

# Multomics Profiling of T-cell Leukemia and Lymphoma Enables Targeted Therapeutic Discovery

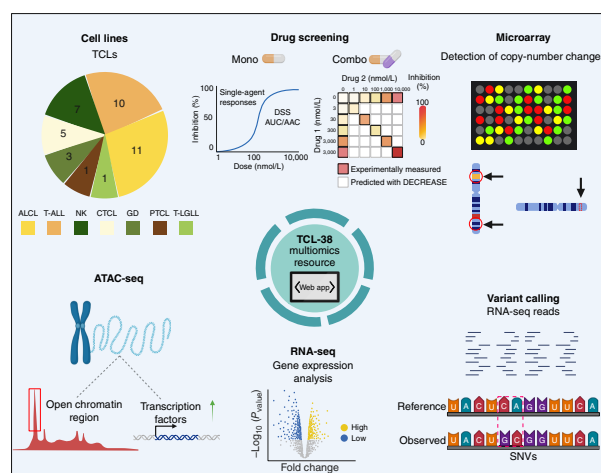
Aleksandr Ianevski<sup>1,2</sup>, Kristen Nader<sup>1,2</sup>, Julia Nguyen<sup>3,4</sup>, Helena Sorger<sup>5,6</sup>, Sanna Timonen<sup>1,7,8</sup>, Edith Julia<sup>9</sup>, Daniel Pölöske<sup>6</sup>, Katrin Spirk<sup>6</sup>, Christina Wagner<sup>6</sup>, Dennis Jungherz<sup>10,11</sup>, Minoru Nakano<sup>3,4</sup>, Sisira Kadambat Nair<sup>3,4</sup>, Philipp Ianevski<sup>1</sup>, Matti Kankainen<sup>7,8,12</sup>, Diogo Dias<sup>1,2,7,8</sup>, Anna Cichońska<sup>13</sup>, Tea Pemovska<sup>14</sup>, Christine Pirker<sup>15,16</sup>, Walter Berger<sup>15,16</sup>, Till Braun<sup>11,17</sup>, Richard Moriggl<sup>6,18</sup>, Emmanuel Bachy<sup>9</sup>, Satu Mustjoki<sup>2,7,8</sup>, Marco Herling<sup>10,11,19</sup>, Heidi A. Neubauer<sup>6</sup>, Benjamin Haibe-Kains<sup>3,4,20,21</sup>, and Tero Aittokallio<sup>1,2,22,23</sup>



## ABSTRACT

T-cell leukemias and lymphomas (TCL) form a heterogeneous group of rare and often aggressive malignancies. Because of the rarity and heterogeneity of TCL subtypes, clinical trials are challenging to conduct, making pharmacogenomic studies in cell line panels critical for the discovery of targeted therapeutics. The scarcity of data repositories with integrated multiomics and drug screening data hinders the preclinical evaluation of drug vulnerabilities and the identification of molecular markers predictive of responses to monotherapies and combinations. To address this gap, we conducted comprehensive pharmacogenomic profiling on a panel of 38 TCL cell lines, representing major clinical TCL subtypes to capture the molecular and phenotypic diversity. The TCL-38 multiomics data resource includes harmonized genetic, molecular, and epigenetic profiles, with comprehensive annotations and standardized drug response assessment of each cell line. This resource, together with machine learning predictions, was leveraged to identify TCL subtype-specific therapeutic vulnerabilities, including single-agent sensitivities and synergistic drug combinations, which were linked to genetic or epigenetic features as potential predictive biomarkers. This integrated and openly available resource (<https://aittokallio.group/tcl38>) could help advance the currently limited treatment options for patients with TCL.

**Significance:** Integrated and harmonized multiomics analyses and drug screening across a heterogeneous panel of T-cell leukemias and lymphomas provide a resource to uncover drug targets and predictive biomarkers to improve patient outcomes.



Created in BioRender. Dias, D. (2025) <https://BioRender.com/7062cja>

<sup>1</sup>Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland. <sup>2</sup>ICAN Digital Precision Cancer Medicine Flagship, University of Helsinki and Helsinki University Hospital, Helsinki, Finland. <sup>3</sup>Princess Margaret Cancer Centre, University Health Network, Toronto, Canada. <sup>4</sup>Department of Medical Biophysics, University of Toronto, Toronto, Canada. <sup>5</sup>Department of Pediatric and Adolescent Surgery, Medical University of Vienna, Vienna, Austria. <sup>6</sup>Centre for Biological Sciences, University of Veterinary Medicine Vienna, Vienna, Austria. <sup>7</sup>Hematology Research Unit Helsinki, University of Helsinki and Helsinki University Hospital Comprehensive Cancer Center, Helsinki, Finland. <sup>8</sup>Translational Immunology Research Program and Department of Clinical Chemistry and Hematology, University of Helsinki, Helsinki, Finland. <sup>9</sup>Lymphoma Immuno Biology Team, Equipe Labellisée LIGUE 2023, Centre International de Recherche en Infectiologie, INSERM U1111-CNRS UMR5308, Faculté de Médecine Lyon-Sud, Hospices Civils de Lyon, Université Claude Bernard Lyon I-ENS de Lyon, Lyon, France. <sup>10</sup>Department of Hematology, Cellular Therapy, Hemostaseology, and Infectious Diseases, University of Leipzig Medical Center, Leipzig, Germany. <sup>11</sup>Department I of Internal Medicine, Center for Integrated Oncology, Aachen-Bonn-Cologne-Duesseldorf, University of Cologne, Cologne, Germany. <sup>12</sup>Laboratory of Genetics, HUS Diagnostic Center, Helsinki and Uusimaa Hospital District, Helsinki, Finland. <sup>13</sup>Harmonic Discovery Inc., New York City, New York. <sup>14</sup>Division of Hematology and Hemostaseology, Department of Medicine I, Medical University of Vienna,

Vienna, Austria. <sup>15</sup>Centre for Cancer Research, Medical University of Vienna, Vienna, Austria. <sup>16</sup>Comprehensive Cancer Center, Medical University of Vienna, Vienna, Austria. <sup>17</sup>Mildred Scheel School of Oncology, Aachen Bonn Cologne Düsseldorf (MSSO ABCD), Faculty of Medicine and University Hospital of Cologne, Cologne, Germany. <sup>18</sup>Department of Biosciences and Medical Biology, Paris Lodron University of Salzburg, Salzburg, Austria. <sup>19</sup>Comprehensive Cancer Center Central Germany, Leipzig-Jena, Germany. <sup>20</sup>Vector Institute for Artificial Intelligence, Toronto, Canada. <sup>21</sup>Structural Genomics Consortium, Toronto, Canada. <sup>22</sup>Department of Cancer Genetics, Institute for Cancer Research, Oslo University Hospital, Oslo, Norway. <sup>23</sup>Oslo Centre for Biostatistics and Epidemiology (OCBE), Faculty of Medicine, University of Oslo, Oslo, Norway.

**Corresponding Author:** Tero Aittokallio, Institute for Molecular Medicine Finland, FIMM, P.O. Box 20 (Tukholmantu 8), University of Helsinki, Helsinki FI-00014, Finland. E-mail: [tero.aittokallio@helsinki.fi](mailto:tero.aittokallio@helsinki.fi)

Cancer Res 2026;86:295-309

doi: 10.1158/0008-5472.CAN-25-0881

This open access article is distributed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) license.

©2025 The Authors; Published by the American Association for Cancer Research

## Introduction

T-cell leukemias and lymphomas (TCL) comprise a heterogeneous group of rare and often aggressive non-Hodgkin lymphomas that originate from T lymphocytes at various stages of maturation. TCLs account for approximately 10% to 15% of all non-Hodgkin lymphomas in the Western hemisphere (1, 2). Each TCL subtype of the current World Health Organization classification is considered a distinct clinicopathologic entity, defined by unique presentation and natural history as well as molecular abnormalities (3). TCL diversity originates from the heterogeneity of T- and NK-cell populations from which they originate. Immature T-cell neoplasms, also known as precursor T-cell neoplasms, include T-cell acute lymphoblastic leukemia (T-ALL); however, we note that maturation status represents a continuum rather than discrete categories. Mature T-cell neoplasms include primary nodal, such as peripheral T-cell lymphoma, not otherwise specified (PTCL-NOS), and anaplastic large cell lymphoma (ALCL); primary extranodal, such as extranodal NK-/T-cell lymphoma-nasal type; primary cutaneous (CTCL), such as mycosis fungoides/Sézary syndrome or cutaneous gamma-delta ( $\gamma\delta$ ) T-cell lymphoma; and leukemic forms, such as T-large granular lymphocytic leukemia (T-LGL) or T-prolymphocytic leukemia (T-PLL). TCLs include both leukemic forms, such as T-PLL, and lymphomatous forms, which together represent the spectrum of T-cell malignancies.

Current treatment options for TCLs have only recently started to become more specific to subtypes. Poly-chemotherapy regimens, such as CHOP (cyclophosphamide, doxorubicin, vincristine, and prednisone), still represent the mainstay of initial therapies, but more targeted therapies have in part supplemented these conventional backbones or have replaced them in relapsed/refractory TCLs. Examples are the anti-CD30 agent brentuximab vedotin, histone deacetylase inhibitors (e.g., romidepsin, belinostat, vorinostat, and tucidinostat), and monoclonal antibodies (e.g., CCR4 inhibitor mogamulizumab and anti-CD52 agent alemtuzumab). The roles of autologous and allogeneic stem cell transplantation for eligible patients are still not entirely defined (4–7). Despite these therapeutic advancements, patient outcomes in most TCLs remain poor (5, 8–10). The rarity and heterogeneity of TCL subtypes pose significant hurdles in conducting sufficiently powered clinical trials for novel treatments in defined TCL subsets. In particular, there is a medical need for a continued discovery and preclinical evaluation of effective and TCL subtype-specific combinations of approved or novel substances, but testing the combinatorial effects is challenged by limited resources of available, well-annotated patient material and representative model systems.

Pharmacogenomic profiling of cancer cell lines has led to the discovery of clinically relevant therapeutic targets, drug combinations, and predictive biomarkers using panels of molecularly characterized systems in various cancer types (11–17). For the rare TCL subtypes, well-characterized cell line panels would offer a valuable platform to identify new targeted treatments and vulnerabilities with companion biomarkers extracted from omics data. Only a few drug screening studies have been performed in TCL cell lines (Supplementary Table S1), and integrating omics data and drug sensitivity profiles from multiple studies is challenging because of technical variations and differences in molecular profiling and screening methodologies. Databases such as DepMap (18), CellMinerCDB (19), and PharmacDB (20, 21) integrate data from multiple pharmacogenomic studies, but these include only a few TCL subtypes with drug response profiles and lack comprehensive, standardized

omics profiles. Similarly, drug combination data resources, such as DrugComb (22, 23) and SYNERGxDB (24), contain only minimal TCL drug combination response data (Supplementary Table S1).

To address these limitations, we established the TCL-38 pharmacogenomic resource, the largest to date and most comprehensive panel of TCL cell lines, featuring harmonized multiomics data, comprehensive annotations, and standardized drug response assessments. Here, we describe the development and use of the TCL-38 resource and an interactive web application (<https://aittokallio.group/tcl38>). In this report, we demonstrate its utility in identifying TCL subtype-specific vulnerabilities, including single-agent sensitivities and synergistic drug combinations. We present several case studies in which novel sensitivities are linked to (epi)genetic features, demonstrating that the TCL-38 multiomics resource has sufficient power for biomarker discovery. Additionally, we show how the pharmacogenomic resource can be used as a unique training dataset for machine learning (ML) models for predicting novel single-drug or drug combination responses that have not yet been experimentally tested across the heterogeneous TCL subtypes.

