical blue line corresponds to a significance threshold for $\% \gamma$ H2AX+ cells in MV411 and the dotted pink line corresponds to a significance threshold for $\% \gamma$ H2AX+ cells in PBMC. Compounds that cross the threshold have darkened colored bars. The compound name (or vehicle) is listed on the far left side of plot colored according to the following system: blue = increases $\% \gamma$ H2AX+ cells past threshold in MV411 and not in PBMC, pink = increases $\% \gamma$ H2AX+ cells past threshold in PBMC and not in MV411, light grey = does not increase $\% \gamma$ H2AX+ cells past threshold in either cell type, dark grey = increases $\% \gamma$ H2AX+ cells past threshold in both cell types. Arrows are displayed to the left of compounds that will be shown in **Figure 4C**. **B)** Scatter plot of $\% \gamma$ H2AX+ cells in MV411 on the x-axis and the log_2 fold ratio of $\% \gamma$ H2AX+ cells in MV411 to $\% \gamma$ H2AX+ cells in PBMC on the y-axis. Each dot corresponds to a compound colored according to rocaglate structural subclass. Compounds that will be shown in **Figure 4C** are circled and labeled. **C)** Contour plots of Side Scatter vs. γ H2AX for Vehicle 3, CMLD010508, CMLD012600, and CMLD013342, respectively from left to right in PBMC (top) and MV411 (bottom). The blue line indicates the γ H2AX+ gate. The $\% \gamma$ H2AX+ cells within the gate is written in blue in the lower right corner. Corresponding structures for each compound are depicted below the compound name.

Thirman et. al. – Figure 5

