Profiling drug sensitivity of leukemic stem cells via bulk-to-single-cell deconvolution

Alexandre Coudray[1,*], Romain Forey[1], Benjamin Bejar Haro[2], Filipe Martins[1,3], Joana Carlevaro-Fita[1], Shaoline Sheppard[1], Sandra Eloise Offner[1], Gioele La Manno[1,#,*], Guillaume Obozinski[1,2,#,*], and Didier Trono[1,#,*]

[1] *EPFL - Ecole Polytechnique Fédérale de Lausanne. 1015 Ecublens, Switzerland*

[2] *SDSC - Swiss Data Science Center. 1015 Ecublens, Switzerland*

[3] *Centre Hospitalier Universitaire Vaudois (CHUV), Hematology Service and Central laboratory, Oncology department, 1011 Lausanne, Switzerland.*

[*] *Corresponding authors*

Compiled September 29, 2024

***Ex-vivo* drug sensitivity screening allows the prediction of cancer treatment effectiveness in a personalized fashion. However, it only provides a readout on mixtures of cells, potentially occulting important information on clinically relevant cell subtypes. To address this shortcoming, we developed a machine-learning framework to decompose drug sensitivity recorded at the bulk level into cell subtype-specific drug sensitivity. We first determined that our method could decipher the cellular composition of bulk samples with top-ranking accuracy across five cancer types compared to state-of-the-art bulk deconvolution methods. We emphasize its effectiveness in the realm of Acute Myeloid Leukemia, where it appears to offer the most precise estimation of leukemic stem cell fractions across three test datasets and three patient cohorts. We then optimized an algorithm capable of estimating cell subtype- and single-cell-specific drug sensitivity, which we evaluated by performing *in-vitro* drug studies and in-depth simulations. We then applied our deconvolution strategy to the beatAML cohort dataset, currently the most extensive database of *ex-vivo* drug screening data. We developed a drug sensitivity profile tailored to specific cell subtypes, focusing on four therapeutic compounds predicted to target leukemic stem cells: the previously known midostaurin and A-674563, as well as SNS-032 and foretinib, which have not been previously linked to leukemic stem cells. Our work provides an attractive new computational tool for drug development and precision medicine.**

INTRODUCTION

Ex-vivo drug sensitivity screening (DSS) allows for the parallel testing of hundreds of compounds on cells extracted from a human tumor, potentially paving the way to personalized therapeutic decisions [1–3]. However, as currently performed, DSS assesses the susceptibility of cell mixtures and not of individual cell subtypes, hence fails to provide information that might be of utmost clinical importance, for instance on the drug sensitivity of rare cancer cells with stem-like properties. Coupling *ex-vivo* DSS with cell purification techniques such as fluorescence-activated cell sorting (FACS) can partly address this shortcoming, but sorting requires large quantities of cells, is limited by the availability of specific cell surface markers, and the procedure itself can impair the viability of some cell subtypes [4]. Therefore, computational methods capable of deconvoluting *ex-vivo* DSS would constitute attractive alternatives, as they would alleviate the need for cell subtypes purification.

Acute Myeloid Leukemia (AML) is a cancer type typically displaying extensive cellular heterogeneity, with the bone marrow of AML patients harboring a mixture of more than twenty cell subtypes, along with approximately 60 recurrently mutated genes and numerous rare mutations [5–7]. Conventional chemotherapies, involving anthracyclines and nucleoside analogs followed by bone marrow transplant, yield limited durable remission (around 20%) [8]. A persistent challenge in AML is drug resistance and relapse, urging a deeper grasp of the biological factors underpinning drug response. Notably, drug response frequently aligns with tumor cell maturation state, with certain drugs effective against less differentiated states [9–15], and others against more differentiated states [9, 12, 13, 16–23]. Recent single-cell sequencing further delineated AML cancer cell subtypes tied to differentiation state, including hematopoietic-stem-cell-like, progenitor-like, GMP-like, promonocyte-like, monocyte-like, and cDC-like types [24], several of which are directly relevant to the oncogenic process. Notably, the hematopoietic-like cancer cell subtype identified by Van Galen et al. (HSC-like) likely encompasses previously characterized leukemic stem cells (LSCs), thought to be responsible for relapse and decreased survival in AML patients [25–28]. Gene expression signatures from these leukemic stem cells (LSCs) also offer valuable prognostic insights [29–31], underscoring their clinical importance. Given this context, we posit that innovative approaches to elucidate the drug sensitivity of distinct cancer cell subpopulations could significantly enhance AML treatment strategies and potentially offer new avenues for targeting LSCs.

In this work, we introduce novel computational techniques that enable the concurrent estimation of cell subtype proportions and the deconvolution of *ex-vivo* drug sensitivity screening data into cell subtype- or single-cell-specific drug sensitivity profiles. We first developed CLIMB, a tool designed to decipher the cell subtype composition and expression within AML bulk samples, using a novel bulk-to-single-cell deconvolution approach. The outcome of CLIMB's deconvolution serves as direct input for our subsequent method, CLIFF. CLIFF employs a probabilistic model to uncover latent variables corresponding to cell subtype survival probability, and thus deconvolute bulk-level drug sensitivity into cell-subtype-specific drug sensitivity profiles.

We benchmarked our methods through the analysis of cancer cell line mixes we generated *in vitro*, and through pseudo-bulk deconvolution based on ten single-cell datasets across five cancer types. We then deployed CLIMB on three cohorts of AML patients, and used CLIFF to deconvolute cell-subtype drug sensitivity based on the beatAML cohort [7]. This led us to predict that the drugs A-674563, midostaurin, SNS-032, and foretinib would be effective in targeting HSC-like cancer cells. Further investigation of single-cell drug sensitivity predictions for venetoclax, a clinically used drug for AML, highlighted a subpopulation of venetoclax-resistant HSC-like cells.

RESULTS

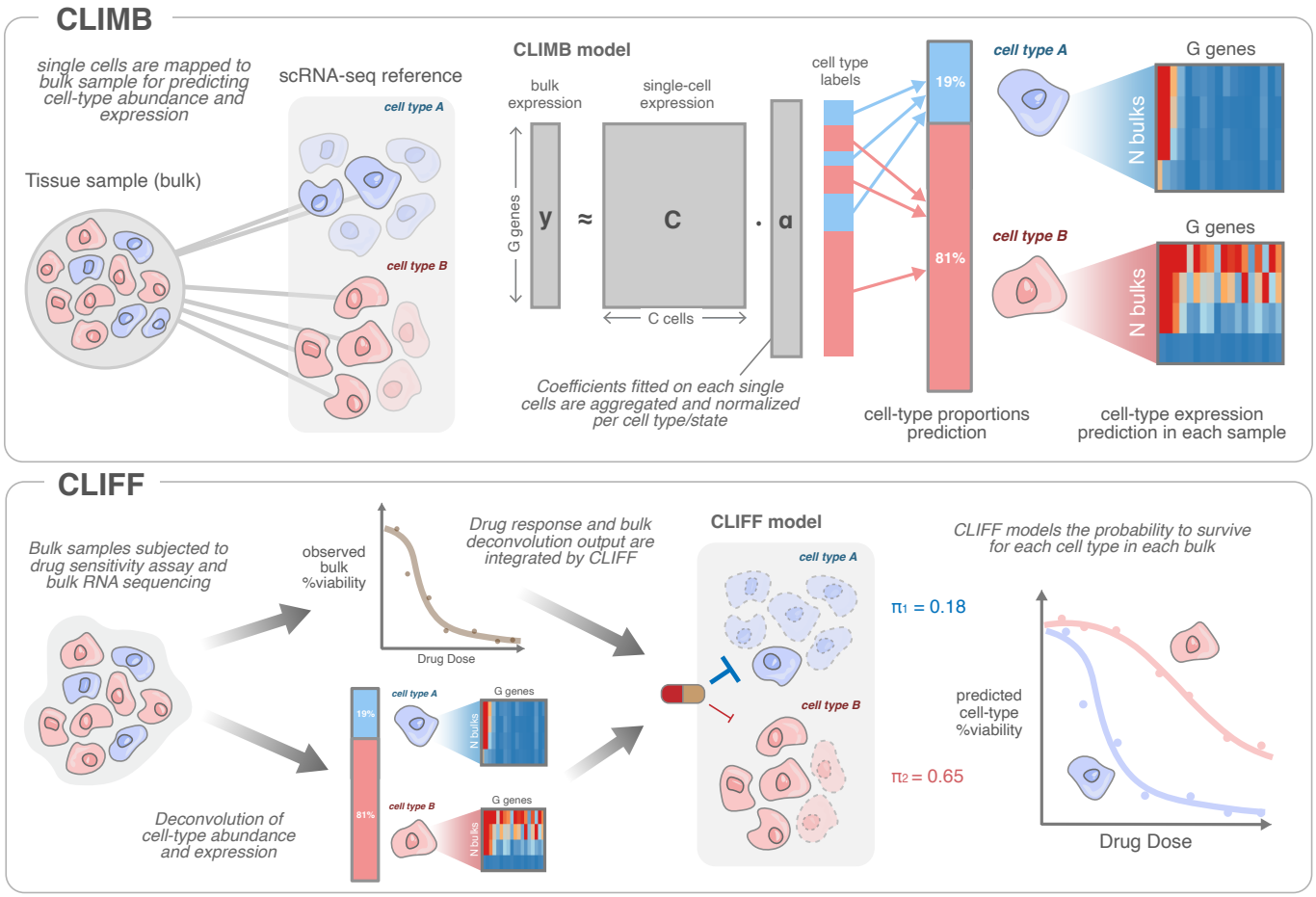


Fig. 1. Experimental Design. (**Top**) CLIMB conducts bulk deconvolution using the full expression matrix from scRNA-seq, fitting coefficients on individual single cells, which are then aggregated at the cell-subtype level, allowing the inference of cell-subtype proportions and expression. (**Bottom**) Presented is the two-step process of our framework 'CLIMB the CLIFF'. Initially, CLIMB deconvolution method is used to estimate cell subtype proportions and expression in each mixed sample. Subsequently, deconvoluted cell subtype abundance and expression are combined with drug sensitivity data to predict cell subtype drug sensitivity. In a nutshell, CLIFF employs a probabilistic model to uncover latent variables corresponding to cell subtype survival probability.

Experimental design

Studying drug sensitivity in cancer cell subpopulations is essential yet challenging, employing various experimental approaches. Patient-derived xenograft models mimic patient heterogeneity and drug responses, but establishing them is resource-intensive. In contrast, *ex-vivo* drug sensitivity screening (DSS) assesses drug

response in surgically extracted cancer cells, preserving their inherent diversity. Combined with fluorescence-activated cell sorting (FACS), DSS allows subpopulation-level analysis. However, FACS requires expensive antibodies and has limitations in isolating cell subtypes. Previous work by Bottomly et al. highlighted AML cell differentiation's impact on drug sensitivity [32], while Zeng et al. linked HSC-like cell proportions (predicted using CiberSortX deconvolution method) to drug responses and outcomes [33]. In contrast, Karakaslar et al. showed that Monocyte-like content correlates with venetoclax resistance [34]. Yet, these studies relied solely on correlations between bulk-level drug sensitivity and cell subtype proportions from deconvolution methods. To enhance drug sensitivity deconvolution at the cellular level, we introduce a novel approach that decomposes bulk-level cell viability data into interpretable cell viability metrics at both cell subtype and single-cell resolutions.

Decoding drug sensitivity in cell subpopulations from heterogeneous mixtures requires knowledge or prediction of cellular composition. To address this challenge, we present CLIMB, a novel method that dissects the cellular composition of bulk samples (Fig. 1, upper part, and SFig. 1A, B). CLIMB employs a unique model that directly maps single cells from a reference dataset to the target bulk expression, aiming to optimize single-cell combinations. This approach makes CLIMB a signature matrix-free deconvolution method, sequentially mapping bulk samples to single cells and further to cell subtypes. Utilizing an iterative deconvolution process with scRNA-seq reference data and bulk RNA-seq expression as input, CLIMB can predict cell-subtype proportion and expression from a bulk RNA-seq sample. We utilized CLIMB to address a pivotal and clinically significant question: Can we accurately evaluate the fraction of leukemic stem cells contained in bulk samples from AML patients?