## Materials and Methods

### Cell culture

MyLa (RRID: CVCL\_2633) and SeAx (RRID: CVCL\_5363) cell lines were a kind gift from Dr. Hinrich Abken. Mac-1 (RRID: CVCL\_H631) and Mac-2a (RRID: CVCL\_H637) cell lines were a kind gift from Dr. Marshall Kadin. TLBR-1 (RRID: CVCL\_L177), TLBR-2 (RRID: CVCL\_A1EY), OCI-Ly13.2 (RRID: CVCL\_8797), DL-40 (RRID: CVCL\_2889), SMZ-1 (RRID: CVCL\_RL84), MOTN-1 (RRID: CVCL\_2127), NKL (RRID: CVCL\_0466), and MTA (RRID: CVCL\_3032) cell lines were a kind gift from Dr. Raphael Koch. The SUP-T13 (RRID: CVCL\_9973) cell line was a kind gift from Dr. Stephen D. Smith. The NK-YS (RRID: CVCL\_8461) and SNK-6 (RRID: CVCL\_A673) cell lines were a kind gift from Dr. John Chan. The KOPT-K1 (RRID: CVCL\_4965) cell line was a kind gift from Dr. Koshi Akahane. Jurkat cells (RRID: CVCL\_0065) were a kind gift from Dr. Florian Grebien. The Karpas 299 (RRID: CVCL\_1324) cell line was a kind gift from Dr. Anna-Lena Illert. The DEL (RRID: CVCL\_1170) cell line was a kind gift from Dr. Georgios Rassidakis. DND-41 (RRID: CVCL\_2022), CCRF-H-SB2 (RRID: CVCL\_1859), ALL-SIL (RRID: CVCL\_1805), Peer (RRID: CVCL\_1913), and Loucy (RRID: CVCL\_1380) cell lines were a kind gift from Dr. A. Thomas Look. H9 (RRID: CVCL\_1240), HH (RRID: CVCL\_1280), SU-DHL-1 (RRID: CVCL\_0538), SUP-M2 (RRID: CVCL\_2209), L82 (RRID: CVCL\_2098), SUP-T11 (RRID: CVCL\_2210), DERL-2 (RRID: CVCL\_2016), DERL-7 (RRID: CVCL\_2017), YT (RRID: CVCL\_1797), KHYG-1 (RRID: CVCL\_2976), and NK-92 (RRID: CVCL\_2142) cell lines were obtained from the German Collection of Microorganisms and Cell Cultures GmbH (DSMZ). HUT-78 (RRID: CVCL\_0337) cells were purchased from CLS Cell Lines Service GmbH. Karpas 384 (RRID: CVCL\_2541) cells were obtained from the European Collection of Cell Cultures. MOLT-4 (RRID: CVCL\_0013) cells were purchased from the ATCC. TCL-38 cell lines were cultured in media containing RPMI-1640 supplemented with 10% heat-inactivated FBS, 10 U/mL penicillin/streptomycin, and 2 mmol/L L-glutamine. The cell lines TLBR-1, NK-92, and NKL were additionally grown in the presence of 10 ng/mL IL2, TLBR2 in 5 ng/mL IL2, DERL-2, DERL-7, NK-YS, SNK-6, KHYG-1, YT, and MOTN-1 in 2.5 ng/mL IL2, and SeAx cells were grown with 5 ng/mL IL2 and IL4. All cell lines were incubated at 37°C with 5% CO<sub>2</sub> in a humidified environment and

were regularly tested and confirmed negative for *Mycoplasma* using PhoenixDx Mycoplasma Mix (Procomcore Biotech). The authenticity of all TCL-38 cell lines was confirmed by short tandem repeat profiling (Microsynth GmbH). Upon thawing cryopreserved cell lines, the cells were kept in culture for no more than 2 months. We note that neither human T-cell lymphotropic virus nor ATLL cell lines were included in the study.

### Single-agent testing

The 172 compounds tested across the 38 cell lines are listed in Supplementary Table S3. The compounds were plated in five concentrations in varying concentration ranges, centered around a compound-specific relevant cellular activity concentration on black clear-bottom 384-well plates (Corning, #3712), using an Echo 550 Liquid Handler (Labcyte). As negative and positive controls, we used 0.1% DMSO and 100  $\mu\text{mol/L}$  benzethonium chloride, respectively. The predrugged plates were stored in pressurized Storage-Pods (Roylan Developments Ltd.) filled with inert nitrogen gas. All subsequent liquid handling was performed using a MultiFlo FX dispenser (BioTek). After 72-hour incubation, cell viability was measured with CellTiter-Glo (Promega) reagent, and luminescence was recorded using a PHERAstar plate reader (BMG Labtech) after 10-minute incubation at room temperature. We re-screened 10 compounds that had poor quality in the initial screen and the SNK-6 cell line with all the 172 compounds. The percent inhibition values from each viability measurement are available in the ORCESTRASet object. The normalized dose-response data were used to calculate drug sensitivity scores (DSS) for each compound and cell line, as described previously (25, 26).

### Drug combination testing

Drug combinations were profiled using a fixed-ratio diagonal design, in which combination responses at pairwise doses of each drug were experimentally tested using seven doses for each compound, using the same CellTiter-Glo cell viability assay as that for the single-agent testing (Supplementary Table S3). The DECREASE (27) ML model was applied to complete the full  $8 \times 8$  dose-response matrices using predicted combination responses. The expected combination responses were calculated based on the zero interaction potency reference model implemented in the Synergy-Finder (28) web tool. Deviations between the observed and expected responses were quantified using the delta score, for which positive values indicate synergistic interactions and negative values denote antagonism. Most synergistic area scores, identified as the most synergistic  $3 \times 3$  dose window within each  $8 \times 8$  dose-response matrix, were used to generate the synergy profiles presented in Fig. 3. All the measured dose-response combination matrices are available in <https://aittokallio.group/tcl38/TCL38combi>.

### RNA sequencing analysis

RNA sequencing (RNA-seq) of the TCL-38 cell lines was done in two batches: the first 16 cell lines were sequenced first and the genome-wide transcriptomic profiles used as input data for the ML models for single-agent responses. The remaining 22 cell lines were sequenced later using the same technology. All cell lines were cultured and harvested in biological triplicates. For consistency, all cells were subcultured into fresh media 24 hours prior to harvesting, and cytokine-dependent lines were additionally supplemented with fresh cytokine 4 hours prior to harvesting. Total RNA was extracted using the RNeasy Plus Micro Kit (Qiagen) according to the manufacturer's instructions. RNA quality was assessed using

the 4150 TapeStation (Agilent), and only samples with an RNA integrity number  $\geq 8$  were included. For mRNA sequencing, polyA-enriched RNA libraries were prepared using the TruSeq Stranded mRNA Library Prep Kit (Illumina) and sequenced on the NovaSeq 6000 platform (Illumina) according to standard protocols.

RNA-seq reads were aligned to the hg38 reference genome using Bowtie2 (version 2.3.5.1) with default parameters. Gene-level counts were quantified using the Rsubread package (version 2.14.2) with the featureCounts function. Technical replicates (three per cell line) were averaged to obtain mean gene expression values. Genes with low expression were filtered out by retaining only those with counts greater than 10 in at least one sample and a total count greater than 25 across all samples. Gene counts were normalized using the DESeq2 package (version 1.32.0) in R. The DESeq function was applied to estimate size factors. Normalized counts were obtained using the counts function with normalized = TRUE and log-transformed log (counts + 1) to generate normalized expression values for downstream analyses.

### Variant calling

Transcript variants and gene fusions in the 38 cell lines were identified as previously described (29). Briefly, RNA-seq data were preprocessed by removing adapter and low-quality sequences and reads with insufficient length. Preprocessed reads were mapped to the human reference genome (hg38). Transcript variants, including TP53 mutations and other single-nucleotide variants, were identified using the GATK Best Practice workflow for RNA-seq variant calling. Fusion genes were identified using FusionCatcher tool from the RNA-seq data (bioRxiv 2014.011650). The complete mutation matrix showing the specific mutation status of each cell line for all the detected variants is available in the TCL38 PSet data object and in interactive TCL-38 web application. To complement the RNA-seq variant calling data, we further integrated publicly available DNA-based sequencing data (damaging mutations and hotspot mutations) from the DepMap (18) database for 28 of the 38 cell lines for which such data were available.

### Assay for transposase-accessible chromatin using sequencing

The genome-wide chromatin accessibility profile was obtained using the previously published protocol (called FAST-ATAC; ref. 30) for live sorted cells with a few adaptations described hereafter. Ten thousand cells per cell line were sorted with a FACSARIA sorter (Becton Dickinson), washed with PBS 1X, and immediately transposed in 50  $\mu\text{L}$  of reaction mixture: 25  $\mu\text{L}$  of  $2 \times$  Tagmentation DNA buffer, 2.5  $\mu\text{L}$  of Tn5 enzyme (Nextera DNA Library Preparation Kit, Illumina), 0.5  $\mu\text{L}$  of 1% digitonin (#G9441, Promega), and 22  $\mu\text{L}$  of nuclease-free water, as previously described (31). The reaction was incubated at 37°C for 30 minutes with agitation at 1,000 rpm, and transposed DNA was purified using a MinElute Reaction Cleanup kit (Qiagen). Assay for transposase-accessible chromatin using sequencing (ATAC-seq) libraries were constructed by ligating Nextera sequencing primers using PCR amplification according to the previously published protocol (30). Briefly, a first round of PCR was assessed in a volume of 50  $\mu\text{L}$  containing NEB Next High Fidelity  $2 \times$  PCR Master Mix (New England Biolabs) with primers (1.25  $\mu\text{mol/L}$  each) with the following thermal cycles: 72°C for 5 minutes and 98°C for 30 seconds, followed by five cycles (98°C for 10 seconds, 63°C for 30 seconds, and 72°C for 1 minute). To avoid over amplification of libraries, 5  $\mu\text{L}$  of the PCR reaction was subjected to qPCR in a volume of 15  $\mu\text{L}$  using SYBR Green I dye

(final  $0.6 \times$  SYBR GreenI, Life Technologies) with the respective primers ( $0.41 \mu\text{mol/L}$  each), used in the first round of PCR. Following qPCR [ $98^\circ\text{C}$  for 30 seconds, followed by 30 cycles ( $98^\circ\text{C}$  for 10 seconds,  $63^\circ\text{C}$  for 30 seconds, and  $72^\circ\text{C}$  for 1 minute)], amplification curves were analyzed and the optimal number of PCR cycles  $N$  for each sample was estimated with cycle thresholds reaching one fourth of the maximum. The remaining  $45 \mu\text{L}$  of PCR reaction was then subjected to a second round of PCR with the following thermal cycles:  $98^\circ\text{C}$  for 30 seconds, followed by  $N$  cycles ( $98^\circ\text{C}$  for 10 seconds,  $63^\circ\text{C}$  for 30 seconds, and  $72^\circ\text{C}$  for 1 minute). The libraries were purified using the Qiagen MinElute PCR Purification Kit and eluted in  $20 \mu\text{L}$  EB buffer ( $10 \text{ mmol/L}$  Tris buffer, pH 8). Libraries were quantified by qPCR using Library Quantification Primer plus KAPA SYBR FAST (Roche Diagnostics) and pooled at  $1 \text{ nmol/L}$ . Samples were sequenced using single-index paired-end ( $75 + 75 \text{ bp}$ ) reads on an Illumina NextSeq 500 instrument in high-output mode.

Sequencing adapters were trimmed with cutadapt version 3.6.50, and reads were aligned with BWA-MEM version 2.2.1 to a custom genome combining GRCh38 (GCA\_000001405.15) and human T-cell lymphotropic virus, type 1 (J02029.1). After alignment, Samtools was utilized to convert BAM files, sort the reads, and generate alignment statistics. To identify open chromatin regions from the aligned ATAC-seq data, MACS2 (v2.2.9.1) was applied for peak calling with specific parameters (`-nomodel --shift -100 --extsize 200 --q value 0.05`). To further analyze the peaks identified by MACS2, Bedtools (v2.27.1) was used to calculate read coverage across genomic regions defined by these peaks using the bedtools coverage command. The EdgeR R package (v3.42.4) was then used for trimmed mean of M-values (TMM) normalization of peak counts. For gene annotation, the ChIPseeker R package (v 1.36.0) was utilized to map accessible peaks to the nearest transcription start sites, considering a range of 3 kb upstream and downstream. This allowed for a comprehensive identification of genes associated with the ATAC-seq peaks. Gene set variation analysis (GSVA; v1.48.3) scores were calculated, and Spearman correlations were performed to relate DSS to specific gene sets.

### Array comparative genomic hybridization

Genomic DNA was extracted from the cell lines using the QIAamp DNA Blood Mini kit (Qiagen). Reference DNA for the analysis was derived from multiple healthy and anonymous male donors (Promega). Labeling and hybridization steps were performed using the SureTag DNA Labeling kit (Agilent Technologies) according to the manufacturer's instructions. Equal amounts of DNA labeled with fluorescent markers were combined and co-hybridized onto a 44 K DNA microarray (G4426B-014950, Agilent Technologies). Following hybridization and washing, scanning was conducted using a G2505B Micro Array Scanner (Agilent Technologies) in adherence to the manufacturer's guidelines. Feature Extraction (version 10.7.3.1) and Agilent Genomic Workbench version 7 were employed for data analysis. The analysis utilized the ADM-1 algorithm with a threshold of 6, and aberration boundaries were defined as  $\pm 0.25$  ( $\log_2$  ratio) with a minimum of three probes per region, as recommended by Agilent Technologies. Copy-number alteration data from the array comparative genomic hybridization (aCGH) analysis were processed to generate binary matrices, in which amplifications and deletions were encoded as 1, if present, and 0, if absent, for each cytoband in each cell line. Statistical association between the binary cytoband amplification profiles and drug response data was calculated across cell lines and was tested with Spearman rank correlation coefficient.

### Multimomics factor analysis

MOFA+ (v1.12.1; ref. 32) was used to integrate the data from DSS, gene expression (RNA-seq), chromatin accessibility (ATAC-seq), mutation profiles, and aCGH. The multimomics factor analysis was performed using the default parameters of multimomics factor analysis (MOFA).

### Drug sensitivity scoring

DSS were calculated from the experimentally measured dose-response curves using an established computational approach that integrates multiple dose-response parameters into a single response metric (25). The DSS calculation involves subsequent normalization steps, in which first the area under the dose-response curve (AUC) over the selected dose range ( $x_1, x_2$ ), for which the responses exceed a minimum activity level ( $A_{\min}$ ), was calculated and normalized using the formula  $DSS = [AUC - A_{\min}(x_2 - x_1)] / [(100 - A_{\min})(C_{\max} - C_{\min})]$ , in which the area below the minimum activity threshold ( $A_{\min} = 10$ ) is subtracted and the result is divided by the maximal response area to account for differences in the tested concentration ranges between compounds ( $C_{\min}, C_{\max}$ ). The DSS was then further normalized by the logarithm of top asymptote ( $R_{\max}$ ),  $DSS/\log_{10}(R_{\max})$ , to penalize for the compounds that show activity only at higher concentrations (potential off-target toxic responses), as described earlier (26).