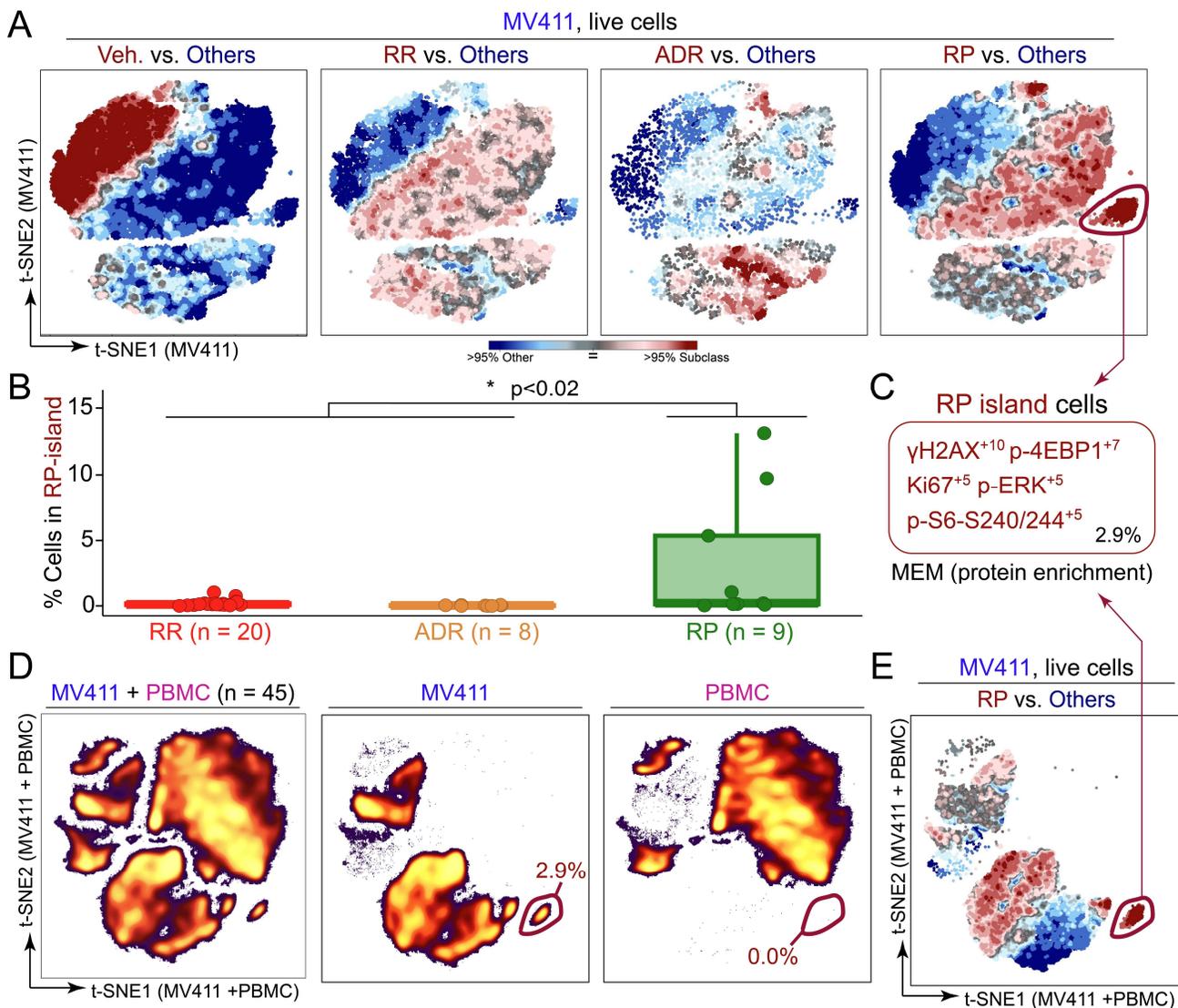


Figure 5 – The RP subfamily had a distinct, leukemia-specific population of cells with high DNA damage response. **A)** T-REX plots depict regions of significant difference between the t-SNE of one rocaglate subclass vs. the pooled set of all three remaining subclasses (including vehicle) for MV411 data. The RP island is circled in red in the “RP vs. Others” plot. **B)** Boxplot depicting the percent of cells in the RP island for MV411 for each rocaglate grouped by rocaglate subclass. The Wilcoxon rank sum test of RP vs. remaining subclass % in gate was conducted and generated a p-value of 0.015. **C)** MEM protein enrichment label generated for the RP island of cells. The percentage of the total population of cells for the RP island is depicted in the lower right corner. **D)** t-SNE plot of the MV411 and PBMC data on a new embedding with corresponding x and y axes (left). The t-SNE plot for both cell types was divided into separate plots for MV411 (middle) and PBMC (right). The percentage of cells in an MV411-specific island is circled in red and depicted in the lower right of the MV411 and PBMC plots. **E)** T-REX plot of RPs vs. the pooled set of all three remaining subclasses (including vehicle) for the MV411 data using the t-SNE performed on the MV411 and PBMC data together shown in **Figure 5D**. The MV411-specific island circled in red was identified to have the same MEM label as the RP island.

Thirman et. al. – Figure 6

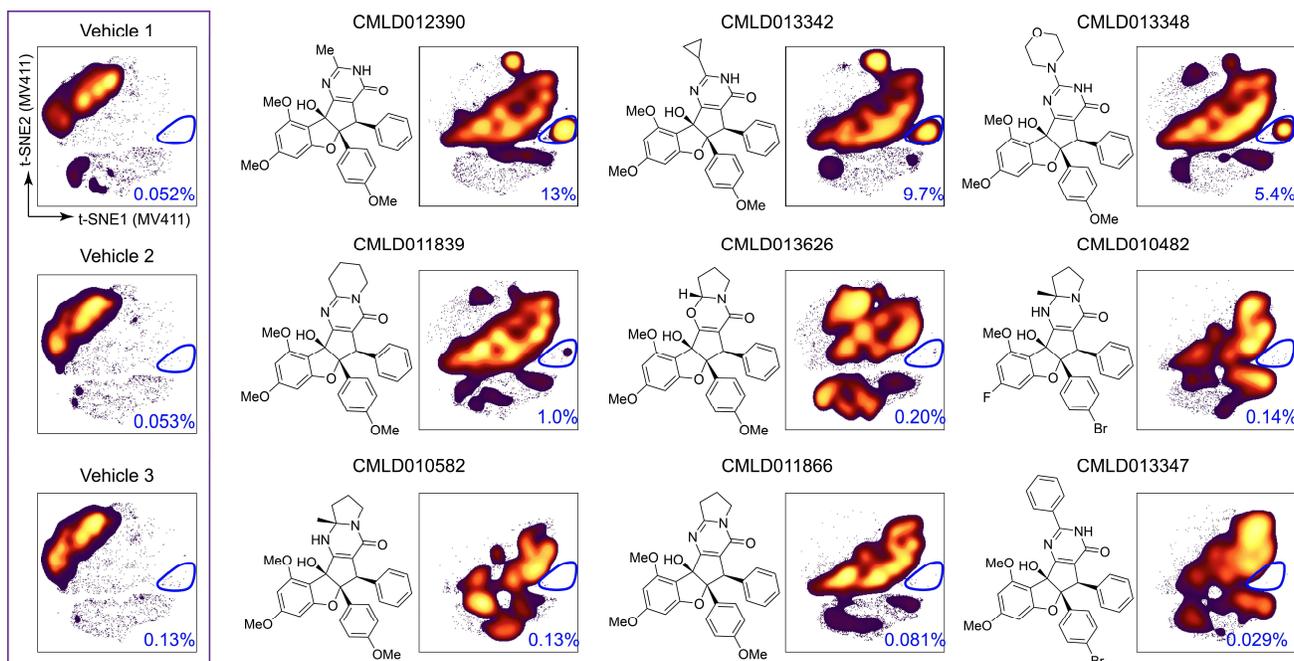