We further developed CLIFF, a probabilistic model that unveils latent variables corresponding to cell subtype cell viability, by leveraging cell-subtype proportions and expressions as input (Fig. 1, bottom, SFig. 1A). CLIFF's framework defines cell viability as the ratio of surviving cells, adopting a Bernoulli distribution assumption for cell fate. CLIFF can thus directly integrate CLIMB's output with cell viability assay data performed at the bulk level. Optionally, CLIFF incorporates somatic mutations as covariates and can predict drug sensitivity in 'high-resolution' mode by deconvoluting cell-subtype drug sensitivity within each bulk sample (SFig. 1B). We then tested CLIFF's capability to predict the drug sensitivity of leukemic stem cells, aiming to assess susceptibilities of this crucial AML-specific cell subtype.

We introduce a dual approach, sequentially employing CLIMB and CLIFF, to dissect the cellular composition of tumor samples and unravel drug sensitivity at the cell subtype level, utilizing only bulk RNA-seq data along with bulk drug sensitivity information. Employing *in vitro* mixtures of human leukemic cell lines (HL60, SUDHL4, K562, and THP1), we assessed our bulk deconvolution in a known cell subtype fraction context, where individual cell line drug sensitivities can be experimentally determined, providing ground truth. By applying bulk RNA-seq alongside bulk *ex-vivo* drug screening data from the beatAML cohort, we translated bulk-level drug sensitivity into cell subtype-specific profiles. Our study then focus in a detailed exploration of

the drug sensitivity specific to the leukemic stem cells, a cell subtype of utmost clinical significance.

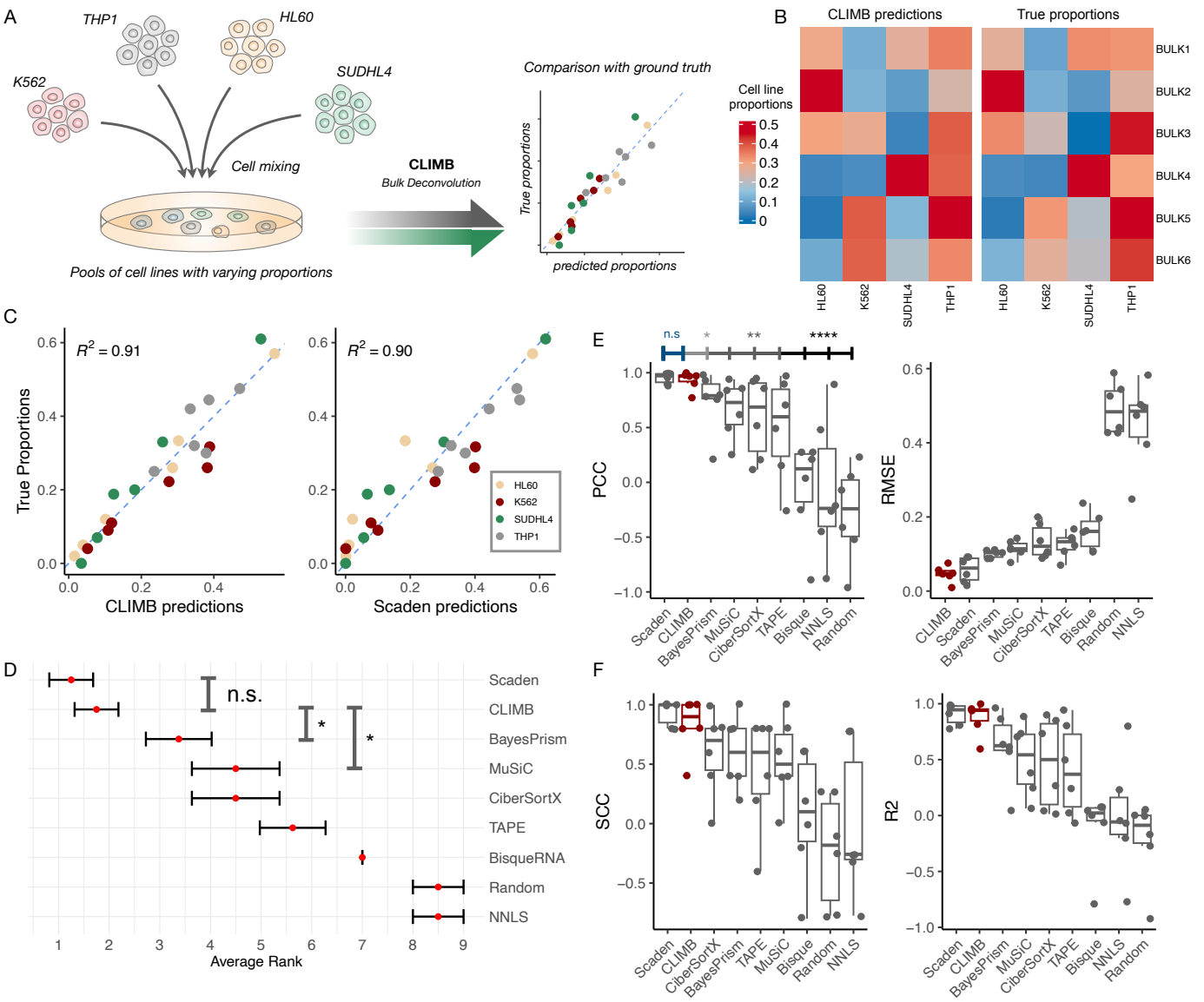


Fig. 2. Cell Line Abundance Deconvolution using an *In-Vitro* Experiment with Known Proportions. (**A**) Experimental design for *in-vitro* experiment involving mixtures of human leukemic cell lines. (**B**) Heatmap displaying predicted cell line proportions by CLIMB (left) and actual proportions (ground truth, right). (**C**) Scatter plot comparing true versus predicted cell line abundances using CLIMB (left) and Scaden (right), with colors representing cell lines (6 bulk samples and 4 cell lines). (**D**) Overall results ranked based on median of four metrics (PCC, SCC, RMSE, and R-square) for bulk sample-level comparison. Wilcoxon rank test used for statistical comparison between CLIMB and other methods. (**E**) Left: Comparison of deconvolution accuracy using Pearson correlation coefficient (PCC) between predicted and true cell line proportions. Statistical significance assessed using Pearson and Filon method with *cocor* R package. Left boxplot depicts PCC *per sample* (4 true versus 4 predicted cell line abundances for 6 bulk samples). Right: same approach using Spearman Correlation Coefficient. (**F**) Left: Similar to (**E**) but for R^2 metric. Right: similar to (**E**) but for root mean square error (RMSE) metric.

Evaluating Bulk Deconvolution Methods for Predicting Cellular Composition in Cancer Cell Line Mixtures

CLIMB infers cell-subtype proportions and can thus predict the composition of a cell mix (Fig. 2A). We generated *in vitro* pools containing various ratios of HL60, SUDHL4, K562, and THP1 human leukemic cells, analyzed by standard bulk RNA sequencing. We could thus use these pre-determined ratios as ground

truth to compare the predictive value of CLIMB (Fig. 2B, C) and a panel of seven bulk deconvolution methods: BayesPrism [35], CiberSortX [36], TAPE [37], Scaden [38], MuSiC [39], BisqueRNA [40], and NNLS [39], using random proportions as a negative control. We then performed single-cell RNA sequencing of a mixture of the same cell lines, and used it as reference for all deconvolution methods with the same cell and gene selection (employing genes with higher estimated biological variance than technical variance, see Methods). The overall ranking among four metrics (average over 4 ranks), placed CLIMB and Scaden as the best deconvolution methods (Fig. 2D), followed by BayesPrism. Notably, CLIMB reached a higher accuracy than other deconvolution methods based on the root mean square error (RMSE, Fig. 2E). Scaden reached similar PCC, SCC, and R-squared, than CLIMB (Fig. 2C, right panel). A wilcoxon test taking four metrics into account showed significant advantage of CLIMB over other deconvolution methods, except Scaden. In summary, CLIMB demonstrated top-ranking accuracy in deconvoluting bulk RNA-seq samples from *in-vitro* cell mixtures, matching the performance of the deep learning-based method Scaden and surpassing other deconvolution methods.

Evaluating Potential Bias Towards Cellular Proportions Contained In Reference Dataset

The deconvolution process of CLIMB, which employs the entire single-cell matrix as input, may introduce bias mirroring the cellular proportions in the reference matrix. To mitigate such bias, CLIMB's model executes three sequential deconvolutions. After deconvoluting cellular proportions during its first iteration, CLIMB subsets cells in the reference matrix based on these proportions, which subsequently serve as a reference for the next deconvolution. This process is iterated twice, progressively adjusting the reference matrix to align its cellular proportions with those of the target bulk sample (SFig. 2, Methods). To demonstrate CLIMB's model accuracy and robustness, we studied how the correlation between cellular proportions in the reference matrix and target bulk samples impacts deconvolution accuracy. We simulated eleven scenarios ranging from strong to zero-correlation contexts, using reference and pseudo-bulk data derived from AML cells in the Van Galen et al. dataset (SFig. 3A). We examined accuracy and resilience to induced bias, which we inferred from the variability in accuracy across our range of bias contexts. By comparing CLIMB (coined 'CLIMB-corrected') with the first iteration of CLIMB algorithm ('CLIMB-init') to a panel of seven deconvolution methods, we observed that CLIMB was the second most accurate method, and the third less biased method (SFig. 3B-D). Interestingly, CLIMB demonstrated enhanced accuracy and resilience to bias over CLIMB-init in most cross-dataset pseudo-bulk analysis, particularly in cases with significant cellular proportion imbalances between reference and target bulk (SFig. 3E-F). Overall, we demonstrate that CLIMB iterative deconvolution exhibits high accuracy while maintaining low bias.