### Drug-cell line specificity analysis

We analyzed the degree of specificity of each drug to the cell lines using statistical scores. Cell line classes with three or fewer cell lines, specifically  $\gamma\delta$  ( $n = 3$ ), T-LGLL ( $n = 1$ ), and PTCL-NOS ( $n = 1$ ), were excluded from the selectivity analyses to ensure robust evaluation. For the remaining cell line classes, two scores—tau and sensitivity enrichment (SE)—were calculated to quantify the drug selectivity index. This drug specificity analysis was modified from the existing scores that analyze gene expression tissue specificity (33); we adapted the tau score (34) for the drug-cell line specificity analysis. The SE score was inspired by the expression enrichment score, which also assesses gene expression tissue specificity (35).

Formally, the tau score for a drug  $i$  is calculated as

$$\tau_i = \frac{1}{G-1} \sum_{g=1}^G \left( 1 - \frac{m_{i,g}}{M_i} \right)$$

in which  $G$  is the number of cell line classes considered (in this analysis,  $G = 4$ ),  $m_{i,g}$  is the mean response (here, average DSS) of drug  $i$  in the cell line class  $g$ , and  $M_i$  is the maximum response across all cell line classes for the drug  $i$ .

The SE score for a drug  $i$  in cell line class  $g$  is calculated as

$$SE_{i,g} = \left( \frac{m_{i,g}}{\sum_{g=1}^G m_{i,g}} \right) \times \left( \frac{\sum_{g=1}^G S_g}{S_g} \right)$$

in which  $m_{i,g}$  is as defined above and  $S_g$  is the total sum of the responses (DSS over all the drugs) in the cell line class.

The tau score reflects the general specificity of a drug across various cell lines, independent of the cell line classes. A higher tau score indicates that the drug has a broad specificity across multiple cell line groups. In contrast, the SE score, tailored to individual cell line classes, quantifies the specificity of a drug response to a particular cell line class. A high SE score signifies that the drug exhibits specificity to that particular cell line group. For example, if a drug has a high tau value, alongside a high  $SE_{i,a}$  and a low  $SE_{i,b}$ , this

indicates that the drug is specific (i.e., has high and unique sensitivity) to the cell line class *a* but not to the class *b*.

Finally, drug selectivity index was for drug *i* was calculated as  $\tau_i \times SE_{i,g}$ , which aims to pinpoint drugs that have both general and targeted specificity, aiding in the identification of treatments that are especially effective for certain cell line classes or contexts.

### Mutation enrichment analysis

To identify enriched mutation drivers in the DSS clusters (three clusters) and TCL cell line classes (seven classes; Supplementary Fig. S1), a Fisher exact test was conducted using the stats package (v.4.3.3), with multiple testing corrections using the Benjamini–Hochberg procedure (36). Significant genes were filtered according to the following criteria: a raw *P* value of less than 0.05, classification as a Tier 1 Hallmark Catalogue of Somatic Mutations in Cancer census gene, and inclusion in the Catalogue of Somatic Mutations in Cancer database for tissue type of leukemia/lymphoma (Tier = 1, Hallmark = 1, and tissue.type contains L).

### GSVA

GSVA scores were calculated using RNA-seq data against the Molecular Signatures Database (37) pathways. The analysis was performed using the GSVA R package (version 1.48.3; ref. 38). All *Homo sapiens* pathways from the Molecular Signatures Database (version 2023.2) were included in the analysis.

### STITCH network analysis

Protein–chemical and protein–protein interactions were predicted using STITCH (Search Tool for Interactions of Chemicals; ref. 39) with default parameters. The analysis utilized the built-in scoring scheme of STITCH for interaction confidence without any parameter modifications.

## Results

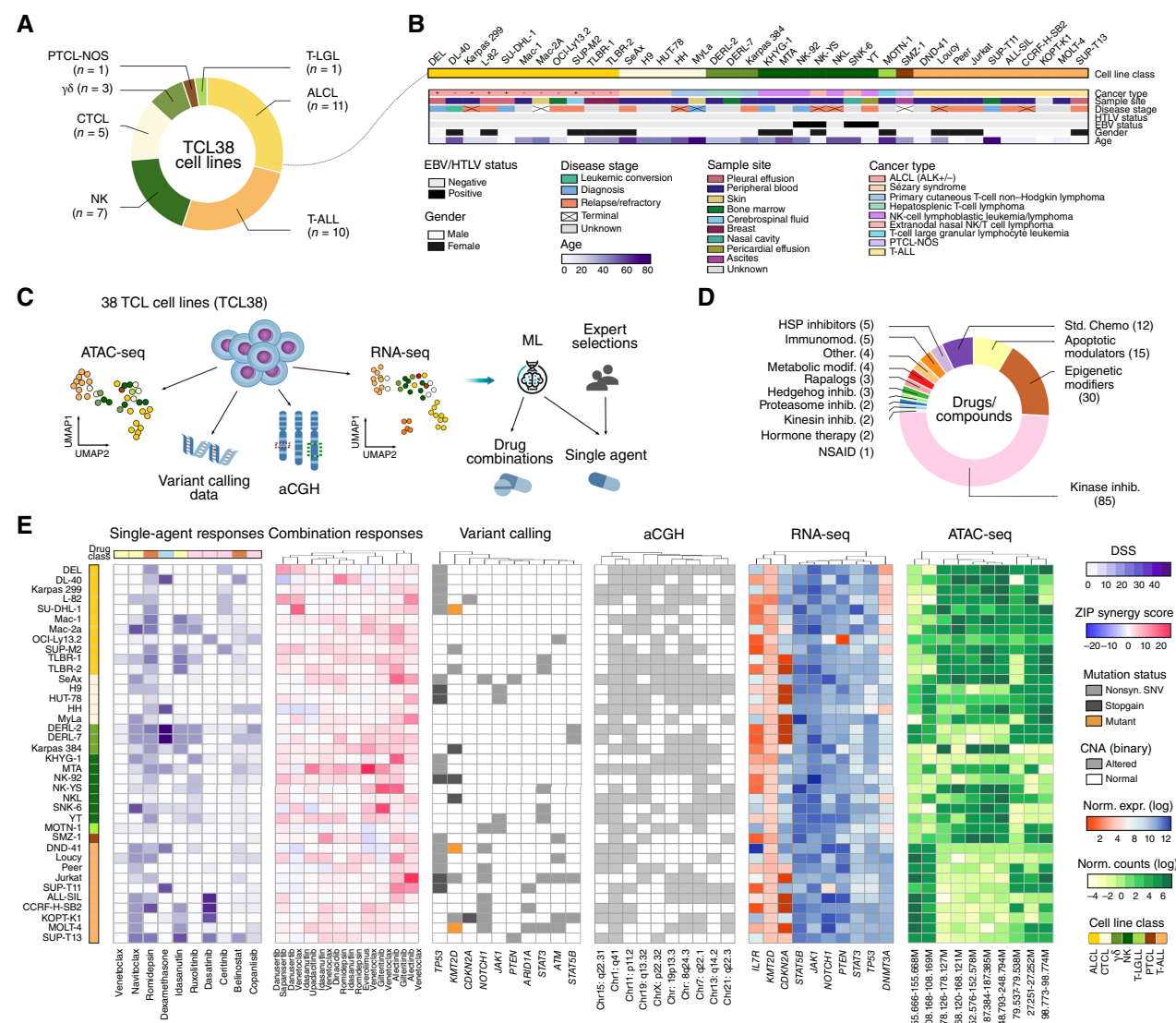
### Establishment of the integrated TCL-38 multiomics resource

To facilitate the discovery of novel therapeutic strategies for TCLs, we established the TCL-38 resource—a comprehensive panel of publicly available TCL cell lines representing virtually all major TCL subtypes (Fig. 1A and B). The selection of the 38 cell lines from the seven TCL classes was primarily driven by their global availability, leveraging collaborations and extensive sourcing efforts. The cell line panel was aimed to portray the genetic, epigenetic, and phenotypic diversity inherent in both mature and immature TCLs, covering both inter- and intra-subtype heterogeneity, thereby providing a robust platform for translational research. Multiple cell lines from each major subtype, including T-ALL, were included to provide sufficient statistical power for ML-based biomarker discovery and to capture the significant heterogeneity that exists within established TCL subtypes. As established cell lines for certain TCL entities are rare (e.g., T-LGLL and PTCL-NOS) and do not exist for others (e.g., T-PLL), we could only include single representative lines (*n* = 1) for these TCL subtypes to ensure comprehensive coverage and to provide a reference for the other cell contexts. We performed extensive multiomics profiling on the TCL-38 cell lines to establish a comprehensive and harmonized dataset that integrates genomic, transcriptomic, epigenomic, and pharmacologic data (Fig. 1C). Specifically, we conducted bulk RNA-seq for gene expression profiling, variant calling, and gene fusion detection, ATAC-seq for characterizing epigenomic landscapes and chromatin accessibility, and aCGH for assessing copy-number alterations.

Functional studies were based on profiling of drug-induced cell viability effects as phenotypic responses to both monotherapies (Supplementary Fig. S1) and combination perturbations (<https://aittokallio.group/tcl38/TCL38combi>). Additionally, we integrated publicly available DNA sequencing data from DepMap (18) for 28 cell lines, including damaging mutations and cancer hotspot mutations (Supplementary Table S2), providing further genomic characterization that complements our experimental profiling.

Our drug screening library comprised 172 unique agents chosen based on their established or potential efficacy as TCL treatments, identified by domain experts, along with compounds predicted to be effective in specific TCL subtypes using ML modeling (Fig. 1D; Supplementary Table S3). Additionally, using gene expression profiles, we predicted synergy profiles for 17 targeted drug combinations in the three largest TCL disease entities (NK, ALCL, and T-ALL). These drug combinations were experimentally tested in all 38 cell lines using the fixed-ratio “diagonal” multidose combination matrix design, in which the experimentally non-measured drug-dose combinations were predicted using the DECREASE ML platform (27). We processed and normalized these genome-wide omics and drug sensitivity data using established protocols (see Materials and Methods; Fig. 1E). The TCL-38 multiomics data resource has been made available as an integrated and fully documented TCL38 PSet data object in the cloud-based ORCESTR platform for transparent and reproducible research (40). Upon future updating and/or extension of these data, the ORCESTR data object enables version handling and sharing of the future versions of the pharmacogenomic resource with the research community. The users can also explore and analyze the multiomics data via an interactive web app (<https://aittokallio.group/tcl38>).

The single-agent responses show relatively high heterogeneity both between and within the established TCL subtypes (Supplementary Fig. S2), indicating that factors beyond the TCL class determine drug responses. A few common gene mutations, such as those encoding for the developmental *NOTCH1* receptor and the transcriptional repressor *SPEN*, were specifically enriched in the TCL cell line classes, and many more, mostly different mutations were uniquely enriched in the DSS-based, phenotypic cell line clusters (Supplementary Table S4). Specifically, *NOTCH1* mutations were identified in nine of 38 cell lines (23.7%) and *SPEN* mutations in eight of 38 cell lines (21.1%), with complete mutation status available for all cell lines in the TCL-38 resource. To uncover shared sources of variation across the different omics layers that contribute to the observed differences in drug sensitivity phenotypes among TCLs, we applied MOFA (32) to the processed multiomics data matrices and drug sensitivity profiles. MOFA identified a total of 15 factors jointly explaining the observed variation in the multiomics TCL-38 data (Supplementary Fig. S3). Factors 1, 4, and 9 collectively explained 12.86% of the variation in DSS, 19.19% in RNA-seq data, and 12.35% in ATAC-seq data. The specific genes and chromatin-accessible sites with the highest factor loadings for these three MOFA factors are provided in Supplementary File S1, revealing the molecular features most strongly associated with the drug sensitivity variation across TCL subtypes. In contrast, genomic features, such as somatic mutations (identified through variant calling) and copy-number aberrations (analyzed via aCGH), exhibited only a minimal joint variation with drug response profiles, accounting for 3.64% and 0.62% of the response variation, respectively. This suggests that transcriptomic and epigenetic gene regulation changes have a markedly greater influence on the drug sensitivity phenotypes in TCLs compared with somatic mutation and copy-number aberrations.



**Figure 1.**

Integrated TCL-38 resource of T-cell lymphoma cell lines and multiomics profiles. **A**, The TCL-38 resource comprises 38 cell lines representing all major TCL subtypes. **B**, Characterization of the TCL cell line panel using information from the Leukemia-Lymphoma Cell Line FactsBook (69) and Cell Model Passports (70). **C**, The omics technologies and assays performed in the multimodal profiling of the 38 cell lines. Note: Both gene expression profiling and variant calling are based on the RNA-seq data. **D**, The drug library with the number of compounds across different compound and target classes indicated. The compounds in the library were selected based on their established efficacy as potential TCL treatments, identified by domain experts, and based on predicted efficacy in specific TCL subtypes, predicted by ML. **E**, Examples of pharmacologic and multiomics profiles included in the TCL-38 resource. Single-agent responses, including 67 approved drugs and 105 investigational compounds, were quantified using DSS (25), and the drug combination synergy was assessed with ZIP synergy score (71). Variant calling heatmap shows the commonly mutated genes, and RNA-seq heatmaps show the expression of those genes that have chromosomal alterations, such as fusions and deletions, and expression of selected drug target genes (full drug response heatmaps are shown in Supplementary Fig. S1). CNA, copy-number alterations; EBV, Epstein-Barr virus; HTLV, human T-cell lymphotropic virus; NK, extranodal NK-/T-cell lymphoma; SNV, single-nucleotide variant; UMAP, Uniform Manifold Approximation and Projection; ZIP, zero interaction potency;  $\gamma\delta$ , cutaneous gamma-delta T-cell lymphoma. Mac-1 and Mac-2a were originally derived from a patient with CTCL, yet they cluster molecularly and functionally with ALCL lines.