Figure 6 - Rocglates contributing to RP-island possessed structural commonalities. t-SNE shown in [Figure 2B](#) of all compounds in MV411 divided based on rocglate well of origin is shown for the set of 9 RPs in order of decreasing percentage of cells in the RP island (t-SNEs for all individual compounds are shown in [Supplementary Figure 2A](#)). The t-SNE plots for the three vehicle wells are shown for reference in the purple box on the left. The percentage of cells in the RP island gate circled in dark blue is included at the bottom right of each plot. The chemical structure corresponding to each rocglate is depicted on the left side of each t-SNE plot (except the three vehicle wells).

Thirman et. al. – Figure 7

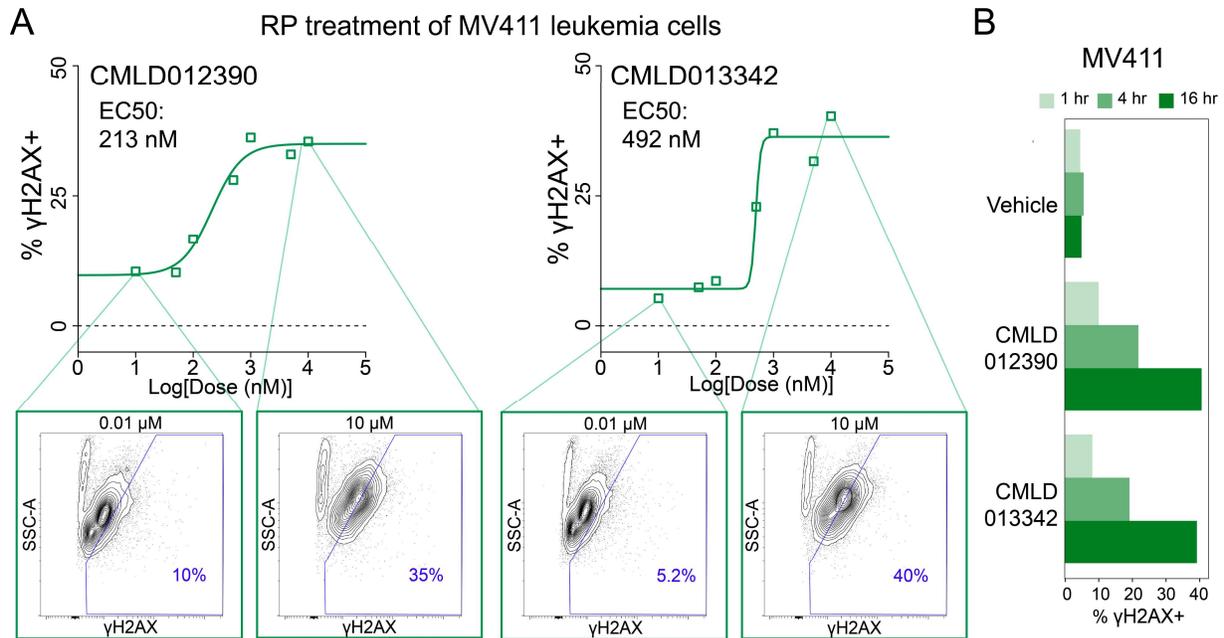


Figure 7 – The top 2 RPs activated γ H2AX with < 500nM potency within 4 hours. A) Dose-response titration curves depicting % γ H2AX+ cells (based on expert gating) vs. $\log[\text{dose}(\text{nM})]$. A half-maximal activating concentration (EC50) curve was fitted to the data and γ H2AX EC50 values are shown in the top left of each plot. Contour plots of Side Scatter Area (SSC-A) vs. γ H2AX at 10 μ M and 0.01 μ M, respectively, with 10% of cells per contour are shown for each compound. A polygon gate is drawn in blue to delineate γ H2AX+ cells, with the % γ H2AX+ cells within the gate depicted in the lower right corner. **B)** Bar plots depicting % γ H2AX+ cells derived from a time course experiment at 1, 4, and 16hr for Vehicle, **CMLD012390**, and **CMLD013342**.