Benchmarking Bulk Deconvolution Methods using scRNA-seq Derived Pseudo-bulk Samples in a Cross-Dataset Context

To extend our in vitro findings, we conducted cross-dataset deconvolution by integrating five pairs of scRNA-seq datasets in five cancer types, totaling ten cross-dataset analyses [24, 41–48]. These cross-dataset analyses

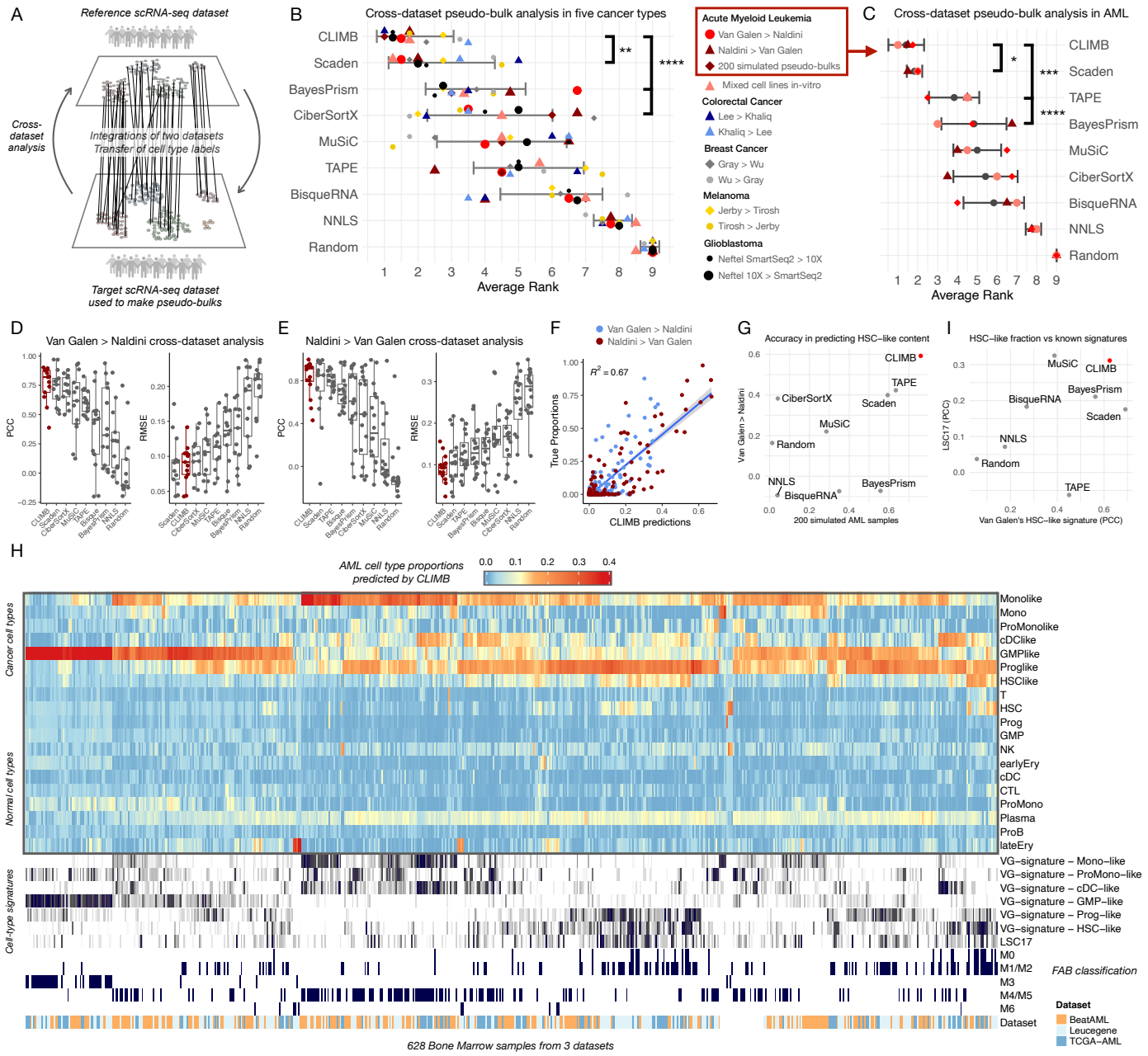


Fig. 3. Cell Subtype Composition Deconvolution in Acute Myeloid Leukemia. (**A**) Experimental design outline for cross-dataset deconvolution analysis. (**B**) Average rank across four metrics (PCC, SCC, R2, and RMSE) for 10 cross-dataset pseudo-bulk deconvolutions, including in-vitro mixes and 200 simulated AML samples, totaling 12 analyses. (**C**) Subset of (**B**), focusing on the four AML context analyses. (**D**) Pseudo-bulk analysis using Van Galen dataset as a reference for Naldini's pseudo-bulks. Boxplots depict PCC/RMSE values for each pseudo-bulk sample (13 samples, 19 cell subtypes' predictions). (**E**) Similar to (**D**), but in the Naldini > Van Galen context, with 16 data points for the 16 samples in the Van Galen dataset. (**F**) Scatter plot comparing true proportions with deconvoluted proportions using CLIMB (left) and Scaden (right) for both Van Galen > Naldini and Naldini > Van Galen analyses. (**G**) Predictive accuracy of HSC-like cell subtype proportions for Van Galen > Naldini analysis, comparing 13 predicted proportions with true proportions. (**H**) Inferred cell-subtype proportions using CLIMB for bone marrow in BeatAML, Leucegene, and TCGA-AML cohorts (N=628), along with scores of seven signatures (LSC17 and six Van Galen signatures) and FAB labels.

first involved transferring cell-subtype labels between datasets. Then, one dataset served as the reference single-cell dataset while the other was used to generate pseudo-bulk samples (Fig. 3A). Importantly, the reference dataset and the pseudo-bulk datasets were constructed using only non-batch-corrected raw counts to preserve a realistic batch effect across datasets. The comprehensive results of this large-scale analysis demonstrated that CLIMB significantly outperformed other deconvolution methods (Fig. 3B), based on an aggregated ranking across four evaluation metrics (PCC, SCC, RMSE, and R2). In the context of Acute Myeloid Leukemia (AML), CLIMB outperformed other methods (Fig. 3C) in (i) 'Van Galen > Naldini' deconvolution, with CLIMB achieving the best PCC, SCC, and R2 (Fig. 3D, F, SFig. 4E), (ii) 'Naldini > Van Galen' deconvolution, where CLIMB achieved the highest PCC, RMSE, and R2 (Fig. 3E, F, SFig. 4G), and (iii) '200 simulated pseudo-bulks', a dataset of 200 simulated bulk samples generated from the Van Galen et al. dataset, where CLIMB demonstrated the top scores for all four metrics (SFig. 4F). Interestingly, CLIMB demonstrated superior performance in predicting the proportion of leukemic stem cells ('HSC-like' in Van Galen dataset) in both analysis with available HSC-like labels (Fig. 3G). Moreover, CLIMB ranked first for CRC cancer analysis (SFig. 4H-J). CiberSortX and CLIMB emerged as the best methods in Melanoma analysis ('Tirosh > Jerby' and 'Jerby > Tirosh', SFig. 4K-M). In the Glioblastoma context, CLIMB was the top-performing method both in the '10x > SmartSeq2' and the 'SmartSeq2 > 10x' analysis (SFig. 4N-P). In breast cancer context, MuSiC performed best, based on 'Gray > Wu' and 'Wu > Gray' analyses (SFig. 4Q-S). Overall, these findings highlight CLIMB's top-ranking accuracy in cross-dataset pseudo-bulk analyses, especially in AML and Glioblastoma contexts, with strong evidence of high accuracy in estimating the HSC-like cell subtype population.

Bulk Deconvolution of Cell Subtype Abundance Across Three AML Patient Cohorts

We performed deconvolution of cell subtype proportions using bulk RNA-seq samples from three AML cohorts: beatAML [7], Leucegene [49], and TCGA-AML [50], using as a reference only cells in pre-treatment conditions from Van Galen dataset [24]. Employing CLIMB, we delineated a landscape of cell subtype proportions in AML (Fig. 3H). To assess deconvolution accuracy in patient samples, we juxtaposed CLIMB-derived proportions with FAB classification [51], a scheme categorizing patients according to blast (cancer cell) differentiation patterns. Specifically, blasts from M0/M1 patients, characterized by minimal differentiation, were expected to correlate with higher proportions of HSC-like and/or Progenitor-like cancer cells, a correspondence verified by CLIMB-derived proportions (Fig. 3H, SFig. 5A). Similarly, M3 samples (PML-RARA classified patients) were anticipated to exhibit enrichment for GMP-like cells, in line with our deconvoluted GMP-like proportions (SFig. 5A). Moreover, we observed monocyte-like cancer cells to be enriched in M4/M5 patients, along with erythrocyte enrichment in M7 patients, both of which were expected (SFig. 5A). To evaluate CLIMB's accuracy in predicting leukemic stem cells, we correlated deconvoluted HSC-like proportions with the LSC17 signature [31]. As expected, the highest correlations were observed with Prog-like and HSC-like cell subtypes (SFig. 5B). Interestingly, proportions of HSC-like cells predicted by CLIMB showed a noticeable alignment with both the HSC-like signature established by Van Galen et al. [24] (SFig. 5C) and the LSC17 signature, when compared to

other deconvolution methods (Fig. 3I). To evaluate the consistency of results across different deconvolution algorithms, we assessed cross-method correlations across all three AML cohorts. We observed that the top three methods—CLIMB, Scaden, and CiberSortX—formed a distinct cluster (SFig. 5D). In summation, CLIMB provides a robust depiction of the cellular landscape of AML patients across three cohorts, laying foundations for deconvoluting cell subtype drug sensitivity.

Benchmarking Bulk Expression Deconvolution into Cell Subtype Expression and Simulated Over-Expression

CLIMB enables the deconvolution of bulk-level RNA expression into cell subtype-specific expression. This can be achieved through either an 'overall' approach, which accounts for expression averaged across all bulks used as input, or a 'high-resolution' mode that disentangles cell subtype expression within each bulk sample. In our in-vitro experiment with leukemic cell lines, CLIMB's 'overall' deconvoluted expression achieved high accuracy when compared to bulk RNA-seq expression obtained from pure cell lines (PCC of 0.94, SFig. 6A-C). BayesPrism displayed similar patterns and accuracy to CLIMB, achieving a PCC of 0.92 (SFig. 6C). Subsequently, we deconvoluted AML cell-subtype expression based on the 'Van Galen > Naldini' cross-dataset framework. Notably, the 'overall' mode of CLIMB demonstrated compelling deconvoluted expression patterns for cancer cell markers (PCC of 0.75, SFig. 6D-F), surpassing the performance of BayesPrism (0.59) and CiberSortX in 'group mode'. We then subjected CLIMB's deconvoluted expression to a 'high-resolution' test, simulating the over-expression of 20 genes in the HSC-like subtype. This induction was applied to 7 out of 13 pseudo-bulk samples, with varying fold changes. The simulated expression was then compared to true expression values, showing CLIMB's deconvoluted expression high accuracy at fold change 10 and higher (SFig. 6G). To assess the precision in detecting differentially expressed genes (DEGs), we conducted a DE analysis between wild-type (WT) and over-expression (OE) conditions. We employed the AUC-ROC metric to evaluate DEG detection accuracy (SFig. 6J-K). BayesPrism exhibited the highest accuracy in recovering DEGs followed by CLIMB (SFig. 6J). As a positive control, we checked that true simulated expression efficiently recovered DEGs (AUC > 0.95 with fold change > 2, SFig. 6J, K). We then computed correlations (PCC) between true and deconvoluted expression (SFig. 6L), showing BayesPrism as the most accurate (median PCC of 0.76), followed by CLIMB (0.66). In summary, CLIMB demonstrates accuracy in expression deconvolution, particularly in 'overall' mode.