Molecular profiling revealed an interesting cell line classification with regard to Mac-1 and Mac-2a cell lines. Although originally derived from a patient with CTCL, presenting with erythroderma and circulating Sézary cells, unsupervised hierarchical clustering of genome-wide transcriptomic profiles demonstrated that these lines cluster with ALCL cell lines rather than other CTCL lines (Fig. 2A) and express ALCL-associated markers including CD30 (*TNFRSF8*)

and *IRF4*. Importantly, Mac-1 and Mac-2a further cluster in terms of drug sensitivity profiles with TLBR-1 and TLBR-2 (Supplementary Fig. S1), which are both established anaplastic lymphoma kinase (ALK)-negative ALCL disease models derived from breast implant-associated lymphomas (41). These convergent molecular findings suggest that Mac-1 and Mac-2a have acquired an ALCL-like molecular phenotype and explain their current use as ALCL

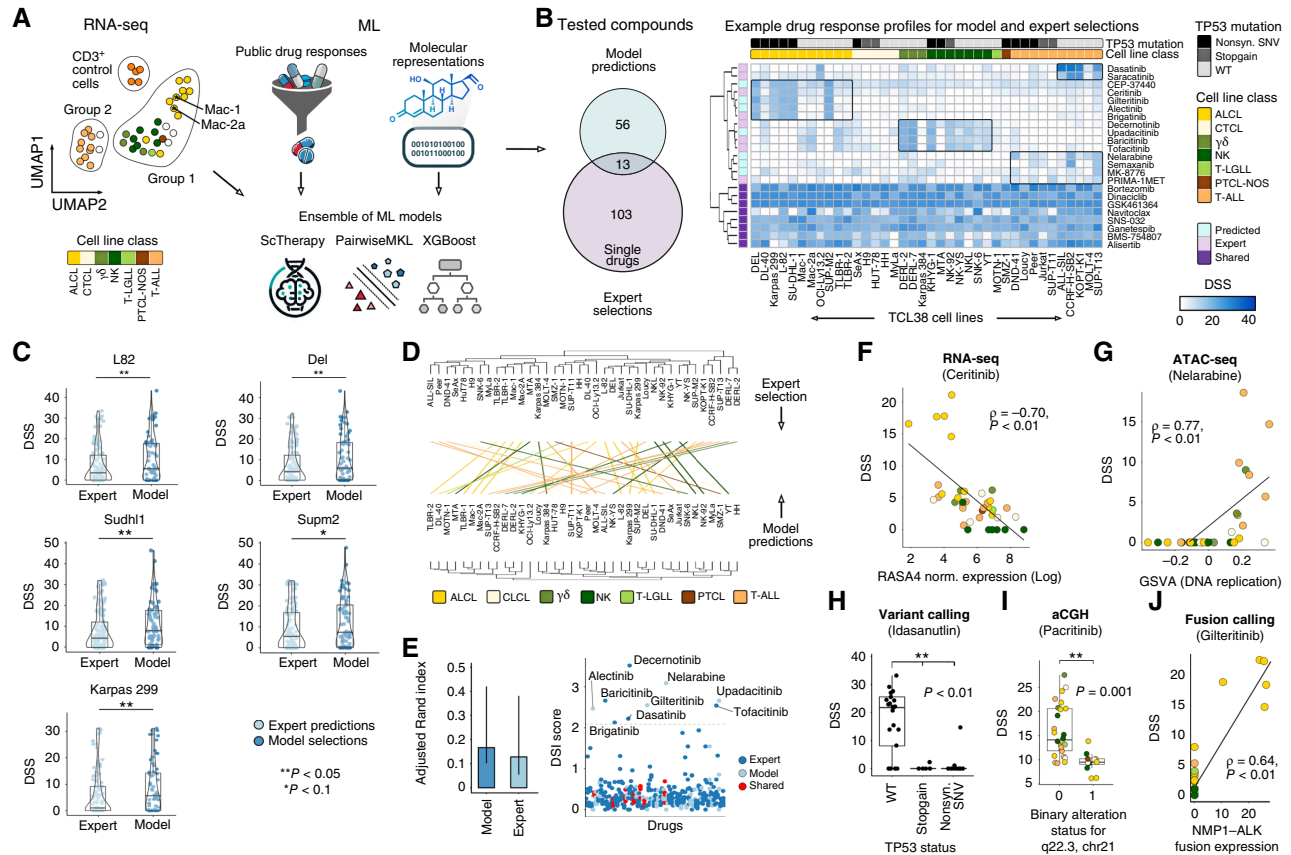


models in functional studies despite their historical CTCL derivation (42, 43).

### Selective single-agent responses predicted by ML models

To investigate potential therapeutic agents for TCLs, we conducted single-agent sensitivity profiling using two distinct compound collections. The first collection comprised 116 compounds selected by domain experts based on their established or potential clinical efficacy in TCL treatment. The second collection included

69 targeted compounds identified through a more unbiased approach, which made use of ML modeling (see Materials and Methods). Specifically, we trained three independent ML models—XGBoost (44), pairwiseMKL (45), and scTherapy (46)—to predict drug sensitivities specific to the different TCL cell line classes (Fig. 2A). These ML models utilized RNA-seq data from 15 cell lines as input features (available at the time of the model development) and single-agent response profiles of 19 single-agents as the outcome variables (Supplementary Table S1), leveraging external



**Figure 2.**

TCL-38 single-agent sensitivity prediction and screening, with links to omics markers. **A**, The models utilized RNA-seq data from 16 TCL-38 cell lines and from CD3<sup>+</sup> cells of five healthy donors available at the time of model training. XGBoost and pairwiseMKL models used drug response data available for 38 drugs across the TCL-38 cell lines, along with their molecular representations, whereas scTherapy used only the RNA-seq profiles as the input data to predict responses to 625 compounds available for testing. **B**, The ensemble ML model predictions confidently identified responses to 69 single agents when using conformal prediction (Supplementary Figs. S4 and S5). These model-selected drugs were then compared and combined with 116 compounds preselected by experts, with 13 compounds overlapping between the two collections (left). In total, 172 compounds were experimentally tested across the 38 cell lines, with two additional entries accounting for duplicates of everolimus and tacrolimus; the right heatmap shows example response profiles, and the full heatmap is provided as Supplementary Fig. S1. **C**, Cell lines with significant differences in DSS between the model-predicted and expert-selected drugs (two-sided Wilcoxon test). **D**, Hierarchical clustering of the TCL-38 cell lines based on DSS profiles from model-predicted and expert-selected compound libraries using the Euclidean distance and Ward clustering method. **E**, Left, comparison of the two clustering solutions, excluding singletons, using adjusted Rand index, with established TCL cell line classes as ground truth solution. Adjusted Rand index ranges between -1 and +1, for which one corresponds to two completely similar clustering solutions. Error bars represent 95% confidence intervals. Right, the drug specificity index quantifies the specificity of a drug to a particular TCL cell line class. The dotted lines correspond to the 95<sup>th</sup> drug specificity index percentile. **F–I**, Links to multionics data identify molecular determinants of drug sensitivity phenotypes. Each point corresponds to a cell line and the colors indicate the TCL cell line class. **F**, RNA-seq data reveals a negative correlation between ceritinib sensitivity and *RASA4* expression levels. **G**, ATAC-seq data reveal a positive correlation between nelarabine sensitivity and DNA replication GSVA score. **H**, Variant calling data explain higher responses of idasanutlin in *TP53* wild-type (WT) cell lines. **I**, aCGH data classify the dual JAK2/FLT3 inhibitor pacritinib sensitivity in terms of the binary alteration status of q22.3, chr21. **J**, Correlation between *NPM1-ALK* fusion expression and increased sensitivity to the FLT3/ALK inhibitor gilteritinib in ALCL cell lines. DSI, drug specificity index; NK, extranodal NK/T-cell lymphoma; SNV, single-nucleotide variant; γδ, cutaneous gamma-delta T-cell lymphoma.

data published before the creation of the TCL38 object for model training. To enhance the ML models' ability to distinguish between cancerous and noncancerous cells, and to focus on cancer cell-selective drug sensitivity predictions, we incorporated RNA-seq data from CD3<sup>+</sup> cells of five healthy donors as reference profiles. Molecular fingerprints of the compounds were used as structural predictors of the drug responses. However, as the scTherapy model does not require drug testing data as its input, its predictions were based merely on the RNA-seq data of the 16 TCL cell lines and five healthy donors (46).

We generated the final set of drug response predictions by integrating the outputs of the three ML models into an ensemble prediction model (Supplementary Fig. S4). We used conformal prediction to categorize ML-predicted responses into low- and high-confidence predictions; for instance, the pan-BCL2/BCL-XL/BCL-w inhibitor navitoclax showed both low-confidence and overall low-sensitivity predictions (Supplementary Fig. S5). Notably, the ensemble ML predictions identified unexpected responses to 56 drugs not suggested by the experts, accounting for 81% of the model-predicted compounds (Fig. 2B, left). Only 13 compounds (8% of all tested agents) overlapped between the expert-selected and model-predicted compound collections; two shared drugs served as technical between-plate replicates in the drug screening assays to assess their reproducibility (Supplementary Fig. S6). Taken together, across both expert- and ML-selected drugs, we identified several cell line-specific responses that clustered with the TCL subtypes (Fig. 2B, middle, black boxes). For instance, Janus kinase (JAK) inhibitors upadacitinib (JAK1/2/3 selective), baricitinib (JAK1/2 selective), and tofacitinib (JAK1/3 selective) showed context-specific response profiles in  $\gamma\delta$  and NK-cell line classes. In contrast, broadly active drugs, such as the proteasome inhibitor bortezomib, showed consistently high activity across multiple cell classes. We note that cell context-specific sensitivities are rarer and therefore more difficult to predict, compared with nonselective drugs that show high activity across various cell lines.

Given that dasatinib, an SRC/ABL tyrosine kinase inhibitor, showed selective efficacy only in a subset of T-ALL cell lines (Fig. 2B), we sought to understand the molecular basis underlying this heterogeneous response pattern. Previous studies in patients with T-ALL have demonstrated that ~30% of patients with T-ALL exhibit unexpected sensitivity to multikinase inhibitor dasatinib, with response correlating with higher levels of phosphorylated SRC, hence establishing a SRC pathway dependency model for dasatinib sensitivity in patients with T-ALL (47). To validate this patient-derived finding using our cell line resource, we analyzed the correlations between kinase expression levels and dasatinib response across T-ALL cell lines. *ABL1* expression emerged as the strongest predictor ( $\rho = 0.68$ ,  $P = 0.035$ ), followed by SRC family members *YES1* ( $\rho = 0.66$ ,  $P = 0.044$ ) and *FYN* ( $\rho = 0.58$ ,  $P = 0.088$ ; Supplementary Fig. S7). These results validate and extend the patient-derived SRC pathway dependency by identifying expression-based biomarkers that can serve as response predictors in T-ALL cases.

To further demonstrate the ability of the TCL-38 resource to recapitulate clinically relevant TCL biology, we examined the well-established, dichotomous drug response paradigm in ALCL. Stratifying the ALCL cell lines based on *ALK* expression levels, we observed a sharp and statistically significant divergence in the sensitivity to two distinct kinase inhibitor classes. In line with extensive clinical data demonstrating the efficacy of ALK inhibitors in ALK-positive disease (48, 49), the cell lines with high *ALK*

expression showed strong sensitivity to agents like alectinib and ceritinib ( $P = 0.0045$  and  $P = 0.008$ , respectively, Supplementary Fig. S8). Conversely, the cell lines with low or absent *ALK* expression (ALK negative) were significantly more sensitive to JAK inhibitors, such as upadacitinib ( $P = 0.027$ ), showing the highest responses in cell models exhibiting a gene expression signature indicative of high JAK-STAT pathway activation (Supplementary Fig. S9).

When comparing the expert-selected drugs and those predicted by the ensemble ML model, we observed that many of the drugs that originated from either of the two collections show relatively selective response patterns across the 38 cell lines, whereas the drugs that were in common between the two collections show more broadly sensitive responses (Fig. 2B; Supplementary Fig. S1). The overall drug sensitivity levels across the cell lines were relatively similar between the expert-selected and model-predicted drugs, as expected, although five of the 38 cell lines showed higher drug responses to the model-predicted drugs (Fig. 2C). Hierarchical clustering of the cell lines based on their DSS profiles revealed relatively distinct response patterns for the two drug collections (Fig. 2D), indicating that they capture complementary aspects of functional TCL phenotypes. Clustering of the cell lines using the two drug collections showed a similar level of concordance with the established TCL cell line classifications (adjusted Rand index = 0.16 vs. adjusted Rand index = 0.13; Fig. 2E, left). Analysis of the drug response specificity further confirmed that the compounds shared between the two collections show relatively nonspecific response patterns across the 38 cell lines, whereas both the expert-selected and model-predicted drugs show more selective and cell context-specific responses (i.e., unique to either a single cell line or class; Fig. 2E, right).