References

- 1 Medina, S., Ihrle, R. A. & Irish, J. M. Learning cell identity in immunology, neuroscience, and cancer. *Seminars in Immunopathology* **45**, 3-16 (2023). <https://doi.org/10.1007/s00281-022-00976-y>
- 2 Irish, J. M. Beyond the age of cellular discovery. *Nature Immunology* **15**, 1095-1097 (2014). <https://doi.org/10.1038/ni.3034>
- 3 Bendall, S. C. *et al.* Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* **332**, 687-696 (2011). <https://doi.org/10.1126/science.1198704>
- 4 Krutzik, P. O., Irish, J. M., Nolan, G. P. & Perez, O. D. Analysis of protein phosphorylation and cellular signaling events by flow cytometry: techniques and clinical applications. *Clin Immunol* **110**, 206-221 (2004). <https://doi.org/10.1016/j.clim.2003.11.009>
- 5 Nolan, G. P. What's wrong with drug screening today. *Nature Chemical Biology* **3**, 187-191 (2007). <https://doi.org/10.1038/nchembio0407-187>
- 6 Krutzik, P. O., Crane, J. M., Clutter, M. R. & Nolan, G. P. High-content single-cell drug screening with phosphospecific flow cytometry. *Nature Chemical Biology* **4**, 132-142 (2008). <https://doi.org/10.1038/nchembio.2007.59>
- 7 Clutter, M. R., Krutzik, P. O. & Nolan, G. P. Phospho-specific flow cytometry in drug discovery. *Drug Discovery Today: Technologies* **2**, 295-302 (2005). [https://doi.org:https://doi.org/10.1016/j.ddtec.2005.08.010](https://doi.org/https://doi.org/10.1016/j.ddtec.2005.08.010)
- 8 Levine, Jacob H. *et al.* Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell* **162**, 184-197 (2015). [https://doi.org:https://doi.org/10.1016/j.cell.2015.05.047](https://doi.org/https://doi.org/10.1016/j.cell.2015.05.047)
- 9 Irish, J. M. & Doxie, D. B. in *High-Dimensional Single Cell Analysis: Mass Cytometry, Multi-parametric Flow Cytometry and Bioinformatic Techniques* (eds Harris G. Fienberg & Garry P. Nolan) 1-21 (Springer Berlin Heidelberg, 2014).
- 10 Krutzik, P. O., Clutter, M. R., Trejo, A. & Nolan, G. P. Fluorescent Cell Barcoding for Multiplex Flow Cytometry. *Current Protocols in Cytometry* **55**, 6.31.31-36.31.15 (2011). <https://doi.org:https://doi.org/10.1002/0471142956.cy0631s55>
- 11 Giudice, V., Feng, X., Kajigaya, S., Young, N. S. & Biancotto, A. Optimization and standardization of fluorescent cell barcoding for multiplexed flow cytometric phenotyping. *Cytometry Part A* **91**, 694-703 (2017). <https://doi.org:https://doi.org/10.1002/cyto.a.23162>
- 12 Manohar, S. *et al.* Combining fluorescent cell barcoding and flow cytometry-based phospho-ERK1/2 detection at short time scales in adherent cells. *Cytometry Part A* **95**, 192-200 (2019). <https://doi.org:https://doi.org/10.1002/cyto.a.23602>
- 13 Schares, H. A. M. *et al.* in *Methods in Cell Biology* (Academic Press, 2023).