Benchmarking Drug Sensitivity Deconvolution with In-Vitro Cell Line Mixtures

In this next step, we integrate drug sensitivity data (measured as a cell viability rate) with CLIMB's bulk deconvolution results. To achieve this, we developed CLIFF, a method that deconvolutes bulk-level cell viability into cell-subtype-specific cell viability (Fig. 1, 4A). We generated drug sensitivity data from *in vitro* mixtures comprising six combinations of four human leukemia cell lines, as previously depicted in Fig. 2 (HL60, SUDHL4, K562, and THP1). These mixtures were exposed to seven drugs commonly employed in AML consolidation therapies (5-Aza-2'-deoxycytidine, cytarabine, imatinib, nilotinib, doxorubicin, rapamycin, mitomycin C). To assess the accuracy of drug sensitivity deconvolution, we determined the ground truth cell

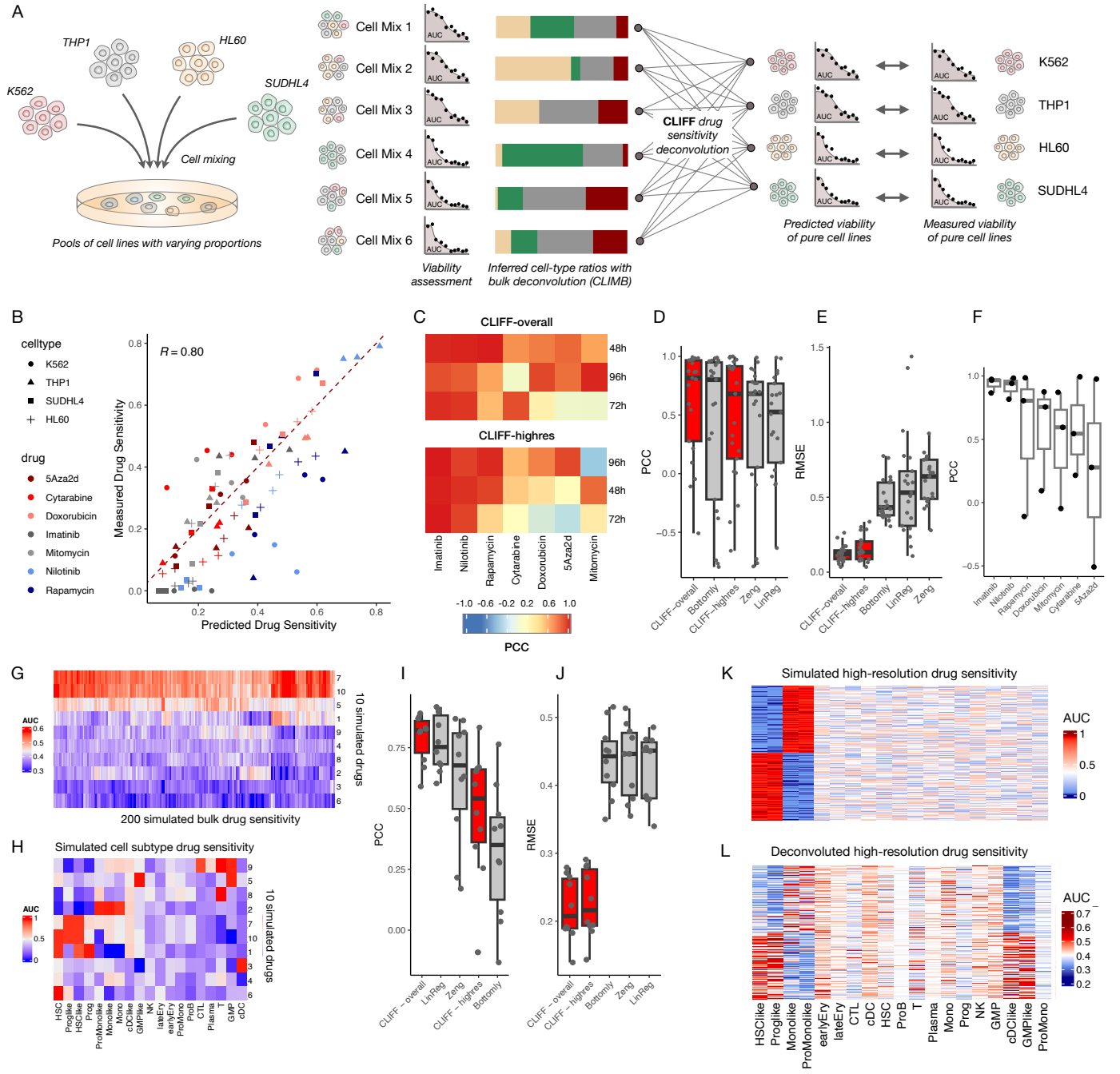


Fig. 4. Deconvolving Cell Line Drug Sensitivity: *In-Vitro* and Simulated AML Context. (**A**) Experimental setup: Known cell line mixtures are deconvoluted to estimate cell abundances, which are integrated with bulk drug sensitivity for predicting individual cell line drug responses. (**B**) Predicted vs. actual cell line drug sensitivity for multiple timepoints, cell lines, and drugs. (**C**) Heatmap displays predictive accuracy as PCCs for each drug and timepoint. (**D**) Method comparison of deconvolution accuracy using PCCs. (**E**) Method comparison of deconvolution accuracy using RMSE values. (**F**) CLIFF's predictive accuracy as PCC, grouped by drugs. (**G**) Drug sensitivity simulation on 200 pseudo-bulk samples derived from cell subtype sensitivities (**H**). (**H**) Simulated cell subtype drug sensitivity pattern for 10 drugs, used for generating bulk drug sensitivity. (**I**) Predictive accuracy as PCC for CLIFF and other methods, deconvoluting cell subtype sensitivity (**G**) using bulk drug sensitivity (**H**). (**J**) Same as (**I**), but using RMSE metric. (**K**) Simulated high-resolution drug sensitivity across pseudo-bulk samples and cell subtypes. (**L**) Deconvoluted high-resolution drug sensitivity based on bulk samples from (**K**).

viability of the four individual cell lines when exposed individually to the same drug panel, doses, and time points. We initially evaluated the theoretical basis of CLIFF's model by reconstructing the experimentally-assessed cell mix viability from the experimentally-determined viability of pure cell lines. This was achieved using CLIFF's model equations (Equations 15, and 16) to correlate both metrics (SFig. 7A), which showed a high Pearson's correlation of 0.88. In the next step, we evaluated CLIFF's deconvoluted cell viabilities against ground truth, showing a PCC of 0.80 across all 21 conditions (Fig. 4B, C). When utilizing high-resolution cell-subtype expression as input ('CLIFF-highres'), the accuracy of CLIFF was diminished (PCC of 0.59, Fig. 4C, SFig. 7B). Subsequently, we conducted a comparison of CLIFF's accuracy with the approach employed by Zeng et al. [34], which relies on a correlation between cell subtype proportions and bulk-level drug sensitivity; with Bottomly et al.'s approach [32], wherein the PC1 of cell subtype signatures is correlated with bulk drug sensitivity; and a linear regression using CLIMB proportions as input and bulk-level cell viability as ouput (LinReg). When assessing drugs and timepoints individually, CLIFF exhibited the highest correlation (median PCC of 0.82, Fig. 4D), followed by Bottomly (0.80), 'CLIFF-highres' (0.68), Zeng (0.68), and LinReg (0.51). Interestingly, CLIFF is the only method that predicts cell viability rather than providing a drug sensitivity score based on correlation metrics, which results in a significantly lower RMSE compared to all other methods, both in 'overall' and 'high-resolution' modes (Fig. 4E). Thus, within a framework comprising six cell mixes encompassing four cell lines, CLIFF exhibited precise deconvolution of drug sensitivity and proved effective to integrate cell subtype proportions estimated via CLIMB.

Simulated *Ex-Vivo* Drug Sensitivity Assay for Benchmarking Deconvolution of Drug Sensitivity

We then simulated bone marrow samples from AML patients to generate a benchmarking dataset similar to the beatAML cohort; as the latter lacks ground truth data for testing deconvolution. Using the Van Galen single-cell dataset, we generated 200 pseudo-bulk RNA-seq samples along with their corresponding bulk cell viability data for ten simulated drugs (Fig. 4G). To do this, we simulated cell viability for 19 cell subtypes across ten drugs (Fig. 4H), which serves as our ground truth for evaluating CLIFF's deconvolution accuracy. We proceeded to assess the accuracy of drug sensitivity deconvolution using the same five approaches mentioned earlier (Fig. 4D-E). In this context, CLIFF-overall demonstrated the highest accuracy, with a median PCC of 0.82 (Fig. 4I), followed by LinReg (0.75), Zeng (0.68), CLIFF-highres (0.54), and Bottomly (0.35). As observed in the in-vitro experiment, CLIFF's advantage, both in overall and high-resolution modes, lies in generating realistic cell viability values, which translates into significantly lower RMSE values compared to every other methods (Fig. 4J). Investigating the effect of the number of input genes on CLIFF's performance, we found that CLIFF-overall maintained consistent accuracy regardless of the gene subset (SFig. 7E), while CLIFF-highres excelled particularly with higher gene counts (>500). Additionally, introducing L2 regularization improved CLIFF-highres predictions, although it had no significant impact on CLIFF-overall predictions (SFig. 7E). The impact of input size on CLIFF's accuracy was examined by varying the number of bulk samples (SFig. 7E). Notably, approximately 25 cell mixes proved sufficient for CLIFF to achieve a PCC > 0.6 when deconvoluting 19

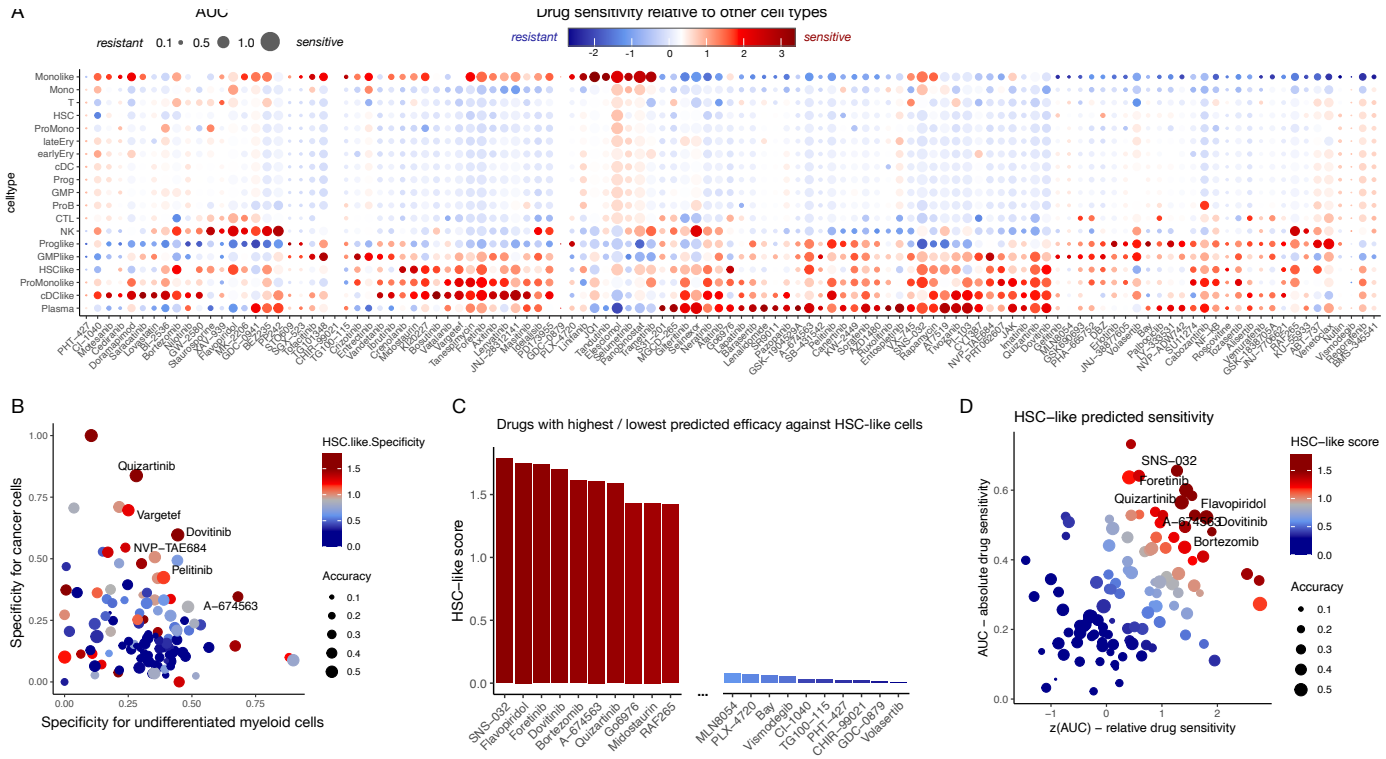


Fig. 5. Deconvolving Cell Subtype-Specific Drug Sensitivity in AML *ex-vivo* Data. (**A**) Deconvoluted drug sensitivity across cell subtypes in the beatAML cohort dataset, with dot size indicating AUC of sensitivity and color representing standardized AUC, indicating relative drug sensitivity among cell subtypes for a specific drug. (**B**) Differentiated (Mono/ProMono/cDC) vs. undifferentiated (HSC/Prog/GMP) myeloid cells' sensitivity (grouped) on the x-axis and normal vs. cancer myeloid cells' sensitivity on the y-axis. (**C**) HSC-like sensitivity score computed from AUC and standardized AUC. (**D**) Drug sensitivity of HSC-like cells shown on a scatter plot with AUC on the x-axis, standardized AUC on the y-axis, and color indicating the combined score. (**E**) High-resolution drug sensitivity landscape for venetoclax across patients, with boxplots showing distribution among patients, and bottom heatmap displaying somatic mutation presence. (**F**) Similar view as (**E**), but for A-674563 compound.

cell subtypes, and about 75 input mixtures were required to attain a PCC > 0.7. By simulating patient-specific mutations (SFig. 7F) that were associated with heightened drug sensitivity in HSC-like/Prog-like cell subtypes and lowered sensitivity in Mono-like/ProMono-like cell subtypes, we examined how well CLIFF-highres could discern patient-level variability (Fig. 4K, L). Our observations showed that CLIFF-highres detected shifts in drug sensitivity between the two sample groups in the correct cell subtypes (HSC-like, Prog-like, Mono-like, and ProMono-like). However, it also introduced false positive predictions in cDC-like and GMP-like subtypes. In summary, based on our benchmarking of CLIFF in both in-vitro experiments and simulated contexts, it appears that CLIFF's 'overall' mode effectively predicts averaged cell-subtype drug sensitivity. Therefore, we will limit the use of CLIFF to its 'overall' mode in the next steps.