Taken together, these results show that the TCL-38 cell lines display heterogeneous drug response profiles, both between and within the established TCL subtypes, and that the expert-selected and model-predicted drug collections portray complementary functional information on the TCL phenotypic diversity in terms of drug-induced cell viability responses.

### Molecular determinants of context-specific drug sensitivities in TCL

To demonstrate how integrating multiomics data in the TCL-38 cell line resource provides insights into the mechanisms underlying drug response phenotypes, we present five specific examples, in which context-specific drug responses are linked to distinct molecular features identified through RNA-seq, ATAC-seq, variant calling, aCGH copy-number alterations, and fusion calling. In the first example, RNA-seq data revealed a negative correlation between the expression of *RASA4* and sensitivity to the ALK inhibitor ceritinib ( $\rho = -0.7$ ,  $P < 0.01$ ; Fig. 2F). Specifically, ceritinib demonstrated an enhanced efficacy, as indicated by higher DSS, in ALCL cell lines characterized by lower expression levels of *RASA4* ( $P = 0.004$ , two-sided Wilcoxon test; Supplementary Fig. S10).

ATAC-seq profiling indicated that increased chromatin accessibility in DNA replication genes correlated with an elevated sensitivity to nelarabine, a nucleoside analog prodrug ( $\rho = 0.77$ ,  $P < 0.01$ ; Fig. 2G). Nelarabine is converted intracellularly into its active form, ara-GTP, which is incorporated into DNA during DNA replication, leading to premature chain termination and inhibition of DNA synthesis (50, 51). Therefore, cell lines with elevated chromatin accessibility at DNA replication genes—indicative of increased replication activity—are more likely to be sensitive to the cytotoxic effects of nelarabine because of higher incorporation of ara-GTP



during DNA synthesis. Notably, this effect was predominantly observed in T-ALL cell lines, a malignancy for which nelarabine is an approved therapy (52). This finding suggests that chromatin accessibility profiles could serve as additional predictive markers for responsiveness to antimetabolite therapies in TCLs.

Variant calling revealed that sensitivity to idasanutlin, an MDM2 inhibitor, was significantly higher in cell lines with wild-type *TP53* compared with those with mutated *TP53* ( $P < 0.001$ , two-sided Wilcoxon test; **Fig. 2H**). Idasanutlin functions by disrupting the MDM2-p53 interaction, thereby activating p53-mediated apoptotic pathways. The presence of functional p53 is essential for this mechanism, as has been established in various cancer types, including T-ALL, T-PLL, acute myeloid leukemia (AML), and neuroblastoma (53–56), explaining the increased efficacy of idasanutlin in *TP53* wild-type TCL cell lines. This finding underscores the importance of *TP53* status as a potential predictive biomarker for idasanutlin responsiveness in TCLs.

Analysis of aCGH data showed that amplification classes (presence 1 and absence 0) in chromosome 21q22.3 were associated with decreased sensitivity to the *JAK2/FLT3* inhibitor pacritinib ( $P = 0.001$ , two-sided Wilcoxon test; **Fig. 2I**). The 21q22.3 region encompasses several critical genes, including *UBASH3A*, *S100B*, *ITGB2*, *ABCG1*, and *DNMT3L*, which play roles in drug transport, epigenetic regulation, signal transduction, and cellular stress responses. For example, *UBASH3A* functions as a negative regulator of T-cell receptor signaling (57), and alterations there can lead to enhanced T-cell activation as well as activation of alternative proliferative and survival pathways independent of *JAK2* and *FLT3* signaling, potentially resulting in decreased sensitivity to pacritinib. Collectively, these genetic alterations in the 21q22.3 region likely disrupt normal cellular functions and activate compensatory mechanisms that diminish the efficacy of pacritinib in treating TCLs.

Finally, our integrative analysis revealed that expression of the *NPM1-ALK* fusion gene correlates with increased sensitivity to gilteritinib, inhibitor of FMS-related receptor tyrosine kinase 3 (*FLT3*), particularly in ALCL cell lines ( $\rho = 0.64$ ,  $P < 0.01$ ; **Fig. 2J**). The *NPM1-ALK* fusion is known to result from the t(2;5) (p23;q35) translocation, which drives oncogenesis in ALCL (58). Gilteritinib is a tyrosine kinase inhibitor approved for the treatment of *FLT3*-mutated AML because of its potent inhibition of *FLT3* (59). Gilteritinib has also shown potency as an *ALK* inhibitor, and our observation suggests that the *NPM1-ALK* fusion status may serve as a predictive biomarker for sensitivity to gilteritinib in ALCL.

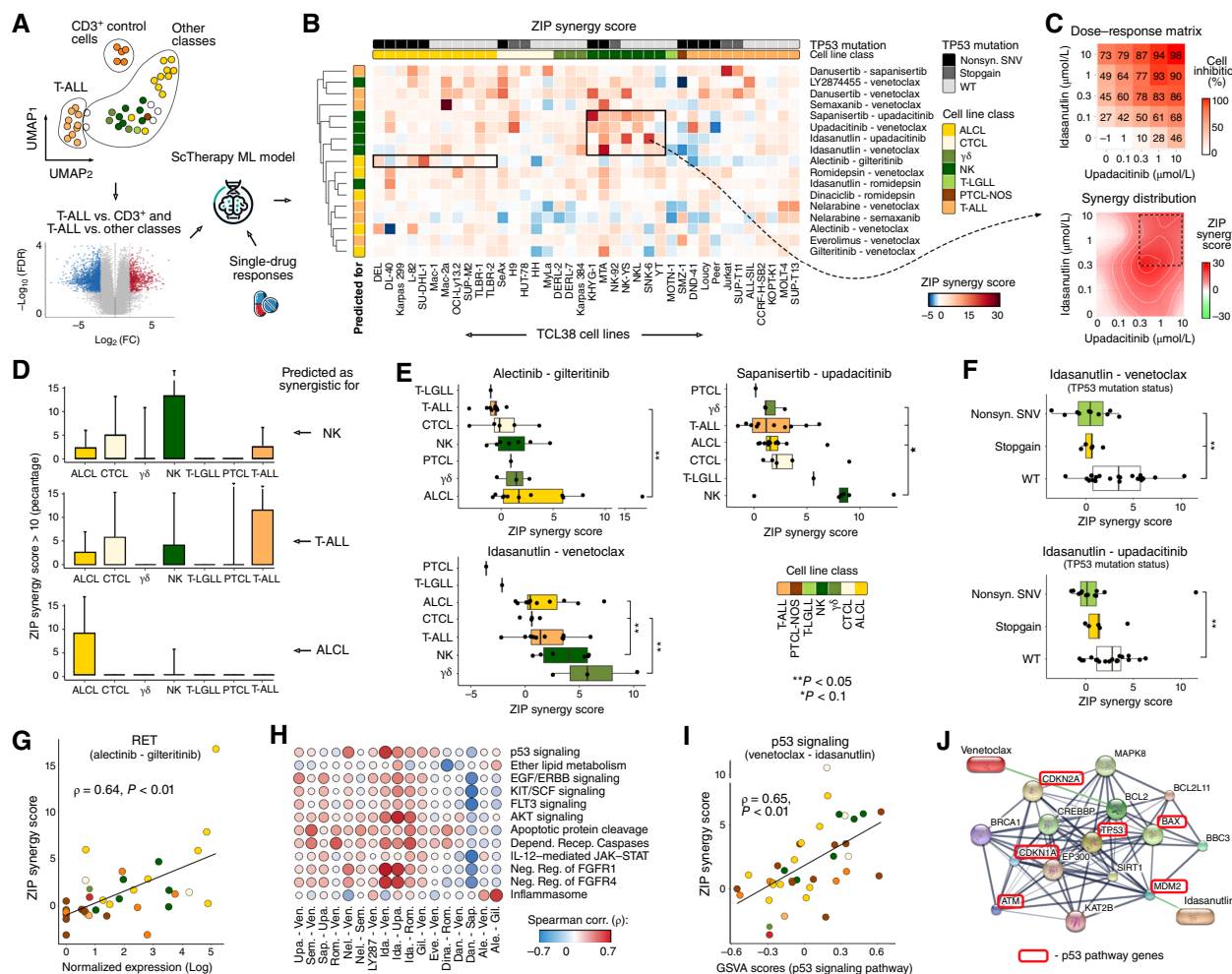
To assess clinical translatability of the drug responses, we compared  $IC_{50}$  values to the clinical pharmacokinetic data from the FDA prescribing information for key compounds (Supplementary Table S5). Clinical  $C_{max}$  values were obtained for romidepsin (377 ng/mL, 0.70  $\mu\text{mol/L}$ ) from Istodax label at 14 mg/m<sup>2</sup> dose; navitoclax (4.0  $\mu\text{mol/L}$ ) calculated from dose-proportional pharmacokinetics at the therapeutic 325 mg dose; nelarabine active metabolite ara-G (31.4  $\mu\text{g/mL}$ , 110.9  $\mu\text{mol/L}$ ) from Arranon label at 1.5 mg/m<sup>2</sup> dose; and gilteritinib (282 ng/mL, 0.51  $\mu\text{mol/L}$ ) from Xospata label at 120 mg daily dose. Notably, romidepsin demonstrated wide clinical relevance, with all 38 cell lines (100%) showing  $IC_{50}$  values less than clinical  $C_{max}$ , indicating broad therapeutic potential across the TCL subtypes. Navitoclax showed strong clinical potential with 30 of 38 lines (79%) at achievable concentrations, supporting its development for TCL treatment. In contrast, nelarabine and gilteritinib exhibited more selective clinical relevance, with eight of 38 (21%) and seven of 38 (18%) cell lines, respectively, showing clinically achievable activity.

Together, these case studies demonstrate how integrating multi-omics data with pharmacologic profiling can uncover potential molecular determinants of drug sensitivity in TCLs, thereby enhancing our understanding of TCL biology and paving the way for more precise therapeutic strategies tailored to the molecular characteristics of individual cancers.

### Prediction of cell class-specific drug combinations based on gene expression

To show how the TCL-38 resource can be used to predict novel and selective drug combinations for a specific TCL cell line class, we utilized our scTherapy model that has previously been shown to provide accurate predictions of multitargeted therapy responses using single-cell transcriptomic profiles in patients with hematologic malignancies (46). The original prediction approach identified multiple subclones of cancer cells using single-cell RNA-seq data, with each subclone targeted by a specific drug, and when these drugs were combined, they led to selective inhibition of the target clones while sparing noncancerous cells. To apply scTherapy to bulk RNA-seq profiles, we first compared the three largest TCL-38 cell line classes with the healthy CD3<sup>+</sup> control cells and then each TCL class against the other cell classes, excluding the one for which the predictions were made (**Fig. 3A**). This pairwise differential expression analysis resulted in confident cell class-specific predictions for approximately 25 drugs per cell line class. We next prioritized drugs that showed moderate single-agent responses across the TCL-38 cell lines (median  $5 < \text{DSS} < 30$ ); this was done to filter out completely inactive drugs and those showing extreme activity alone as those are not expected to lead to combination synergies (12). This DSS filtering led to a total of 17 drug combinations identified for the three largest TCL cell line classes (**Table 1**; Supplementary Table S2). Notably, although some combinations involved common drugs, such as venetoclax, each combination was unique to its specific cell line class. In general, venetoclax showed a relatively low efficacy as monotherapy in most of the TCL cell lines (Supplementary Fig. S2), hence requiring combination partners that boost its activity in specific cell contexts.

We next experimentally tested the 17 model-predicted combinations across the 38 cell lines (**Fig. 3B**), using  $8 \times 8$  dose combination matrices for each combination and cell line (**Fig. 3C**). The experimental validations confirmed relatively heterogeneous combination responses across the TCL-38 cell line classes, analogous to the single-agent responses. For instance, the group of four combinations predicted for NK cells was specifically synergistic in most of the NK cell lines (**Fig. 3B**, black rectangle). Similarly, the alectinib-gilteritinib combination was specifically synergistic in the ALCL class even though there was a degree of within-class variability in the synergy levels. However, when considering only the strongly synergistic combinations (zero interaction potency  $\geq 10$ ), we observed a higher proportion of synergistic combinations in the cell line classes for which the predictions were originally made (**Fig. 3D**). This indicates a surprisingly high class specificity in the model predictions given the rather small number of training cell lines per cell line class. Especially for the ALCL class, the experimental synergies were highly unique to that particular class, perhaps because of this class possessing the largest number of cell lines ( $n = 11$ ). The experiments also validated selected combinations that showed differential synergy in specific TCL cell line classes (**Fig. 3E**). For instance, the idasanutlin-venetoclax combination showed enhanced synergy in NK and  $\gamma\delta$  classes. This combination was recently shown to induce selective synergy in patients with T-PLL (54), and it is currently



**Figure 3.**

TCL-38 resource enables context-specific drug combination predictions. **A**, Drug combination prediction with the scTherapy model used as inputs differentially expressed genes from pairwise cell class comparisons, such as T-ALL vs noncancer CD3<sup>+</sup> cells and other TCL classes. UMAP, Uniform Manifold Approximation and Projection. **B**, Experimental validation of the synergy predictions. The zero interaction potency synergy score was calculated over the full 8 × 8 dose combination matrix showing maximal synergy. The annotations on the left indicate the three largest classes for which a drug combination was predicted. The black rectangle indicates two class-specific groups of drug combinations. **C**, An example of measured 5 × 5 dose combination response submatrix (upper heatmap), and the corresponding synergy distribution (lower plot), for the combination between idasanutlin and upadacitinib in SNK-6 NK-/T-cell lymphoma cell line. The dotted box shows the 3 × 3 dose window with maximal synergy. **D**, The proportion of combinations with strong synergistic responses (zero interaction potency ≥ 10) across the cell line classes, with error bars indicating 95% confidence interval. The annotations on the right indicate the three cell line classes for which the predictions were made. **E**, Three selected combinations that showed a differential synergy in specific cell line classes. Statistical testing with two-sided Wilcoxon test. **F**, Two idasanutlin combinations demonstrated higher synergy in cell lines with wild-type (WT) p53. Statistical testing with two-sided Wilcoxon test. **G**, The correlation between alectinib-gilteritinib combination synergy with the expression of *RET* (rearranged during transfection) across the 38 cell lines. **H**, GSEA scores calculated based on RNA-seq data for the Molecular Signatures Database pathways and correlated with the synergy scores. **I**, The correlation between p53 signaling GSEA score with the idasanutlin-venetoclax combination synergy. **J**, STRING network analysis of the p53 pathway connects the primary targets of idasanutlin and venetoclax. NK, extranodal NK-/T-cell lymphoma; SNV, single-nucleotide variant; ZIP, zero interaction potency; γδ, cutaneous gamma-delta T-cell lymphoma.

being tested in phase I clinical trials for relapsed/refractory patients with AML (NCT02670044).