- 14 Giudice, V., Feng, X., Kajigaya, S., Young, N. S. & Biancotto, A. Optimization and standardization of fluorescent cell barcoding for multiplexed flow cytometric phenotyping. *Cytometry. Part A : the journal of the International Society for Analytical Cytology* **91**, 694-703 (2017). <https://doi.org/10.1002/cyto.a.23162>
- 15 Krutzik, P. O., Clutter, M. R., Trejo, A. & Nolan, G. P. Fluorescent cell barcoding for multiplex flow cytometry. *Current protocols in cytometry / editorial board, J. Paul Robinson, managing editor ... [et al.] Chapter 6*, Unit 6 31 (2011). <https://doi.org/10.1002/0471142956.cy0631s55>
- 16 Bodenmiller, B. *et al.* Multiplexed mass cytometry profiling of cellular states perturbed by small-molecule regulators. *Nature biotechnology* **30**, 858-867 (2012). <https://doi.org/10.1038/nbt.2317>
- 17 Manohar, S. *et al.* Combining fluorescent cell barcoding and flow cytometry-based phospho-ERK1/2 detection at short time scales in adherent cells. *Cytometry. Part A : the journal of the International Society for Analytical Cytology* **95**, 192-200 (2019). <https://doi.org/10.1002/cyto.a.23602>
- 18 Earl, D. C. *et al.* Discovery of human cell selective effector molecules using single cell multiplexed activity metabolomics. *Nat Commun* **9**, 39 (2018). <https://doi.org/10.1038/s41467-017-02470-8>
- 19 Frischbutter, S., Schultheis, K., Patzel, M., Radbruch, A. & Baumgrass, R. Evaluation of calcineurin/NFAT inhibitor selectivity in primary human Th cells using bar-coding and phospho-flow cytometry. *Cytometry. Part A : the journal of the International Society for Analytical Cytology* **81**, 1005-1011 (2012). <https://doi.org/10.1002/cyto.a.22204>
- 20 Peter, S. C. *et al.* in *Encyclopedia of Bioinformatics and Computational Biology* (eds Shoba Ranganathan, Michael Gribskov, Kenta Nakai, & Christian Schönbach) 661-676 (Academic Press, 2019).
- 21 Amir, E.-a. D. *et al.* viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. *Nature Biotechnology* **31**, 545-552 (2013). <https://doi.org/10.1038/nbt.2594>
- 22 Van der Maaten, L. & Hinton, G. Visualizing data using t-SNE. *Journal of machine learning research* **9** (2008).

- 23 Belkina, A. C. *et al.* Automated optimized parameters for T-distributed stochastic neighbor embedding improve visualization and analysis of large datasets. *Nature Communications* **10**, 5415 (2019). <https://doi.org/10.1038/s41467-019-13055-y>
- 24 Becht, E. *et al.* Dimensionality reduction for visualizing single-cell data using UMAP. *Nature Biotechnology* **37**, 38-44 (2019). <https://doi.org/10.1038/nbt.4314>
- 25 McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. (2018). <https://doi.org/10.48550/arxiv.1802.03426>
- 26 Kuchroo, M. *et al.* Multiscale PHATE identifies multimodal signatures of COVID-19. *Nature Biotechnology* **40**, 681-691 (2022). <https://doi.org/10.1038/s41587-021-01186-x>
- 27 Barone, S. M. *et al.* Unsupervised machine learning reveals key immune cell subsets in COVID-19, rhinovirus infection, and cancer therapy. *eLife* **10**, e64653 (2021). <https://doi.org/10.7554/eLife.64653>
- 28 Diggins, K. E., Greenplate, A. R., Leelatian, N., Wogsland, C. E. & Irish, J. M. Characterizing cell subsets using marker enrichment modeling. *Nature Methods* **14**, 275-278 (2017). <https://doi.org/10.1038/nmeth.4149>
- 29 Kuralt, V. & Frlan, R. Navigating the Chemical Space of ENR Inhibitors: A Comprehensive Analysis. *Antibiotics (Basel)* **13** (2024). <https://doi.org/10.3390/antibiotics13030252>