Deconvoluting Bulk-Level Drug Sensitivity to Cell Subtype Sensitivity Using beatAML Cohort

Utilizing cell subtype proportions and expression data estimated by CLIMB from the beatAML cohort, we integrated this information with bulk *ex-vivo* drug sensitivity data using CLIFF, which revealed a heterogeneous landscape of drug sensitivity patterns among distinct cell subtypes (Fig. 5A). We assessed the findings based on 14 studies that documented cell-subtype level drug sensitivity and featured 22 pairs of drug-to-cell subtype interactions (Supplementary Tab. 1). CLIFF's predictions aligned with existing literature for venetoclax [12, 13], erlotinib [16], A-674563 [17], AZD1480 [18, 19], and crenolanib [20], indicating elevated sensitivity in HSC-like and Prog-like cancer cells. Interestingly, midostaurin [21] and selinexor [22] were predicted to target HSC-like cell subtypes, but not progenitor cancer cells (Prog-like). We also documented drug sources that targeted differentiated blasts. Doramapimod, predicted by CLIFF to target monocyte-like cells, aligns with previous evidence of heightened activity in M4/M5 patients showing expression of CD14 monocyte marker [11]. Similar agreement was observed for trametinib [12] and selumetinib [9], known to target differentiated blasts, and projected by CLIFF to exhibit strong affinity for monocyte-like cancer cells. Bromodomain and Extraterminal Domain Inhibitors (BETi) were also shown by Romine et al. [10] to target differentiated blasts, which corroborates CLIFF's predictions on BET-inhibitor JQ1, predicted to preferentially target monocyte-like cancer cells. Ruxolitinib [12], however, did not display a clear pattern, as it was predicted to target pro-monocyte-like but not monocyte-like cancer cells. CLIFF's predictions also indicate a notable resistance of monocyte-like cells to venetoclax, validated by four distinct studies [11–15]. Among the 22 documented instances of cell subtype-specific drug sensitivity, CLIFF demonstrated 16/22 cases with clear concordance (72.7%), 3/22 with partial alignment, and 3/22 in direct contradiction (13.6%). Furthermore, we evaluated the concordance between CLIFF's predictions and results reported by Bottomly et al. using the same dataset [32], revealing an overall correlation of 0.44 for all drugs (SFig. 8A), with stronger correlations observed for established drug-to-cell subtype associations (SFig. 8B).

Subsequently, we focused on evaluating drug sensitivity in HSC-like cells, ranking drugs according to their potential effectiveness against this particular cell subtype (Fig. 5B-D). Among the top ten drugs, only A-

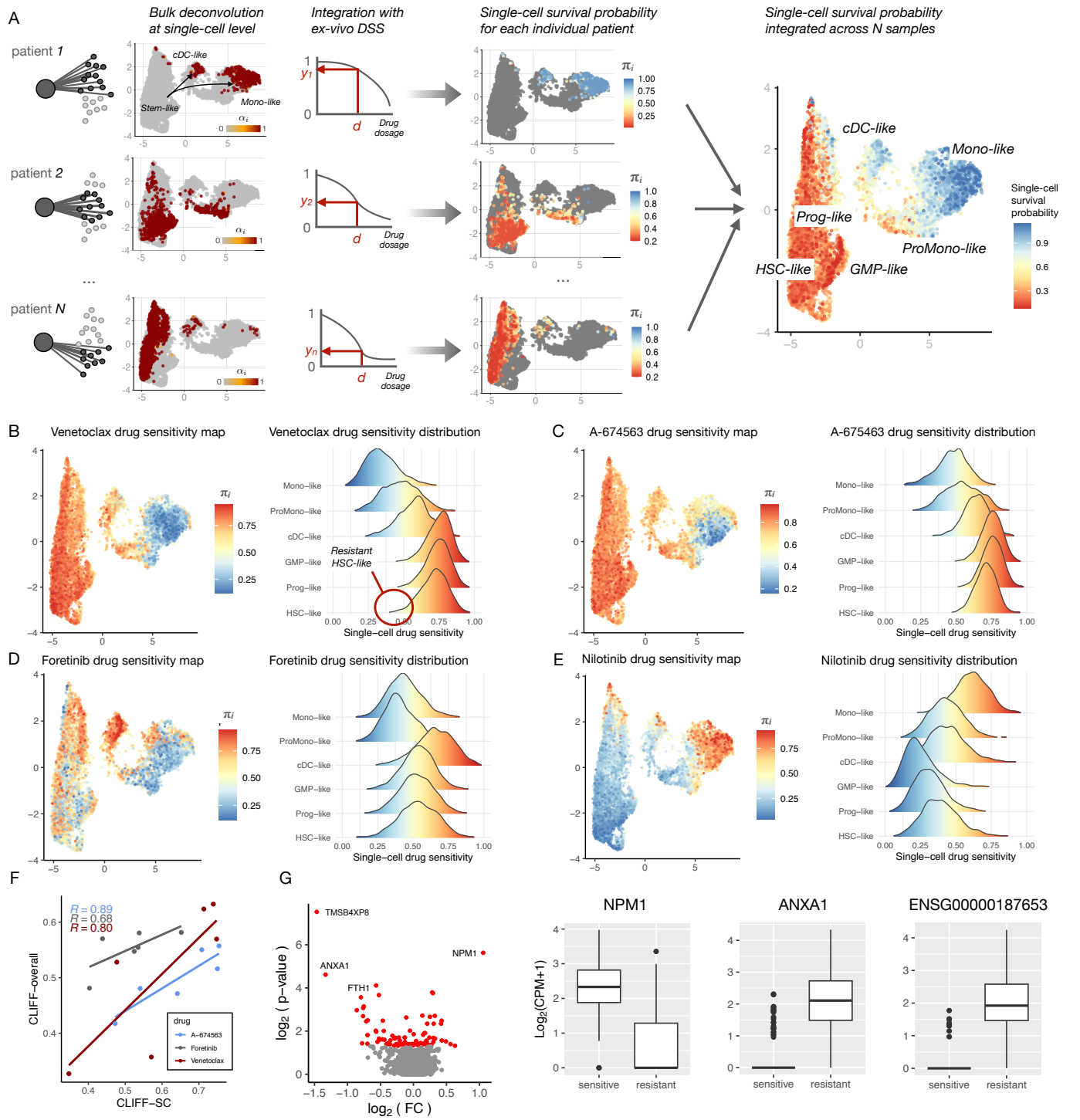


Fig. 6. Single-Cell Resolution Drug Sensitivity Prediction. (**A**) Framework overview: CLIMB deconvolution generates coefficients for each single cell (left), integrated with *ex-vivo* drug screening data (DSS) for single-cell survival probabilities (right). (**B**) Single-cell survival probability for venetoclax, based on UMAP projections of cancer cells from Van Galen et al. [24] (left), and survival probability distributions among six cancer cell subtypes (right). (**C**) Similar view for A-674563 single-cell drug sensitivity. (**D**) Similar view for foretinib single-cell drug sensitivity, highlighting upper tail of resistant HSC-like cells in red circle. (**E**) Similar view for nilotinib single-cell drug sensitivity. (**F**) Volcano plot showing gene expression comparison between venetoclax-resistant (25%) and sensitive cells (25%) with p-value versus log fold change. (**H**) Genes over-expressed in venetoclax-resistant population of HSC-like cells.

674563 and midostaurin had available validation from the literature (ref Tab). The eight undocumented drugs predicted to be the most effective against HSC-like cells were, in order, SNS-032, flavopiridol, foretinib, dovitinib, bortezomib, quizartinib, Go6976, and RAF265. In summary, the combined use of CLIMB and CLIFF successfully unveiled a comprehensive landscape of drug sensitivity at the cell subtype level in AML, along with the ability to predict drug sensitivity in the clinically relevant HSC-like cell subtype.

Single-Cell Resolution Prediction Identifies Resistant Cancer Cell Subpopulations

We further developed CLIFF-SC, which computes drug sensitivity prediction at single-cell resolution (Fig. 6A, Methods, Supplemental Methods), leveraging coefficients derived by CLIMB for each individual cell. We focused on three drugs predicted to target the HSC-like compartment: venetoclax, A-674563, and foretinib, and one drug with an opposite pattern, nilotinib, which targets monocyte-like cancer cells (Fig. 5A). We visualized single-cell survival probability on 2D projections (UMAP) using all cancer cells from Van Galen dataset [24], which showed considerable heterogeneity in distributions across the four drugs and the six cancer cell subtypes (Fig. 6B-E). We observed high concordance between average cell subtype sensitivity predicted by CLIFF-SC and the one predicted by CLIFF-overall (Fig. 6F). Interestingly, we observed a long upper tail of seemingly resistant HSC-like cells for venetoclax compound (Fig. 6B, right panel, circled in red). We thus decided to study in more detail this population of resistant HSC-like. We performed differential expression analysis between resistant and sensitive HSC-like populations, which were clustered using single-cell venetoclax sensitivity predictions (25% most sensitive versus 25% most resistant, Fig. 6G). We could show a downregulation of NPM1 expression and upregulation of ANXA1 and TMSB4XP8 in the 25% with highest predicted resistance to venetoclax (Fig. 6H). These data suggest that expression of these genes could be predictive of venetoclax sensitivity at single-cell level.

DISCUSSION

The beatAML consortium achieved the most extensive *ex-vivo* drug sensitivity screening (DSS) to date across all cancer types [7], paving the way for deciphering cell subtype-specific drug sensitivity. Our study proposes a method to address this challenge, potentially replacing resource-intensive investigations commonly required for assessing cell population-level drug sensitivity in large patient cohorts. We present a pioneering approach to deconvolute bulk drug sensitivity data into its constituent cell subtypes, employing a method capable of retrieving realistic AUC metrics for cell subtype-level drug sensitivity. Through extensive evaluation in cell line mixes and acute myeloid leukemia (AML), we predict drug sensitivity for 19 cell subtypes, 115 anti-cancer drugs, and 286 patients, culminating in a comprehensive dataset comprising over 344,318 data points derived from the beatAML dataset.

Bulk deconvolution is a cost-effective method to uncover cell composition in patient cohorts. To this end, we introduce CLIMB, a novel algorithm that outperforms current standards in various contexts, achieving accuracy comparable to deep learning-based methods. CLIMB's advantage lies in its use of the complete

single-cell expression matrix, unlike other methods that assume uniform cell subtype expression across bulk samples. This is crucial, as cancer cell expression variability among patients is expected. Based on or cross-dataset pseudo-bulk study, CLIMB effectively addresses batch effects. Notably, CLIMB excels in estimating proportions of HSC-like cells, a subtype linked to poor AML prognosis and relapse, enhancing predictions of their drug sensitivity. We anticipate CLIMB's utility across diverse biological scenarios, particularly in deciphering cancer sample cellularity.