### Molecular determinants of context-specific drug combination synergies in TCLs

To identify molecular determinants of synergy, we again made use of the integrated TCL-38 omics profiles. We found that idasanutlin combinations with either venetoclax or upadacitinib demonstrated

significantly higher synergy in cell lines with wild-type p53, indicating that functional p53 enhances the cooperative effects of these combinations ( $P < 0.05$ , two-sided Wilcoxon test; Fig. 3F). To understand why such cell context-specific synergies occur, we correlated the mRNA expression of the targets of drugs in combinations with the synergy scores. We considered the alectinib-gilteritinib combination as an example, in which both drugs are clinically used: alectinib is approved for the treatment of ALK-positive metastatic non-small cell lung cancer and

**Table 1.** Drug combinations identified for the largest TCL-38 cell line classes using the scTherapy ML model.

| Cell line class                            | Drug combination           | Drug 1 MoA                             | Drug 2 MoA                                            |
|--------------------------------------------|----------------------------|----------------------------------------|-------------------------------------------------------|
| Primary extranodal NK/T-cell lymphoma (NK) | Venetoclax + LY-2874455    | BCL2-selective inhibitor               | FGFR inhibitor                                        |
|                                            | Venetoclax + idasanutlin   | BCL2-selective inhibitor               | p53-MDM2 inhibitor                                    |
|                                            | Romidepsin + idasanutlin   | HDAC inhibitor                         | p53-MDM2 inhibitor                                    |
|                                            | Upadacitinib + MLN0128     | JAK1-selective inhibitor               | mTOR inhibitor                                        |
|                                            | Venetoclax + upadacitinib  | BCL2-selective inhibitor               | JAK1-selective inhibitor                              |
| ALCL                                       | Idasanutlin + upadacitinib | p53-MDM2 inhibitor                     | JAK1-selective inhibitor                              |
|                                            | Venetoclax + alectinib     | BCL2-selective inhibitor               | ALK inhibitor                                         |
|                                            | Venetoclax + gilteritinib  | BCL2-selective inhibitor               | FLT3/AXL inhibitor                                    |
|                                            | Alectinib + gilteritinib   | ALK inhibitor                          | FLT3/AXL inhibitor                                    |
|                                            | Venetoclax + romidepsin    | BCL2-selective inhibitor               | HDAC inhibitor                                        |
| T-ALL                                      | Dinaciclib + romidepsin    | CDK inhibitor                          | HDAC inhibitor                                        |
|                                            | Venetoclax + everolimus    | BCL2-selective inhibitor               | Binds FKBP12 and causes inhibition of mTORC1          |
|                                            | Venetoclax + danusertib    | BCL2-selective inhibitor               | Aurora, Ret, TrkA, and FGFR1 inhibitor                |
|                                            | Venetoclax + nelarabine    | BCL2-selective inhibitor               | Nucleoside analog and DNA and RNA synthesis inhibitor |
|                                            | Venetoclax + semaxanib     | BCL2-selective inhibitor               | VEGFR inhibitor                                       |
|                                            | Semaxanib + nelarabine     | VEGFR inhibitor                        | Nucleoside analog and DNA and RNA synthesis inhibitor |
|                                            | Danusertib + MLN0128       | Aurora, Ret, TrkA, and FGFR1 inhibitor | mTOR inhibitor                                        |

NOTE: Supplementary Table S3 lists the concentration used in the multidose combination testing.  
Abbreviation: MoA, mode of action.

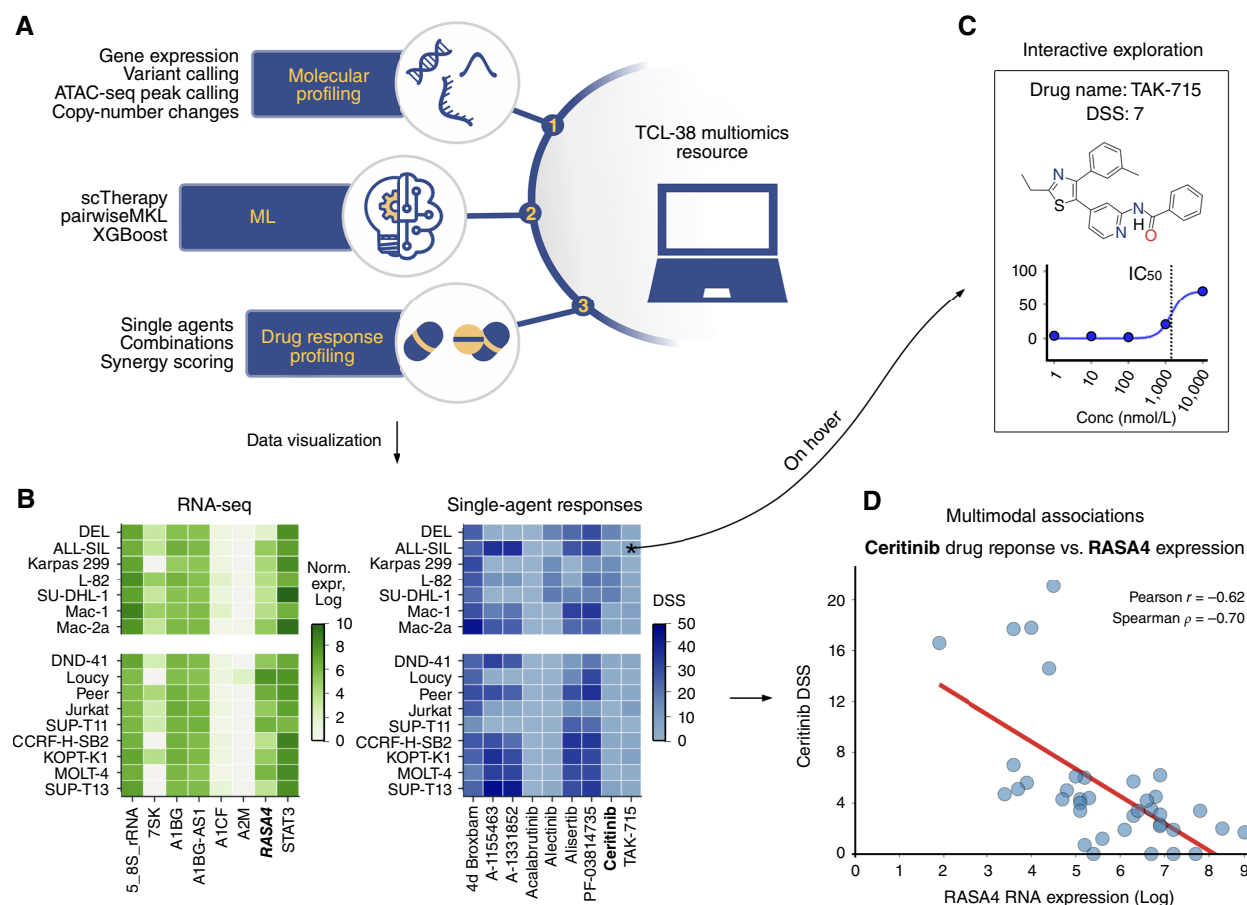
gilteritinib is used to treat adult AML with an *FLT3* mutation. In this case, the scTherapy model made an unexpected prediction of their combination synergy, specifically in the ALCL cell context, and this combination indeed exhibited ALCL-specific synergy in the dose-response experiments (Fig. 3E). Interestingly, the measured combination synergy correlated with the expression of the *RET* proto-oncogene across the 38 cell lines ( $\rho = 0.64$ ,  $P < 0.01$ ; Fig. 3G). Specifically, higher *RET* expression was associated with increased levels of measured alectinib-gilteritinib synergy, highlighting *RET* expression status as a key molecular determinant in this cellular context.

As a cell context-specific combination synergy is a rare event (estimated to be less than 5% of all possible pairwise combinations), and single genes are rarely predictive of such selective synergies (12), we next performed pathway enrichment analyses. These multi-gene analyses revealed again a relatively heterogeneous distribution of pathway-level expression determinants of TCL class-specific synergies across the 17 combinations (Fig. 3H). In the context of p53 signaling, we found that its expression enrichment is correlated with the idasanutlin-venetoclax combination effects ( $\rho = 0.65$ ,  $P < 0.01$ ; Fig. 3I). The MDM2 inhibitor idasanutlin promotes p53-dependent apoptosis through upregulation of proapoptotic proteins such as *PUMA*, *NOXA*, and *BAX*. Venetoclax complements this apoptotic effect by inhibiting BCL2, a protein that normally binds and sequesters proapoptotic proteins, such as *BAX* and *BAK*, preventing them from inducing cell death (Fig. 3J). When used together in combination, idasanutlin and venetoclax treatment removes the key inhibitory checkpoints on apoptosis, resulting in a synergistic increase in cell death, particularly in cell lines with wild-type p53 (Fig. 3F) and higher baseline levels of p53 pathway expression (Fig. 3I). This finding could not have been made without the access to a large enough panel of TCL cell lines, with comprehensive molecular profiles, showcasing the discovery value of the TCL-38 resource.

### Extending, exploring, and analyzing the multiomics data: interactive TCL-38 web app

The TCL-38 multiomics data resource has been made publicly available as an integrated and fully documented TCL38 PSet data object (60) in the cloud-based ORCESTR platform (40). The PSet object allows for combining, comparing, and conducting meta-analyses of pharmacogenomic datasets from multiple sources beyond TCL-38, including in-house data as well as publicly available datasets from PharmacDB, DrugComb, and SYNERGxDB (Supplementary Table S1). Such integrated analyses enable an extension of the therapeutic discovery to other cell contexts, cancer types, and compound libraries, beyond those profiled in the current version of TCL-38. Upon future updates and/or extensions of these data, the ORCESTR data object enables version handling and sharing of the new versions of the TCL-38 resource with the community, in line with the FAIR principles (Findability, Accessibility, Interoperability, and Reusability). For instance, the cell type selectivity of the compounds can be studied beyond the current TCL subtypes, using drug responses observed in other cell lines, such as B-cell malignancies, or in completely distinct cell contexts, such as solid tumor lines. Complementing these resources, FIMM has provided drug response profiles for more than 500 anticancer compounds—including those in the TCL-38 library—tested in healthy bone marrow samples, which can serve as reference data for cancer-selective sensitivity scoring and de-prioritizing compounds with broadly toxic effects (25, 61).

In addition, end users can explore and analyze the TCL-38 multiomics data through an interactive web application hosted on FIMM servers (<https://aittokallio.group/tcl38>), which is freely accessible without the need for a login (Fig. 4A). The web platform facilitates dynamic, side-by-side comparisons of the diverse multiomics data and drug response profiles. For example, Fig. 4B highlights such paired visualizations in the form of adjacent



**Figure 4.**

TCL-38 multiomics resource for interactive data exploration and analysis. **A**, Schematic overview of the integrated components in the web platform, encompassing molecular profiling data (gene expression, variant calling, ATAC-seq peak calling, and copy-number changes), ML approaches (scTherapy, pairwiseMKL, and XGBoost), and drug response profiling (single agents and drug combinations). All the data are freely accessible through an interactive web application. **B**, Example visualizations from the web interface, enabling side-by-side comparisons of multiple data types. The left heatmap displays RNA-seq-normalized log expression values of user-selected genes and the right heatmap shows single-agent DSSs of user-selected compounds for the same set of cell lines. **C**, Mouse-over functionality offers detailed exploration of individual drug responses as demonstrated here by the TAK-715 response curve popup. **D**, A multimodal association analysis illustrates an example correlation analysis between ceritinib drug response and *RASA4* expression levels across the 38 cell lines (the points).

heatmaps that display RNA-seq expression and drug response profiles for the user-selected set of cell lines. Interactive mouse-over functionality allows for detailed inspection of the individual drug-response curves (e.g., for TAK-715, **Fig. 4C**), whereas integrated multimodal association analyses offer means for identifying potential biomarkers for single-agent and drug combination responses, illustrated here by correlating ceritinib sensitivity with *RASA4* expression profiles (**Fig. 4D**). Such interactive analyses and open data sharing enable the community to make full use of this resource for advancing therapeutic discovery for TCLs and other malignancies.