- 30 Boby, M. L. *et al.* Open science discovery of potent noncovalent SARS-CoV-2 main protease inhibitors. *Science* **382**, eabo7201 (2023). <https://doi.org/10.1126/science.abo7201>
- 31 Scott, D. E., Coyne, A. G., Hudson, S. A. & Abell, C. Fragment-based approaches in drug discovery and chemical biology. *Biochemistry* **51**, 4990-5003 (2012). <https://doi.org/10.1021/bi3005126>
- 32 Lu King, M. *et al.* X-Ray crystal structure of rocaglamide, a novel antileulemic 1H-cyclopenta[b]benzofuran from *Aglaia elliptifolia*. *Journal of the Chemical Society, Chemical Communications*, 1150-1151 (1982). <https://doi.org/10.1039/C39820001150>
- 33 Gerard, B., Jones, G. & Porco, J. A. A Biomimetic Approach to the Rocaglamides Employing Photogeneration of Oxidopyryliums Derived from 3-Hydroxyflavones. *Journal of the American Chemical Society* **126**, 13620-13621 (2004). <https://doi.org/10.1021/ja044798o>
- 34 Zhang, W. *et al.* Chemical Synthesis Enables Structural Reengineering of Aglaroxin C Leading to Inhibition Bias for Hepatitis C Viral Infection. *Journal of the American Chemical Society* **141**, 1312-1323 (2019). <https://doi.org/10.1021/jacs.8b11477>
- 35 Zhang, W. *et al.* Intercepted Retro-Nazarov Reaction: Syntheses of Amidino-Rocaglate Derivatives and Their Biological Evaluation as eIF4A Inhibitors. *Journal of the American Chemical Society* **141**, 12891-12900 (2019). <https://doi.org/10.1021/jacs.9b06446>
- 36 Chu, J. *et al.* Rocaglates Induce Gain-of-Function Alterations to eIF4A and eIF4F. *Cell Reports* **30**, 2481-2488.e2485 (2020). <https://doi.org/10.1016/j.celrep.2020.02.002>
- 37 Chu, J. *et al.* Amidino-Rocaglates: A Potent Class of eIF4A Inhibitors. *Cell Chemical Biology* **26**, 1586-1593.e1583 (2019). <https://doi.org/10.1016/j.chembiol.2019.08.008>
- 38 Krutzik, P. O., Crane, J. M., Clutter, M. R. & Nolan, G. P. High-content single-cell drug screening with phosphospecific flow cytometry. *Nature chemical biology* **4**, 132-142 (2008). <https://doi.org/10.1038/nchembio.2007.59>
- 39 Irish, J. M. *et al.* Single cell profiling of potentiated phospho-protein networks in cancer cells. *Cell* **118**, 217-228 (2004). <https://doi.org/10.1016/j.cell.2004.06.028>
- 40 Irish, J. M., Czerwinski, D. K., Nolan, G. P. & Levy, R. Altered B-cell receptor signaling kinetics distinguish human follicular lymphoma B cells from tumor-infiltrating nonmalignant B cells. *Blood* **108**, 3135-3142 (2006). <https://doi.org/10.1182/blood-2006-02-003921>
- 41 Irish, J. M. *et al.* B-cell signaling networks reveal a negative prognostic human lymphoma cell subset that emerges during tumor progression. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 12747-12754 (2010). <https://doi.org/10.1073/pnas.1002057107>
- 42 Bartkowiak, T. *et al.* An immunosuppressed microenvironment distinguishes lateral ventricle-contacting glioblastomas. *JCI Insight* **8** (2023). <https://doi.org/10.1172/jci.insight.160652>
- 43 Greenplate, A. R. *et al.* Computational Immune Monitoring Reveals Abnormal Double-Negative T Cells Present across Human Tumor Types. *Cancer Immunol Res* **7**, 86-99 (2019). <https://doi.org/10.1158/2326-6066.CIR-17-0692>
- 44 Earl, D. C. *et al.* Discovery of human cell selective effector molecules using single cell multiplexed activity metabolomics. *Nature Communications* **9**, 39 (2018). <https://doi.org/10.1038/s41467-017-02470-8>
- 45 Uxa, S. *et al.* Ki-67 gene expression. *Cell Death & Differentiation* **28**, 3357-3370 (2021). <https://doi.org/10.1038/s41418-021-00823-x>

- 46 Miller, I. *et al.* Ki67 is a Graded Rather than a Binary Marker of Proliferation versus Quiescence. *Cell Rep* **24**, 1105-1112.e1105 (2018). <https://doi.org/10.1016/j.celrep.2018.06.110>