By considering latent variables and employing an Expectation-Maximization algorithm, CLIFF deconvolves bulk-level cell viability into its sub-population-level components (Method, Supplemental Methods). We propose that the primary strength of the CLIFF method, when compared to prior methodologies [32–34], lies in its capacity to generate biologically interpretable metrics of cell viability at the cellular level. This stands in contrast to correlative techniques that merely provide a score linking cell subtypes and drugs without delivering a meaningful biological measure. We demonstrated CLIFF abilities first with engineered mixes of leukemic cell lines and pseudo-bulk analysis. We explored multiple avenues for extending CLIFF's 'overall' mode—which provides an averaged prediction across all bulk samples taken as input—with two higher resolution strategies. On one hand, CLIFF-highres, which uses cell subtype expression and mutation data as input, delivering cell viability predictions for each sample. In our benchmarking, CLIFF-highres generally showed decreased accuracy compared to CLIFF-overall, likely due to the limited number of input samples. In contrast, CLIFF-SC employs a full-sized single-cell dataset as input, substituting cell subtypes with single cells in CLIFF's original model, thereby allowing the prediction of single-cell drug sensitivity. It proved to be a straightforward method to enhance CLIFF's resolution, providing valuable single-cell maps of drug sensitivity. We then turned to the beatAML dataset, using CLIFF-overall model enhanced with mutation data as covariates. In this setting, we could reveal an extensive landscape of cell subtype drug sensitivity. Also, we predicted A-674563, Midostaurin, SNS-032, and foretinib, among other compounds, to be efficient at targeting HSC-like cells. We validated our findings by relying on 14 publications that documented cell-subtype level drug sensitivity through functionnal experiments in AML patient's samples, for 22 pairs of drug to cell subtype interactions [9, 11–13, 16–23], with showed overall good agreement with our predictions (72.1%).

Our approach has some limitations. First, since ground truth cell viability data is not available for patient samples, it is challenging to assess our deconvolution results on AML cohorts. This limitation led us to use a small compilation of literature as a comparison point. Second, any weakness in CLIMB's model could potentially affect CLIFF's predictive accuracy. We believe CLIFF would benefit significantly from being used in contexts where cellular proportions can be directly observed and measured rather than predicted through deconvolution. For example, using a cohort of patients with available scRNA-seq data and cell viability measures. Third, we rely on cell viability observed ex-vivo from bone marrow samples in 96-well cell viability plates, an environment which might bias the viability of fragile cell subtypes.

By integrating diverse data modalities, our combined approach seeks to enhance cellular composition predic-

tions and unravel drug sensitivity at the level of individual cell subpopulations. Here, we make a pioneering effort, as far as our knowledge extends, to advance drug sensitivity deconvolution to novel resolutions, encompassing personalized and single-cell-level predictions. As the integration of single-cell RNA-seq and ex-vivo drug sensitivity assays gains traction for diagnostic applications, the development of suitable computational methodologies becomes imperative to fully leverage the wealth of information from these joint data sources.

METHODS

Gene and cell selection

In our approach to gene subset selection prior to bulk deconvolution, we utilized the Scran::ModelGenVar function on the expression matrix derived from single-cell RNA sequencing (scRNA-seq) data. By applying this methodology, we identified genes with estimated biological variance surpassing technical variance, as indicated by positive biological variance outcomes from ModelGenVar. This strategy eliminated the reliance on arbitrary thresholds for the determination of variable or deferentially expressed genes and was consistently employed across all analyses in this study.

Regarding cell selection, in the context of our in-vitro experiment, we observed cells forming five clusters, with four clusters being distinctly attributed to the four mixed cell lines. The fifth cluster lacked a clear marker expression pattern, exhibited smaller size (only few cells) compared to the other clusters, and displayed elevated mitochondrial read content. Consequently, we excluded this ambiguous cluster from our analysis, while retaining the remaining cells in four clusters that were attributed their known cell line labels according to their marker expression pattern.

Regarding the re-analysis of publicly available scRNA-seq datasets, we harnessed the cell selection information provided in the processed data tables, abstaining from further cell filtering based on specific criteria. In instances where datasets encompassed over 30,000 cells, we subsampled them to reach a count of 30,000 cells, adopting the set.seed(1) option for reproducibility. This enabled us to execute the Seurat integration tool without memory constraints. The datasets subjected to subsampling included Naldini et al., Gray et al., Wu et al., and Khaliq et al.

Prototypical bulk deconvolution model

The goal of a bulk deconvolution model is to fit coefficients w_k corresponding to each cell-subtype k through a linear model. Fitted coefficients directly relates to cell-subtype proportions. Thus, bulk deconvolution solves the following problem :

$$\mathbf{x}_n \approx \mathbf{S} \cdot \mathbf{w}_n \quad \textbf{(1)}$$

with $\mathbf{x}_n$ representing a vector of gene expression of length G, for a given mixture n. The matrix $\mathbf{S}$ is a G by K matrix containing the average expression for each cell-subtype k and each gene g. To obtain proportions w_{nk} that sum up to 1, the model can either impose a constraint $\sum_k w_{nk} = 1$, or the coefficients w_{nk} can be normalized after fitting. In this model, the genes are our observations, and corresponds to genes shared between $\mathbf{x}_n$ and $\mathbf{S}$ or to a subset of marker genes. Note that every mixture n are resolved with an independent linear model, and that the average expression of cell-subtypes is usually derived from a scRNA-seq that does not contain mixture n. Thus, we assume that $\mathbf{S} \approx \mathbf{S}_n$. Of note, this simple model do not apply for advanced deconvolution algorithms like CiberSortX, BayesPrism, Scaden, BLADE, or TAPE.

CLIMB: CELL-LEVEL LINEAR MODEL THROUGH BULK-TO-SINGLE-CELL DECONVOLUTION

We developed the CLIMB method, which proposes an innovative approach to solve a bulk deconvolution problem, by applying a mapping from each bulk sample to each single-cell from a reference dataset. We thus fit the coefficients α_i for each single-cell i in an intial run of CLIMB, and thus we find the best linear combination of single-cell c_i to predict bulk expression $\mathbf{x}_n$. Thus, we use the relation $\mathbf{x}_n \sim \mathbf{C} \cdot \boldsymbol{\alpha}_n$, shown below in more details :

$$\begin{bmatrix} x_{n1} \\ \vdots \\ x_{nG} \end{bmatrix} \propto \underbrace{\begin{bmatrix} c_1^1 & \dots & c_1^C \\ \vdots & \ddots & \vdots \\ c_G^1 & \dots & c_G^C \end{bmatrix}}_{\mathbf{C}^{(G\times C)}} \cdot \begin{bmatrix} \alpha_n^1 \\ \vdots \\ \alpha_n^C \end{bmatrix} \tag{2}$$

where the vector $\mathbf{x}_n$ represent the bulk expression matrix for a mixture n, $\mathbf{C}$ is the raw counts matrix from scRNA-seq, and $\boldsymbol{\alpha}_n$ is the vector of coefficients fitted for each single-cell i for a given mixture n. Thus, CLIMB loss function is defined by :

$$L(\boldsymbol{\alpha}_n \mid \mathbf{C}, \mathbf{x}_n) = \min_{\boldsymbol{\alpha}} \frac{1}{2G} \left|\mathbf{x}_n - \mathbf{C} \cdot \boldsymbol{\alpha}_n\right|^2 \quad \text{, with } \alpha_{ni} \in [0, +\infty] \tag{3}$$

Thus, we impose a non-negativity constraint on the coefficients, and no regularization. We optimize α_i coefficients associated with each single-cell i with glmnet R library [52], with the options lower.limits = 0, lambda = 0, and scale = TRUE, intersect=FALSE. We indeed observed no improvements of the deconvolution with neither l_1 nor l_2 regularization.

We then group and normalize coefficients α_{ni} obtained with (3) to obtain the cell-subtype proportions w_{nk} for each cell-subtype k and mixture n :

$$w_{nk} = \frac{\sum_{i=1}^{C} \tau_{ki} \cdot \alpha_{ni}}{\sum_{i=1}^{C} \alpha_{ni}} \tag{4}$$

With $\boldsymbol{\tau}_k$ being an indicator variable of length C, equal to 1 whenever cell i is of type k.

Empirical Bayes Sub-Sampling Mitigates Bias Induced by Reference Matrix

The equations 2 3, 4 yield biased estimates of cellular proportions towards the cellular proportions contained in the reference single-cell matrix **C**. This bias acts as a prior, impacting CLIMB's predictions. To enhance cellular composition depiction, we'll correct **C** by informed cell sub-sampling, creating a new matrix via an Empirical Bayes approach estimating the prior from the data.

We use the initial CLIMB-derived cellular proportions from equations 2 3, 4 to sub-sample half of the cells in reference matrix C based on our predicted biased proportions. Each sub-sample ensures a minimum of 50 cells (or 2% of total cell count) per cell type. In total, five sub-sampling are done, and this process is repeated twice. Ultimately, the proportions from the second run become the final CLIMB output (stored in 'props.corrected'). More details in algorithm **??**.

Additionally, a gene selection aids CLIMB's deconvolution. First, dataset-specific top-variable genes are selected based on cell type average expression per gene. Genes with significant fold change compared to the average expression and compared to the second top-expressed cell type, are chosen (threshold set to 95% and 75% fold changes quantiles). Cell types with top expression are recorded, allowing a maximum of 100 genes per cell type. This initial gene selection is followed by bulk sample-specific gene selection: Genes are ranked according to their maximum cell type expression multiplied by their expression in the target bulk sample, and the top 500 genes are selected. These are thus bulk sample-specific. CLIMB's iterative algorithm is summarized in SFig.2.

Deconvolution of cell-subtype expression

CLIMB allows the prediction of cancer cell expression in a two step manner. Let first define our target *high-resolution* expression $\hat{\mathbf{S}}_{(ngk)}$, which is the cell-subtype-specific and bulk sample-specific expression for each cell-subtype k. We assume that for a given group of samples, the *grouped* expression $\hat{\mathbf{S}}_{(gk)}$ relates to the *high-resolution* expression with the following formula:

$$s_{gk}^{(grouped)} = \frac{1}{N}\sum_{n=1}^{N} s_{ngk}^{(high-res)} \tag{5}$$

First, we take advantage of the bulk to single-cell mapping to directly derive the cell-subtype expression from the bulk-to-single-cell deconvolution or *mapping*. We can simply sum up the expression of cells mapped by CLIMB to get a *high-resolution* cell-subtype-specific and bulk-specific expression $s_{ngk}^{(mapping)}$:

$$s_{ngk}^{(mapping)} = \sum_{i=1}^{C} \alpha_{in} \cdot c_{ig} \cdot \tau_{ik} \tag{6}$$

with α_{in} being the coefficient fitted by GLMNET on each single-cell, linking each cell i to each bulk n. c_{ig} denotes the single-cell expression of cell i and gene g, and τ_{ik} is an indicator vector equal to 1 if cell i is of type k.