## Discussion

TCLs form a heterogeneous group of malignancies that can arise in lymphoid tissues, such as lymph nodes, spleen, and thymus, as well as in extranodal sites, and may involve the bone marrow and peripheral blood at various stages of maturation. Accordingly, each TCL subtype is considered a distinct clinicopathologic entity,

defined by unique pathophysiology presentation and genetic/molecular differences, and therefore TCL-specific treatment options are required for improved patient outcomes. The most common TCL subtypes are categorized into immature and mature T-cell neoplasms, which were combined because of similarities and transdifferentiation cancer cell mimicry being recognized as a new hallmark of cancers (62). Toward preclinical discovery of molecularly targeted therapeutics, we established the TCL-38 resource that includes 38 patient-derived cell lines, representing the major TCL subtypes. All the cell lines were authenticated, checked to be free from infectious disease status, extensively profiled using multiomics technologies, and screened against a library of 172 agents containing 67 approved drugs (39%) and 105 investigational compounds (61%). The drug library comprises agents chosen based on their established or potential efficacy as TCL treatments, along with compounds predicted to be effective using ML models trained on earlier, scarce pharmacogenomic data. This led to the identification of both broadly effective and TCL subtype-specific responses. Our novel

profiling data additionally revealed that cell line classification based solely on the tissue of origin may not always reflect the correct molecular phenotype and cellular context of the cell lines. Mac-1 and Mac-2a cell lines exemplify this complexity, showing clear ALCL-like molecular features despite their CTCL origin.

We demonstrated in several case studies, in which novel sensitivities were linked to molecular or (epi)genetic features, that the multiomics resource has sufficient power for biomarker discovery. Our findings include both validation of established clinical mechanisms (such as the SRC pathway dependency for dasatinib and the ALK-stratified ALCL responses), as well as novel discoveries (including the correlation between ceritinib sensitivity and *RASA4* expression and the RET-mediated combination synergies). Similarly, we found that p53 signaling was correlated with the idasanutlin–venetoclax combination effects across all 38 cell lines. This demonstrates the potential of the TCL-38 resource to generate testable hypotheses beyond direct drug–target relationships. In particular, the observed negative correlation between the *RASA4* expression and ceritinib sensitivity in ALCL cell lines suggests potential mechanistic insights worthy of further investigation. Given that ALK is known to activate various downstream signaling pathways, including the RAS–MAPK pathway (63, 64), our results suggest an interplay between ALK signaling and *RASA4* function in the ALCL cells. Reduced expression of *RASA4*, which encodes a RAS GTPase-activating protein that negatively regulates RAS-mediated signaling (65), could potentially lead to enhanced RAS pathway activity. This might increase cancer cell dependence on ALK-mediated proliferation and survival, thereby enhancing sensitivity to ALK inhibition by ceritinib. When combined with ML, this study further demonstrated how the harmonized pharmacogenomic resource can be leveraged as a unique training dataset for ML models to predict novel single-drug or drug combination responses that have not yet been experimentally tested across these or other TCL subtypes.

This integrated and harmonized data resource is expected to empower the research community to uncover novel drug targets and predictive biomarkers, ultimately advancing the treatment of TCLs. In particular, the TCL-38 resource serves as a crucial complement to drug screening in primary patient material by addressing their inherent limitations, such as restricted availability and inconsistent sample and data quality. With the standardized multiomics profiles across the heterogeneous TCL cell contexts, the TCL-38 cell line panel enables robust and reproducible drug sensitivity profiling, as well as detailed TCL subtype-specific analyses. Based on our previous work on patient samples from TCL subtypes, we believe that the careful selection of the 38 cell lines to depict the molecular and phenotypic diversity of TCLs enhances the relevance of the resource to translational applications. Notably, our findings align with previous studies in mature T-cell malignancies, reinforcing vulnerabilities to inhibitors targeting MDM2 (66), CDK (67), HDAC (68), and BCL2 family members (67). Importantly, findings from the TCL-38 resource validate the key associations observed in primary samples, such as the link between *TP53* status and MDM2 inhibitor sensitivity observed in primary T-PLL samples (54), underscoring its potential to enhance biomarker discovery and accelerate targeted therapy development for TCLs.

The TCL-38 multiomics data resource has been made freely available as an integrated and fully documented TCL38 PSet data object in the cloud-based ORCESTR platform for transparent and reproducible research. Upon future updating and/or extension of these data, the ORCESTR data object enables version handling and

sharing of the future versions of the pharmacogenomic resource with the community in line with the FAIR principles. The interactive web app enables the community to explore the multiomics profiles and to identify mechanistic hypotheses for the discovery of investigational molecules and molecular biomarkers for further validation, as well as novel repurposing opportunities among already approved drugs or combinations not yet used in TCL treatment. As with any cell line resource, the translational value of the *in vitro* drug sensitivities and molecular markers will be confirmed once these are advanced to the next stages of either *in vivo* or clinical studies. Based on the case studies presented in this work and their alignment with the discoveries made in primary patient material, we anticipate that this integrated and open-data resource will ultimately advance the currently limited treatment options for patients with TCLs.

## Data Availability

The TCL38 data object has been deposited in ORCESTR to maximize FAIR dissemination at <https://www.orcestra.ca/pset/675c3be4136b907873280cb8>. The raw sequencing data have been made publicly available at the Sequence Read Archive (RRID: SCR\_004891; PRJNA1332417; <https://www.ncbi.nlm.nih.gov/sra/PRJNA1332417>). The processed TCL38 profiling data have been uploaded to Zenodo (<https://zenodo.org/records/14908028>; ref. 60). The interactive TCL38 web app is freely accessible at <https://aittokallio.group/tcl38>, without any login requirements. To complement our multiomics profiling, we integrated publicly available DNA-based sequencing data from the DepMap database (18); for 28 of the 38 TCL cell lines, we obtained both damaging mutations and hotspot mutations. These data provide additional genomic context for investigators using the TCL-38 resource. All other raw data are available upon request from the corresponding author.

## Authors' Disclosures

J. Nguyen reports grants from Canadian Cancer Society during the conduct of the study. A. Cichońska reports other support from Harmonic Discovery outside the submitted work. T. Pemovska reports grants from Krebsforschungslauf and the Initiative Krebsforschung (Comprehensive Cancer Center Vienna-Medical University of Vienna) during the conduct of the study; other support from Exalt FlexCo outside the submitted work; a patent for EP24158079.4 pending and licensed to Exalt FlexCo and a patent for EP 24169882.8 pending and licensed to Exalt FlexCo; and other support from Exalt FlexCo (co-founder and ownership of shares). E. Bachy reports grants from European ERA-NET Grant during the conduct of the study, as well as personal fees and nonfinancial support from Roche, Takeda, AbbVie, Bristol Myers Squibb, Novartis, Kite/Gilead Sciences, and Miltenyi Biomedicine and personal fees from Incyte and Innate Pharma outside the submitted work. S. Mustjoki reports grants from ERA PerMed JAK-STAT consortium, Cancer Foundation Finland, Research Council of Finland, Sigrid Juselius Foundation, Gyllenberg Foundation, Helsinki Institute of Life Science, Finnish special governmental subsidy for health sciences, research, and training, EP Permed consortium Immune-TME, and Helsinki Institute of Life Sciences during the conduct of the study, as well as grants and personal fees from Novartis, grants from Bristol Myers Squibb and Pfizer, and personal fees from Dren Bio outside the submitted work. H.A. Neubauer reports grants from Austrian Science Fund (FWF) under the frame of ERA PerMed (JAKSTAT-TARGET) and EP PerMed (ImmuneT-ME) and Austrian Science Fund SFB F06109 during the conduct of the study. B. Haibe-Kains reports personal fees from Break Through Cancer, Shriners Children Hospital, CQDM, and AACR Cancer Research outside the submitted work. No disclosures were reported by the other authors.

## Authors' Contributions

**A. Ianevski:** Conceptualization, data curation, software, formal analysis, validation, investigation, visualization, methodology, writing—original draft, writing—review and editing. **K. Nader:** Data curation, software, formal analysis, visualization, methodology, writing—original draft, writing—review and editing. **J. Nguyen:** Resources, data curation, software, methodology. **H. Sorger:** Resources, data curation. **S. Timonen:** Resources, data curation, validation. **E. Julia:** Resources, data curation, methodology. **D. Pölöske:** Resources, data curation, methodology.

**K. Spirk:** Resources, data curation, methodology, writing-review and editing. **C. Wagner:** Resources, data curation, methodology. **D. Junghertz:** Resources, data curation, methodology. **M. Nakano:** Resources, data curation, software. **S. Kadambat Nair:** Resources, data curation, software. **P. Ianevski:** Resources, software, visualization, methodology. **M. Kankainen:** Formal analysis, methodology, writing-review and editing. **D. Dias:** Visualization. **A. Cichońska:** Formal analysis, methodology. **T. Pemovska:** Investigation, writing-review and editing. **C. Pirker:** Resources, data curation, methodology, writing-review and editing. **W. Berger:** Resources, data curation, methodology. **T. Braun:** Resources, supervision, methodology, writing-original draft, writing-review and editing. **R. Moriggi:** Resources, supervision, funding acquisition, writing-review and editing. **E. Bachy:** Resources, supervision, funding acquisition, writing-review and editing. **S. Mustjoki:** Resources, supervision, funding acquisition, writing-original draft, writing-review and editing. **M. Herling:** Resources, supervision, funding acquisition, writing-original draft, writing-review and editing. **H.A. Neubauer:** Conceptualization, resources, supervision, funding acquisition, writing-original draft, writing-review and editing. **B. Haibe-Kains:** Conceptualization, resources, supervision, funding acquisition, writing-review and editing. **T. Aittokallio:** Conceptualization, supervision, funding acquisition, investigation, writing-original draft, writing-review and editing.

## Acknowledgments

The authors thank Petra Mayer and Tony Andreas Müller for sample handling. The drug assays were carried out at the FIMM High Throughput Biomedicine Unit, hosted by the University of Helsinki and supported by HiLIFE and Biocenter Finland. CSC IT Center for Science is acknowledged for data storage and

computational resources. T. Aittokallio is supported by Research Council of Finland (grants 340141, 344698, 345803, and 367855), under the frame of ERA PerMed JAKSTAT-TARGET consortium; Research Council of Norway (project 357095), under the frame EP PerMed as part of the European ImmuneT-ME consortium; the Cancer Foundation of Finland; the Norwegian Cancer Society (grants 216104 and 273810); the Sigrid Jusélius Foundation; and iCAN – Digital Precision Cancer Medicine Flagship (iCAN-MULTIDRUG). S. Mustjoki is supported by European Union's Horizon 2020 Research and Innovation Programme (ERA PerMed JAKSTAT-TARGET), Research Council of Finland under the frame EP PerMed as part of the European ImmuneT-ME consortium, Cancer Foundation Finland, Sigrid Jusélius Foundation, the Helsinki Institute of Life Science (HiLIFE) Fellow grants, Signe and Ane Gyllenberg Foundation, and the Finnish special governmental subsidy for health sciences, research, and training. H.A. Neubauer is supported by Austrian Science Fund Special Research Area grant (SFB-P06109) under the frame of ERA PerMed (JAKSTAT-TARGET) and EP PerMed (ImmuneT-ME). T. Pemovska is supported by Comprehensive Cancer Center Forschungsförderung der Initiative Krebsforschung, Medical University of Vienna.

## Note

Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org/>).

Received February 27, 2025; revised August 15, 2025; accepted October 24, 2025; posted first October 30, 2025.