- 47 Magnuson, B., Ekim, B. & Fingar, D. C. Regulation and function of ribosomal protein S6 kinase (S6K) within mTOR signalling networks. *Biochem J* **441**, 1-21 (2012). <https://doi.org/10.1042/bj20110892>
- 48 Roux, P. P. *et al.* RAS/ERK Signaling Promotes Site-specific Ribosomal Protein S6 Phosphorylation via RSK and Stimulates Cap-dependent Translation *. *Journal of Biological Chemistry* **282**, 14056-14064 (2007). <https://doi.org/10.1074/jbc.M700906200>
- 49 Mah, L. J., El-Osta, A. & Karagiannis, T. C. γ H2AX: a sensitive molecular marker of DNA damage and repair. *Leukemia* **24**, 679-686 (2010). <https://doi.org/10.1038/leu.2010.6>
- 50 Igelmann, S., Neubauer, H. A. & Ferbeyre, G. STAT3 and STAT5 Activation in Solid Cancers. *Cancers (Basel)* **11** (2019). <https://doi.org/10.3390/cancers11101428>
- 51 Wingelhofer, B. *et al.* Pharmacologic inhibition of STAT5 in acute myeloid leukemia. *Leukemia* **32**, 1135-1146 (2018). <https://doi.org/10.1038/s41375-017-0005-9>
- 52 Kotecha, N. *et al.* Single-cell profiling identifies aberrant STAT5 activation in myeloid malignancies with specific clinical and biologic correlates. *Cancer cell* **14**, 335-343 (2008). <https://doi.org/10.1016/j.ccr.2008.08.014>
- 53 Roskoski, R., Jr. ERK1/2 MAP kinases: structure, function, and regulation. *Pharmacol Res* **66**, 105-143 (2012). <https://doi.org/10.1016/j.phrs.2012.04.005>
- 54 Mendoza, M. C., Er, E. E. & Blenis, J. The Ras-ERK and PI3K-mTOR pathways: cross-talk and compensation. *Trends Biochem Sci* **36**, 320-328 (2011). <https://doi.org/10.1016/j.tibs.2011.03.006>
- 55 Hans, F. & Dimitrov, S. Histone H3 phosphorylation and cell division. *Oncogene* **20**, 3021-3027 (2001). <https://doi.org/10.1038/sj.onc.1204326>
- 56 Gingras, A. C. *et al.* Regulation of 4E-BP1 phosphorylation: a novel two-step mechanism. *Genes Dev* **13**, 1422-1437 (1999). <https://doi.org/10.1101/gad.13.11.1422>
- 57 Takata, M. *et al.* Tyrosine kinases Lyn and Syk regulate B cell receptor-coupled Ca²⁺ mobilization through distinct pathways. *The EMBO Journal* **13**, 1341-1349-1349 (1994). [https://doi.org:https://doi.org/10.1002/j.1460-2075.1994.tb06387.x](https://doi.org/https://doi.org/10.1002/j.1460-2075.1994.tb06387.x)
- 58 Blix, E. S. *et al.* Phospho-specific flow cytometry identifies aberrant signaling in indolent B-cell lymphoma. *BMC Cancer* **12**, 478 (2012). <https://doi.org/10.1186/1471-2407-12-478>
- 59 Geben, L. C. *et al.* Dephosphorylation of 4EBP1/2 Induces Prenatal Neural Stem Cell Quiescence. *bioRxiv*, 2023.2002.2014.528513 (2023). <https://doi.org/10.1101/2023.02.14.528513>
- 60 Balsamo, J. A. *et al.* An immunogenic cell injury module for the single-cell multiplexed activity metabolomics platform to identify promising anti-cancer natural products. *J Biol Chem* **298**, 102300 (2022). <https://doi.org/10.1016/j.jbc.2022.102300>
- 61 Chu, J., Cencic, R., Wang, W., Porco, J. A., Jr & Pelletier, J. Translation Inhibition by Rocaglates Is Independent of eIF4E Phosphorylation Status. *Molecular Cancer Therapeutics* **15**, 136-141 (2016). <https://doi.org/10.1158/1535-7163.Mct-15-0409>
- 62 Chu, J. *et al.* CRISPR-Mediated Drug-Target Validation Reveals Selective Pharmacological Inhibition of the RNA Helicase, eIF4A. *Cell Reports* **15**, 2340-2347 (2016). [https://doi.org:https://doi.org/10.1016/j.celrep.2016.05.005](https://doi.org/https://doi.org/10.1016/j.celrep.2016.05.005)
- 63 Manier, S. *et al.* Inhibiting the oncogenic translation program is an effective therapeutic strategy in multiple myeloma. *Sci Transl Med* **9** (2017). <https://doi.org/10.1126/scitranslmed.aal2668>