Although this is potentially sufficient in cases where scRNA-seq reference dataset maps all cell states present in bulk samples, it can be insufficient to detect genes deferentially-expressed (DE) at cell-subtype level. As DE genes will likely induce error on the cell-subtype expression deconvolution, we use the error itself to fit our model. We can also make the reasonable assumption that DE genes only originates in cancer cells. Thus, for a given gene g and a given bulk n, we have:

$$y_{ng} \sim \hat{y}_{ng}^{(mapping)} + \epsilon_{ng} \quad \rightarrow \quad y_{ng} - \hat{y}_{ng}^{(mapping)} = \epsilon_{ng} \tag{7}$$

Thus, we use $\epsilon_{ng} = y_{ng} - \hat{y}_{ng}^{(mapping)}$ as input for a model to predict cancer cell-subtype expression. We decompose our error ϵ_{ng} between the error originated from DE genes $\epsilon_{ng}^{(DE)}$, and an additional error term ϵ_{ng}^{*}. Similar as other methods, we use the information from all bulk samples N to deconvolute the high-resolution cell-subtype specific expression $\hat{s}_{ngk}^{(DE)}$. We consider that $\epsilon_{ng}^{(DE)}$ only originates from cancer cells $C^{(C)}$ of type $K^{(C)}$, thus

$$\epsilon_{ng} \sim \underbrace{\sum_{k=1}^{K^{(C)}} \left(w_{nk} \cdot \hat{s}_{ngk}^{(DE)} \right)}_{\epsilon_{ng}^{(DE)}} + \epsilon_{ng}^{*} = \sum_{i=1}^{C^{(C)}} \left(\alpha_{in} \cdot \hat{c}_{ig} \right) + \epsilon_{ng}^{*} = \mathbf{A} \cdot \hat{\mathbf{c}}_{g} + \epsilon_{ng}^{*} \tag{8}$$

Where $\hat{\mathbf{c}}_{g}$ is the vector of coefficients we wish to predict. We used Glmnet with options lower.limits = 0, lambda = 0, and scale = TRUE, fitting the following loss function:

Our final prediction of cell-subtype expression for a cancer cell-subtype k, a given bulk n, and a given gene g is:

$$\hat{s}_{ngk}^{(high-res)} = \underbrace{\hat{s}_{ngk}^{(mapping)}}_{\text{from scRNA-seq}} + \underbrace{\hat{s}_{ngk}^{(DE)}}_{\text{learned}} = \sum_{i=1}^{C} \alpha_{in} \cdot (c_{ig} + \hat{c}_{ig}) \cdot \tau_{ik} \tag{9}$$

Following the assumptions made above, we set $\hat{c}_{ig}$ to zero for normal cells, which imply that $\hat{s}_{ngk}^{(high-res)} = \hat{s}_{ngk}^{(mapping)}$ for normal cell subtypes.

CLIFF - Cell fate inference for drug sensitivity deconvolution

Here we propose to infer the drug sensitivity of individual cell subtypes using only drug sensitivity data observed on mixed cell populations, considering K different cell populations in N mixed samples, with pro-

portions than can be defined as $\mathbf{w}_n = [w_{n1}, ..., w_{nK}]$, with $w_{nk} \in [0,1]$ and $\sum_{k=1}^{K} w_{nk} = 1$. Drug sensitivity observed in the bulk samples we wish to demix are represented by a vector $\mathbf{y}^{(N)}$, with $y_n \in [0,1]$ being the empirical surviving rate of cells, and for sample n. Therefore the target matrix we wish to predict is $\hat{\mathbf{Y}}^{N\times D}$, and the predicted values of drug sensitivity is assumed to depend on the underlying drug sensitivity of cell subtypes $\pi_{nk} \in [0,1]$ for patient n, and cell-subtype k, represented as a surviving rate of cells.

To model drug sensitivity, we assume that single-cell survival is represented as a binary output $\delta_{ni} \in \{0,1\}$ that follows a Bernoulli distribution. Thus, every single-cell i in patient n exposed to a drug can either survive ($\delta_{ni} = 1$) or die ($\delta_{ni} = 0$). Thus a cell i has a probability $P(\delta_{ni} = 1)$ to survive and $1 - P(\delta_{ni} = 1)$ to die. Let's say we are interested to know the survival probability $P(\delta_{ni} = 1)$, given a drug dose. We could resolve the log-likelihood of the Bernoulli distribution for all the cells of a single sample n :

$$\mathcal{L}_n = \sum_{i=1}^{C_n} \left[\delta_{ni} \cdot \log P(\delta_{ni} = 1) + (1 - \delta_{ni}) \cdot \log(1 - P(\delta_{ni} = 1))\right] \tag{10}$$

It turns out that in our problem, δ_{ni} cannot be directly observed. We are indeed only able to observe the overall ratio of cells that survives among all cells C_n for a given patient n, which we define as y_n. It relates to the values δ_{ni} with :

$$y_n = \frac{1}{C_n} \sum_{i=1}^{C_n} \delta_{ni} \tag{11}$$

Thus, by multiplying (10) by $1/C_n$, if we assume that every cell i share the same average probability $P(\delta_{nid} = 1)$, we can re-write (10) and obtain the following log-likelihood :

$$y_n \cdot \log P(\delta_{ni} = 1) + (1 - y_n) \cdot \log(1 - P(\delta_{ni} = 1)) \tag{12}$$

As stated above, the core of the problem is to assume different survival probability for our single-cells. To slightly simplify this, we can consider a cell-subtype-specific survival probability. Thus, we consider the survival probability observed in our mixed population of cells $P(\delta_{nid} = 1)$, as a mixture of K probabilities $P(\delta_{nid}^{(k)} = 1)$, one for each cell subtype. Let introduce π_{kn}, a variable describing the survival probability of cell-subtype k:

$$\pi_{kd} = P(\delta_{ni}^{(k)} = 1) \tag{13}$$

We also define ν_{nd}, which corresponds to the expected number of cells surviving in patient n with drug dose d, and assume it follows a Binomial distribution, with y_{nd} being the ratio of surviving cells among cells in C_n. We further define $\hat{y}_{nd}$ as an average of cell-subtype survival probabilities π_{kdn} constituting the bulk, weighted with cell-subtype proportions w_{nk}:

$$\nu_n = C_n \cdot \underbrace{\sum_{k=1}^{K} w_{nk} \cdot \pi_{kn}}_{\hat{y}_n} \tag{14}$$

with w_{nk} the cell-subtype proportions. $\hat{y}_n$ is the predicted cell surviving rate in the sample n, and thus the target we wish to predict through a cross-validation procedure. We can say that if we let :

$$\hat{y}_n = \sum_k w_{nk} \cdot \pi_{kn} \quad \text{we have that} \quad \hat{y}_n \sim \text{Ber}(\hat{y}_n). \tag{15}$$

Additionally, we impose the cell-subtype-specific survival probabilities π_{kdn} to be non-linearly related with cell-subtype gene expression $\mathbf{s}_{kn}$ (a vector of length G) through a sigmoid function:

$$\pi_{kn} = \sigma(\beta^T \mathbf{s}_{kn} + \Theta_{mn}) \tag{16}$$

Note that the sigmoid in (16) also optionaly depends on Θ_{mn}, which is an intersect associated with each mutation m in each patient n, attributed to the value 1 in each cancer cell type and to 0 in normal cell types. To clarify the notation, we set that $\tilde{\boldsymbol{\beta}} = \{\boldsymbol{\beta}, \boldsymbol{\theta}\}$, with $\boldsymbol{\theta}$ being a vector containing coefficients associated with each mutation m.

When running CLIFF in *overall* mode, there are two assumptions that we use to simplify the problem. First, we assume that every sample n share the same cell-subtype expression s_{gk}, for a given gene g and cell-subtype k. Thus, $s_{gk} \approx \mathbb{E}_N[s_{gkn}]$. Second, we assume that cells of the same cell-subtypes across the N samples share the same sensitivities π_k. We thus assume that π_k correspond to the expected drug sensitivity over N samples, and therefore that $\pi_k \approx \mathbb{E}_N[\pi_{kn}]$.

Thus, the overall log-likelihood for our problem is :

$$\mathcal{L}(\tilde{\boldsymbol{\beta}} \mid \mathbf{Y}, \mathbf{W}, \mathbf{S}) = \max_{\tilde{\boldsymbol{\beta}}} \sum_{n=1}^{N} \sum_{d=1}^{D} \left[y_{nd} \log \left(\sum_{k=1}^{K} w_{nk} \cdot \pi_{kn}(\tilde{\boldsymbol{\beta}}) \right) + (1 - y_{nd}) \log \left(1 - \sum_{k=1}^{K} w_{nk} \cdot \pi_{kn}(\tilde{\boldsymbol{\beta}}) \right) \right] \tag{17}$$

As this log-likelihood expression contains the log of a sum, it is difficult to optimize with respect to $\tilde{\boldsymbol{\beta}}$. Therefore, we use an Expectation-Maximization to iteratively define lower bound functions that we optimize with respect to $\tilde{\boldsymbol{\beta}}$, which allows solving (17).

Expectation-Maximization algorithm to solve CLIFF

By using the concavity of the log through Jensen inequality, we can define a lower bound of (17) (more details on this in the Supplemental Methods) :

$$\begin{aligned}\mathcal{L}_B = &\sum_{n=1}^{N}\sum_{k=1}^{K} q_{nk}^{(1)} \cdot y_n \cdot \log(\pi_{kn}(\tilde{\beta})) + q_{nk}^{(0)} \cdot (1-y_n) \cdot \log(1-\pi_{kn}(\tilde{\beta})) \\ &- \sum_{n=1}^{N}\left[y_n \cdot KL(\mathbf{q}_n^{(1)} \| \mathbf{w}_n) + (1-y_n) \cdot KL(\mathbf{q}_n^{(0)} \| \mathbf{w}_n)\right]\end{aligned} \tag{18}$$

In order to solve (17), we will iterate over the Expectation step or E-step where we update $q_{nk}^{(0)}$ and $q_{nk}^{(1)}$ in (18), and the Maximization step or M-step, where we maximize (18) with respect to $\tilde{\beta}$. We can therefore define our Maximization algorithm as follow :

E-step

During E-step, we define our lower bound function at step t, $\mathcal{L}_B^{(t)}$ by defining $q_{nk}^{(0),(t)}$ and $q_{nk}^{(1),(t)}$ using $\pi_{kn}^{(t-1)}$ defined at previous step :

$$\begin{aligned} q_{nk}^{(1),(t)} &= \frac{w_{nk} \cdot \pi_{kn}^{(t-1)}}{\sum_{k=1}^{K} w_{nk} \cdot \pi_{kn}^{(t-1)}} \\ q_{nk}^{(0),(t)} &= \frac{w_{nk} \cdot (1-\pi_{kn}^{(t-1)})}{\sum_{k=1}^{K} w_{nk} \cdot (1-\pi_{kn}^{(t-1)})} \end{aligned} \tag{19}$$

This gives us the lower bound function of (18), tight to the original log-likelihood (17) when $q_{nk}^{(0),(t)}$ and $q_{nk}^{(1),(t)}$ satisfies (19). For a precise description of how $q_{nk}^{(0),(t)}$ and $q_{nk}^{(1),(t)}$ were defined, see Supplemental Methods.

M-step

During M-step, we optimize (18) with respect to $\tilde{\beta}$, considering the variables $\mathbf{q}^{(0),(t)}$ and $\mathbf{q}^{(0),(t)}$ defined during E-step as constants. Also, the two last term of (18) do not depend on $\tilde{\beta}$, thus we can solve the following problem which will produce the same optimal $\tilde{\beta}$ coefficients than if solving (18) :

$$\mathcal{L}_B = \max_{\tilde{\beta}} \sum_{n=1}^{N}\sum_{k=1}^{K} q_{nk}^{(1)} \cdot y_n \cdot \log(\pi_{kn}(\tilde{\beta})) + q_{nk}^{(0)} \cdot (1-y_n) \cdot \log(1-\pi_{kn}(\tilde{\beta})) \tag{20}$$

Interestingly, (20) can be solved with standard libraries using a weighted logistic regression model, where we define weights as being $\gamma_k^{(1)} = \sum_{n=1}^{N} q_{nk}^{(1)} \cdot y_n$ and $\gamma_k^{(0)} = \sum_{n=1}^{N} q_{nk}^{(0)} \cdot (1-y_n)$.