## References

- Vose JM. Peripheral T-cell non-Hodgkin's lymphoma. *Hematol Oncol Clin North Am* 2008;22:997–1005.
- Yoon SE, Song Y, Kim SJ, Yoon DH, Chen TY, Koh Y, et al. Comprehensive analysis of peripheral T-cell and natural killer/T-cell lymphoma in Asian patients: a multinational, multicenter, prospective registry study in Asia. *Lancet Reg Health West Pac* 2021;10:100126.
- Cree IA. The WHO classification of haematolymphoid tumours. *Leukemia* 2022;36:1701–2.
- Lunning MA. Treatment of peripheral T-cell lymphoma: many shades of gray. *Oncol Williston Park N* 2015;29:545–50.
- Broccoli A, Argnani L, Zinzani PL. Peripheral T-cell lymphomas: focusing on novel agents in relapsed and refractory disease. *Cancer Treat Rev* 2017;60:120–9.
- Jabbour E, Pui CH, Kantarjian H. Progress and innovations in the management of adult acute lymphoblastic leukemia. *JAMA Oncol* 2018;4:1413–20.
- Hristov AC, Tejasvi T, Wilcox RA. Cutaneous T-cell lymphomas: 2021 update on diagnosis, risk-stratification, and management. *Am J Hematol* 2021;96:1313–28.
- Lansigan F, Horwitz SM, Pinter-Brown LC, Rosen ST, Pro B, Hsi ED, et al. Outcomes for relapsed and refractory peripheral T-cell lymphoma patients after front-line therapy from the COMPLETE registry. *Acta Haematol* 2020;143:40–50.
- Ngu H, Parkin S, Villa D, Scott D, Gerrie A, Toze C, et al. TCL-323 outcome of relapsed and refractory peripheral T-Cell lymphoma (PTCL) with intention for curative therapy incorporating high-dose chemotherapy and hematopoietic stem cell transplant (HDC/SCT). *Clin Lymphoma Myeloma Leuk* 2022;22:S399–400.
- Stefoni V, Pellegrini C, Argnani L, Corradini P, Dodero A, Orsucci L, et al. Brentuximab vedotin in the treatment of relapsed/refractory CD30+ peripheral T-cell lymphoma: a FIL phase 2 study. *Hematol Oncol* 2022;40:307–9.
- Trastulla L, Noorbakhsh J, Vazquez F, McFarland J, Iorio F. Computational estimation of quality and clinical relevance of cancer cell lines. *Mol Syst Biol* 2022;18:e11017.
- Jaaks P, Coker EA, Vis DJ, Edwards O, Carpenter EF, Leto SM, et al. Effective drug combinations in breast, colon and pancreatic cancer cells. *Nature* 2022;603:166–73.
- Shoemaker RH. The NCI60 human tumour cell line anticancer drug screen. *Nat Rev Cancer* 2006;6:813–23.
- Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The cancer cell line encyclopedia enables predictive modelling of anti-cancer drug sensitivity. *Nature* 2012;483:603–7.
- Basu A, Bodycombe NE, Cheah JH, Price EV, Liu K, Schaefer GI, et al. An interactive resource to identify cancer genetic and lineage dependencies targeted by small molecules. *Cell* 2013;154:1151–61.
- Garnett MJ, Edelman EJ, Heidorn SJ, Greenman CD, Dastur A, Lau KW, et al. Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature* 2012;483:570–5.
- Corsello SM, Nagari RT, Spangler RD, Rossen J, Kocak M, Bryan JG, et al. Discovering the anti-cancer potential of non-oncology drugs by systematic profiling. *Nat Cancer* 2020;1:235–48.
- Tsherniak A, Vazquez F, Montgomery PG, Weir BA, Kryukov G, Cowley GS, et al. Defining a cancer dependency map. *Cell* 2017;170:564–76.e16.
- Reinhold WC, Wilson K, Elloumi F, Bradwell KR, Cerbelli M, Varma S, et al. CellMinerCDB: NCATS is a web-based portal integrating public cancer cell line databases for pharmacogenomic explorations. *Cancer Res* 2023;83:1941–52.
- Smirnov P, Kofia V, Maru A, Freeman M, Ho C, El-Hachem N, et al. PharmacDB: an integrative database for mining in vitro anticancer drug screening studies. *Nucleic Acids Res* 2018;46:D994–1002.
- Feizi N, Nair SK, Smirnov P, Beri G, Eeles C, Esfahani PN, et al. PharmacDB 2.0: improving scalability and transparency of in vitro pharmacogenomics analysis. *Nucleic Acids Res* 2022;50:D1348–57.
- Zagidullin B, Aldahdooh J, Zheng S, Wang W, Wang Y, Saad J, et al. DrugComb: an integrative cancer drug combination data portal. *Nucleic Acids Res* 2019;47:W43–51.
- Zheng S, Aldahdooh J, Shadbahr T, Wang Y, Aldahdooh D, Bao J, et al. DrugComb update: a more comprehensive drug sensitivity data repository and analysis portal. *Nucleic Acids Res* 2021;49:W174–84.
- Seo H, Tkachuk D, Ho C, Mammoliti A, Rezaie A, Madani Tonekaboni SA, et al. SYNERGxDB: an integrative pharmacogenomic portal to identify synergistic drug combinations for precision oncology. *Nucleic Acids Res* 2020;48:W494–501.
- Chen Y, He L, Ianevski A, Ayuda-Durán P, Potdar S, Saarela J, et al. Robust scoring of selective drug responses for patient-tailored therapy selection. *Nat Protoc* 2024;19:60–82.
- Yadav B, Pemovska T, Szwajda A, Kuleskiy E, Kontro M, Karjalainen R, et al. Quantitative scoring of differential drug sensitivity for individually optimized anticancer therapies. *Sci Rep* 2014;4:5193.
- Ianevski A, Giri AK, Gautam P, Kononov A, Potdar S, Saarela J, et al. Prediction of drug combination effects with a minimal set of experiments. *Nat Mach Intell* 2019;1:568–77.
- Ianevski A, Giri AK, Aittokallio T. SynergyFinder 3.0: an interactive analysis and consensus interpretation of multi-drug synergies across multiple samples. *Nucleic Acids Res* 2022;50:W739–43.



29. Kumar A, Kankainen M, Parsons A, Kallioniemi O, Mattila P, Heckman CA. The impact of RNA sequence library construction protocols on transcriptomic profiling of leukemia. *BMC Genomics* 2017;18:629.
30. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr Protoc Mol Biol* 2015;109:21.29.1–21.29.9.
31. Corces MR, Buenrostro JD, Wu B, Greenside PG, Chan SM, Koenig JL, et al. Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nat Genet* 2016;48:1193–203.
32. Argelaguet R, Arnol D, Bredikhin D, Deloro Y, Velten B, Marioni JC, et al. MOFA+: a statistical framework for comprehensive integration of multimodal single-cell data. *Genome Biol* 2020;21:111.
33. Kryuchkova-Mostacci N, Robinson-Rechavi M. A benchmark of gene expression tissue-specificity metrics. *Brief Bioinform* 2017;18:205–14.
34. Yanai I, Benjamin H, Shmoish M, Chalifa-Caspi V, Shklar M, Ophir R, et al. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics* 2005;21:650–9.
35. Yu X, Lin J, Zack DJ, Qian J. Computational analysis of tissue-specific combinatorial gene regulation: predicting interaction between transcription factors in human tissues. *Nucleic Acids Res* 2006;34:4925–36.
36. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol* 1995;57:289–300.
37. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst* 2015;1:417–25.
38. Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 2013;14:7.
39. Szklarczyk D, Santos A, von Mering C, Jensen LJ, Bork P, Kuhn M. STITCH 5: augmenting protein-chemical interaction networks with tissue and affinity data. *Nucleic Acids Res* 2016;44:D380–384.
40. Mammoliti A, Smirnov P, Nakano M, Safikhani Z, Eccles C, Seo H, et al. Orchestrating and sharing large multimodal data for transparent and reproducible research. *Nat Commun* 2021;12:5797.
41. Lechner MG, Megiel C, Church CH, Angell TE, Russell SM, Sevell RB, et al. Survival signals and targets for therapy in breast implant-associated ALK-anaplastic large cell lymphoma. *Clin Cancer Res* 2012;18:4549–59.
42. Chen J, Zhang Y, Petrus MN, Xiao W, Nicolae A, Raffeld M, et al. Cytokine receptor signaling is required for the survival of ALK- anaplastic large cell lymphoma, even in the presence of JAK1/STAT3 mutations. *Proc Natl Acad Sci U S A* 2017;114:3975–80.
43. Cillessen SAGM, Hijmering NJ, Moesbergen LM, Vos W, Verbrugge SE, Jansen G, et al. ALK-negative anaplastic large cell lymphoma is sensitive to bortezomib through Noxa upregulation and release of Bax from Bcl-2. *Hematologica* 2015;100:e365–8.
44. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York (NY): ACM; 2016. p. 785–94.
45. Cichonska A, Pahikkala T, Szedmak S, Julkunen H, Airola A, Heinonen M, et al. Learning with multiple pairwise kernels for drug bioactivity prediction. *Bioinformatics* 2018;34:i509–18.
46. Janevski A, Nader K, Driva K, Senkowski W, Bulanova D, Moyano-Galceran L, et al. Single-cell transcriptomes identify patient-tailored therapies for selective co-inhibition of cancer clones. *Nat Commun* 2024;15:8579.
47. Frimantas V, Dobay MP, Rinaldi A, Tchinda J, Dunn SH, Kunz J, et al. Ex vivo drug response profiling detects recurrent sensitivity patterns in drug-resistant acute lymphoblastic leukemia. *Blood* 2017;129:e26–37.
48. Rindone G, Aroldi A, Bossi E, Verga L, Zambrotta G, Tarantino S, et al. A monocentric analysis of the long-term safety and efficacy of crizotinib in relapsed/refractory ALK+ lymphomas. *Blood Adv* 2023;7:314–6.
49. Fukano R, Mori T, Sekimizu M, Choi I, Kada A, Saito AM, et al. Alectinib for relapsed or refractory anaplastic lymphoma kinase-positive anaplastic large cell lymphoma: an open-label phase II trial. *Cancer Sci* 2020;111:4540–7.
50. Sanford M, Lyseng-Williamson KA. Nelarabine. *Drugs* 2008;68:439–47.
51. Rodriguez CO Jr, Stelrecht CM, Gandhi V. Mechanisms for T-cell selective cytotoxicity of arabinosylguanine. *Blood* 2003;102:1842–8.
52. DeAngelo DJ, Yu D, Johnson JL, Coutre SE, Stone RM, Stopeck AT, et al. Nelarabine induces complete remissions in adults with relapsed or refractory T-lineage acute lymphoblastic leukemia or lymphoblastic lymphoma: cancer and leukemia group B study 19801. *Blood* 2007;109:5136–42.
53. Johansson KB, Zimmerman MS, Dmytrenko IV, Gao F, Link DC. Idasanutlin and navitoclax induce synergistic apoptotic cell death in T-cell acute lymphoblastic leukemia. *Leukemia* 2023;37:2356–66.
54. von Jan J, Timonen S, Braun T, Jiang Q, Janevski A, Peng Y, et al. Optimizing drug combinations for T-PLL: restoring DNA damage and P53-mediated apoptotic responses. *Blood* 2024;144:1595–610.
55. Lakoma A, Barbieri E, Agarwal S, Jackson J, Chen Z, Kim Y, et al. The MDM2 small-molecule inhibitor RG7388 leads to potent tumor inhibition in p53 wild-type neuroblastoma. *Cell Death Discov* 2015;1:15026.
56. Lehmann C, Friess T, Birzele F, Kiialainen A, Dangel M. Superior anti-tumor activity of the MDM2 antagonist idasanutlin and the Bcl-2 inhibitor venetoclax in p53 wild-type acute myeloid leukemia models. *J Hematol Oncol* 2016;9:50.
57. Ge Y, Paisie TK, Chen S, Concannon P. UBASH3A regulates the synthesis and dynamics of TCR-CD3 complexes. *J Immunol* 2019;203:2827–36.
58. Morris SW, Kirstein MN, Valentine MB, Dittmer KG, Shapiro DN, Saltman DL, et al. Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma. *Science* 1994;263:1281–4.
59. Perl AE, Martinelli G, Cortes JE, Neubauer A, Berman E, Paolini S, et al. Gilteritinib or chemotherapy for relapsed or refractory FLT3-Mutated AML. *N Engl J Med* 2019;381:1728–40.
60. Nguyen J, Joseph J, Boccalon M, Aittokallio T, Haibe-Kains B. TCL38 PSet v1.3 [Internet]. Geneva (Switzerland):Zenodo; 2025 [cited 2025 Feb 18]. Available from: <https://zenodo.org/records/14733210>.
61. Potdar S, Janevski F, Janevski A, Tanoli Z, Wennerberg K, Seashore-Ludlow B, et al. Breeze 2.0: an interactive web-tool for visual analysis and comparison of drug response data. *Nucleic Acids Res* 2023;51:W57–61.
62. Hanahan D. Hallmarks of cancer: new dimensions. *Cancer Discov* 2022;12:31–46.
63. Hrustanovic G, Olivas V, Pazarentzos E, Tulpule A, Asthana S, Blakely CM, et al. RAS-MAPK dependence underlies a rational polytherapy strategy in EML4-ALK-positive lung cancer. *Nat Med* 2015;21:1038–47.
64. Hallberg B, Palmer RH. The role of the ALK receptor in cancer biology. *Ann Oncol* 2016;27(Suppl 3):iii4–15.
65. Wang Y, Qi YX, Qi Z, Tsang SY. TRPC3 regulates the proliferation and apoptosis resistance of triple negative breast cancer cells through the TRPC3/RASA4/MAPK pathway. *Cancers* 2019;11:558.
66. Schrader A, Crispatsu G, Oberbeck S, Mayer P, Pützer S, von Jan J, et al. Actionable perturbations of damage responses by TCL1/ATM and epigenetic lesions form the basis of T-PLL. *Nat Commun* 2018;9:697.
67. Andersson EI, Pützer S, Yadav B, Dufva O, Khan S, He L, et al. Discovery of novel drug sensitivities in T-PLL by high-throughput ex vivo drug testing and mutation profiling. *Leukemia* 2018;32:774–87.
68. Pützer S, Varghese L, von Jan J, Braun T, Giri AK, Mayer P, et al. Reinstated p53 response and high anti-T-cell leukemia activity by the novel alkylating deacetylase inhibitor tinostamustine. *Leukemia* 2020;34:2513–8.
69. Drexler HG. The leukemia-lymphoma cell line facts book [Internet]. London: Academic Press; 2001.
70. van der Meer D, Barthorpe S, Yang W, Lightfoot H, Hall C, Gilbert J, et al. Cell model passports—a hub for clinical, genetic and functional datasets of pre-clinical cancer models. *Nucleic Acids Res* 2019;47:D923–9.
71. Yadav B, Wennerberg K, Aittokallio T, Tang J. Searching for drug synergy in complex dose-response landscapes using an interaction potency model. *Comput Struct Biotechnol J* 2015;13:504–13.