- 64 Zhang, X. *et al.* Targeting translation initiation by synthetic rocaglates for treating MYC-driven lymphomas. *Leukemia* **34**, 138-150 (2020). <https://doi.org/10.1038/s41375-019-0503-z>
- 65 Lehman, S. L. *et al.* Inhibition of the Translation Initiation Factor eIF4A Enhances Tumor Cell Radiosensitivity. *Mol Cancer Ther* **21**, 1406-1414 (2022). <https://doi.org/10.1158/1535-7163.Mct-22-0037>
- 66 Naineni, S. K. *et al.* Exploring the targeting spectrum of rocaglates among eIF4A homologs. *RNA (Cambridge)* **29**, 826-835 (2023). <https://doi.org/10.1261/rna.079318.122>
- 67 Lindqvist, L. M. *et al.* Translation inhibitors induce cell death by multiple mechanisms and Mcl-1 reduction is only a minor contributor. *Cell Death Dis* **3**, e409 (2012). <https://doi.org/10.1038/cddis.2012.149>
- 68 Kokpol, U., Venaskulchai, B., Simpson, J. & Weavers, R. T. Isolation and X-ray structure determination of a novel pyrimidinone from *Aglaia odorata*. *Journal of the Chemical Society, Chemical Communications*, 773-774 (1994). <https://doi.org/10.1039/C39940000773>
- 69 Wang, S.-K. & Duh, C.-Y. Cytotoxic Cyclopenta[b]benzofuran Derivatives from the Stem Bark of *Aglaia formosana*. *Planta Med* **67**, 555-557 (2001). <https://doi.org/10.1055/s-2001-16471>

- 70 Wu, P.-F. *et al.* Three New Aglain Derivatives from *Aglaia odorata* Lour. and Their Cytotoxic Activities. *Chemistry & Biodiversity* **19**, e202101008 (2022). [https://doi.org:https://doi.org/10.1002/cbdv.202101008](https://doi.org/10.1002/cbdv.202101008)
- 71 Levi-Galibov, O. *et al.* Heat Shock Factor 1-dependent extracellular matrix remodeling mediates the transition from chronic intestinal inflammation to colon cancer. *Nat Commun* **11**, 6245 (2020). [https://doi.org:10.1038/s41467-020-20054-x](https://doi.org/10.1038/s41467-020-20054-x)
- 72 Bonner, W. M. *et al.* GammaH2AX and cancer. *Nat Rev Cancer* **8**, 957-967 (2008). [https://doi.org:10.1038/nrc2523](https://doi.org/10.1038/nrc2523)
- 73 Chen, M. *et al.* Dual targeting of DDX3 and eIF4A by the translation inhibitor rocaglamide A. *Cell Chemical Biology* **28**, 475-486.e478 (2021). [https://doi.org:10.1016/j.chembiol.2020.11.008](https://doi.org/10.1016/j.chembiol.2020.11.008)
- 74 Iwasaki, S. *et al.* The Translation Inhibitor Rocaglamide Targets a Bimolecular Cavity between eIF4A and Polypurine RNA. *Molecular Cell* **73**, 738-748.e739 (2019). [https://doi.org:10.1016/j.molcel.2018.11.026](https://doi.org/10.1016/j.molcel.2018.11.026)
- 75 Hernández-Díaz, T., Valiente-Echeverría, F. & Soto-Rifo, R. RNA Helicase DDX3: A Double-Edged Sword for Viral Replication and Immune Signaling. *Microorganisms* **9** (2021). [https://doi.org:10.3390/microorganisms9061206](https://doi.org/10.3390/microorganisms9061206)
- 76 Greenplate, A. R. *et al.* Myelodysplastic Syndrome Revealed by Systems Immunology in a Melanoma Patient Undergoing Anti-PD-1 Therapy. *Cancer Immunol Res* **4**, 474-480 (2016). [https://doi.org:10.1158/2326-6066.CCR-15-0213](https://doi.org/10.1158/2326-6066.CCR-15-0213)
- 77 Leelatian, N. *et al.* Unsupervised machine learning reveals risk stratifying glioblastoma tumor cells. *eLife* **9**, e56879 (2020). [https://doi.org:10.7554/eLife.56879](https://doi.org/10.7554/eLife.56879)
- 78 Diggins, K. E., Gandelman, J. S., Roe, C. E. & Irish, J. M. Generating Quantitative Cell Identity Labels with Marker Enrichment Modeling (MEM). *Curr Protoc Cytom* **83**, 10.21.11-10.21.28 (2018). [https://doi.org:10.1002/cpcy.34](https://doi.org/10.1002/cpcy.34)