We optimize $\tilde{\beta}$ coefficients through a ten-fold cross-validation (taking 90% of the samples N in the training set, and 10% in the testing set), and using early-stopping to terminate the EM algorithm. Optimized $\tilde{\beta}$ coefficients

allow to compute optimized values for π_{kn}, our estimate of cell-subtype drug sensitivity.

CLIMB-SC - bulk deconvolution at single-cell level

The bulk deconvolution performed by CLIMB has the particularity of estimating coefficient directly on single-cells from a reference scRNA-seq dataset. Instead of transforming these coefficients into cell-subtype proportions as in Methods (3), we can compute normalized coefficients $\tilde{\alpha}_{ni}$ with the following formula :

$$\tilde{\alpha}_{ni} = \frac{\alpha_{ni}}{\sum_{i=1}^{C} \alpha_{ni}} \tag{21}$$

With α_{ni} being the coefficients fitted with GLMNET library. We therefore call this *single-cell bulk deconvolution*, which links every bulk RNA-seq n to every single-cells i from a scRNA-seq dataset, through what we will call *single-cell proportions* $\tilde{\alpha}_{ni}$. These *single-cell proportions* are thus normalized coefficients fitted by CLIMB, as shown above in (21). Thus we obtain that $\sum_{i=1}^{C} \tilde{\alpha}_{ni} = 1$. As $\tilde{\alpha}_{ni}$ sum up to 1, they can be used by CLIFF as if they were cell-subtype proportions.

CLIFF-SC for drug sensitivity inference at single-cell level

By slightly modifying CLIFF method and by linking it with the ouput of CLIMB-SC, we propose CLIFF-SC, a method providing drug sensitivity estimates at single-cell resolution. If the input dataset is composed of paired bulk RNA-seq and drug sensitivity screening data, a so-called *bulk deconvolution at single-cell level* is performed using a reference scRNA-seq dataset (cf. CLIMB-SC above).

Thus, as in (15), we model the observed ratio of cells y_{nd} that survives in a mixed population of cells as an average of single-cell survival probabilities π_{id}, associated with the bulk sample n, weighted by the coefficients $\tilde{\alpha}_{ni}$:

$$\hat{y}_{nd} = \sum_{i=1}^{C} \tilde{\alpha}_{ni} \cdot \pi_{id} \tag{22}$$

With C being the number of cells contained in a scRNA-seq reference datasets. Thus, we model the single-cell drug sensitivity π_{id}, which is also modeled with a sigmoid function, similar to (16) :

$$\pi_{id} = \sigma(\beta^T \mathbf{c}_i + \theta_d) \tag{23}$$

Where $\mathbf{c}_i$ is the ith row of the single-cell expression matrix $\mathbf{C}$, thus $\mathbf{c}_i$ is a vector of length G, representing the gene expression of cell i. Similarly as in the cell-subtype-centered model, we can derive a log-likelihood function :

$$\mathcal{L}(\tilde{\boldsymbol{\beta}} \mid \mathbf{Y}, \mathbf{W}, \mathbf{C}) = \max_{\tilde{\boldsymbol{\beta}}} \sum_{n=1}^{N} \sum_{d=1}^{D} \left[y_{nd} \log \left(\sum_{i=1}^{C} \tilde{\alpha}_{ni} \cdot \pi_{id}(\tilde{\boldsymbol{\beta}}) \right) + (1 - y_{nd}) \log \left(1 - \sum_{i=1}^{C} \tilde{\alpha}_{ni} \cdot \pi_{id}(\tilde{\boldsymbol{\beta}}) \right) \right] \tag{24}$$

A lower bound can be derived from (24), which gives :

$$\begin{aligned} \mathcal{L}_B = & \sum_{d,n}^{D,N} \sum_{i=1}^{C} q_{ndi}^{(1)} \cdot y_{nd} \cdot \log(\pi_{id}(\tilde{\boldsymbol{\beta}})) + q_{ndi}^{(0)} \cdot (1 - y_{nd}) \cdot \log(1 - \pi_{id}(\tilde{\boldsymbol{\beta}})) \\ & - \sum_{d,n}^{D,N} \left[y_{nd} \cdot KL(\mathbf{q}_{nd}^{(1)} \| \mathbf{w}_n) + (1 - y_{nd}) \cdot KL(\mathbf{q}_{nd}^{(0)} \| \mathbf{w}_n) \right] \end{aligned} \tag{25}$$

Then, the Expectation-Maximization algorithm can initialized with all coefficients $\tilde{\boldsymbol{\beta}}$ at 0. This implies that $\pi_{id}^{(t=0)} = 0.5$ for all i, d. During E-step, we compute values for $q_{ndi}^{(0)}$ and $q_{ndi}^{(1)}$ variables :

$$\begin{aligned} q_{ndi}^{(1),(t)} &= \frac{\tilde{\alpha}_{ni} \cdot \pi_{id}^{(t-1)}}{\sum_{i=1}^{C} \tilde{\alpha}_{ni} \cdot \pi_{di}^{(t-1)}} \\ q_{ndi}^{(0),(t)} &= \frac{\tilde{\alpha}_{ni} \cdot (1 - \pi_{id}^{(t-1)})}{\sum_{i=1}^{C} \tilde{\alpha}_{ni} \cdot (1 - \pi_{id}^{(t-1)})} \end{aligned} \tag{26}$$

Then, we optimize $\tilde{\boldsymbol{\beta}}$ coefficients at each step (t), during the so-called M-step :

$$\mathcal{L}_B = \max_{\beta,\theta} \sum_{d,n}^{D,N} \sum_{i=1}^{C} q_{ndi}^{(1)} \cdot y_{nd} \cdot \log(\pi_{id}(\tilde{\boldsymbol{\beta}})) + q_{ndi}^{(0)} \cdot (1 - y_{nd}) \cdot \log(1 - \pi_{id}(\tilde{\boldsymbol{\beta}})) \tag{27}$$

And we run the EM algorithm in a similar way than shown above for CLIFF.

Simulation of pseudo-bulk samples from scRNA-seq

We generated pseudo-bulk samples by summing up single-cell expression vectors per gene to obtain an artificially generated "pseudo"-bulk RNA-seq. As the precise amount of cells from each cell subtype added to the mix is known, we thus precisely know the cell-subtype proportions of the mixture and can benchmark deconvolution algorithms efficiently. We generated two datasets of pseudo-bulk in this work, both based on single-cells found in the Van Galen et al. dataset.

First, we used all single cells from the same sample (thus, the same patient), which we summed up to obtain a patient-specific pseudo-bulk RNA-seq sample. Thus, we define a variable ϕ_{ni} as a binary output variable indicating if a cell i belongs to patient n. Thus, a pseudo-bulk samples is a vector of length G containing the gene expression summed over all cells from patient n :

$$\tilde{\mathbf{x}}_n = \sum_{i=1}^{C} \phi_{ni} \cdot \mathbf{c}_i \tag{28}$$

where $\mathbf{c}_i$ is a vector of length G representing the gene expression of cell i. We can therefore easily get the cell-subtype proportion w_{nk}, by simply counting the number of cells from each cell subtypes and dividing by the total number of cells in patient n. For this we introduce an indicator vector $\boldsymbol{\tau}_k$, which is of length C and equal to 1 whenever cell i is of type k.

$$w_{nk} = \frac{\sum_{i=1}^{C} \phi_{ni} \cdot \tau_{ik}}{\sum_{i=1}^{C} \phi_{ni}} \tag{29}$$

Van Galen et al. dataset offers only a limited number of patients (16 AML patients at day 0 pre-treatment). To simulate an *in-silico* a dataset of similar size to the beatAML cohort, we generated 200 pseudo-bulk samples from the Van Galen et al. dataset allowing a random mix of cells from multiple patients. First, we generated 200 random cell-subtype proportions using a Dirichlet distribution. Thus, we assume that cell-subtype proportions $\mathbf{w}_n \sim \mathrm{Dir}(\boldsymbol{\alpha})$, with $\boldsymbol{\alpha}$ being a vector of length K, containing one value for each cell subtype k. We fitted $\boldsymbol{\alpha}$ using cell-subtype proportions in Van Galen et al. dataset with the R library *feralaes/dampack*. Then, we generated cell-subtype proportions for 200 pseudo-bulks, with the R library *DirichletReg* and the $\boldsymbol{\alpha}$ previously fitted. Cells from Van Galen et al. dataset were selected with probabilities corresponding to cell-subtype proportions. To use a realistic number of cells to generate our pseudo-bulk samples, we used a gamma distribution fitted on Van Galen et al. dataset using *fitdistr* function from R library *MASS*. In the end, we generated a matrix of pseudo-bulk expression $\bar{\mathbf{X}}$ of size $(N \times G)$ with N = 200, together with associated true cell-subtype proportions $\bar{\mathbf{W}}$ of size $(N \times K)$.

Simulation of drug sensitivity screening data

To simulate drug screening data, we assumed that the drug sensitivity observed on a mix of cells corresponds to the sum of the cell-subtype drug sensitivity constituting its bulk weighted by their respective proportions. Additionally, we added Gaussian noise with a signal-to-noise ratio of 100. Thus, the simulated survival rate for pseudo-bulk n, in cell-subtype k exposed to drug dose d was generated with the following formula, given the proportions :

$$\bar{\pi}_{kd} = \frac{1}{1 + \exp(\zeta_f \cdot d + c)} \tag{30}$$

with ζ_f being a randomly generated number, with $\zeta_f \in [1, 2]$ if f = 1 (low sensibility cell-subtypes - resisting cells) and $\zeta_f \in [5, 6]$ if f = 2 (high sensibility cell-subtypes). Here we did not use the drug dose d, only the integer number related to drug dose $d = \{0, 1, ..., 6\}$, as we considered seven drug doses. We arbitrarily defined three cell subtypes with high sensitivity and the rest having low sensitivity. The constant c added was set to c = –8 as it provided visually convincing results. Thus $\bar{\pi}_{kd}$ follows a sigmoid function that depends on drug

dose d but in a way that is slightly different than in our model. Then, we generated the survival rate at the bulk level using the same formula as in our model but adding gaussian noise to it :

$$\bar{y}_{nd} = \sum_{k=1}^{K} \bar{w}_{nk} \cdot \bar{\pi}_{kd} + \epsilon \quad \text{with} \quad \epsilon \in \mathcal{N}(0, \sigma^2) \tag{31}$$

and σ was chosen to generate a signal-to-noise ratio arbitrarily chosen at 100.

Area under the curve for summarizing drug sensitivity of multiple data points

As CLIFF produces D predictions for each cell subtype, and CLIFF-SC produces D predictions for every single cell, it is desired to summarize these D data points into a single value. The goal is to summarize cell-subtype or single-cell drug sensitivity. To summarize drug sensitivity from D different drug doses, we normalize drug doses so that the minimum dose = 0 and the maximum dose = 1. As the cell survival rate is $\in [0, 1]$, we can compute an area under the curve (AUC) which will be $\in [0, 1]$. Conversely, we get an AUC of 1.0 if all cell survives and an AUC of 0.0 if all cells die, taking every dose into account. We computed the AUC using the function *computeAUC* from the R package *PharmacoGx*.

Culture conditions of cancer cell lines

Human cancer cell lines (K562, HL60, THP1, SUDHL4) were purchased from the American Type Culture Collection (ATCC; Rockville, MD, USA) and cultured at 37°C in a humidified 5% CO2 environment in RPMI 1640 (HCC38) supplemented with 2mM L-glutamine, 2 g/L D-glucose, and 10% fetal bovine serum (FBS; Gibco ThermoFisher).

Pharmaceutical compounds

Stock solutions for every agent (Table2) were prepared using DMSO (Sigma-Aldrich; stored at 80°C), further diluted in H2O to the appropriate concentration, plated in Greiner 96-well flat bottom (REF) and the plates were stored at 20°C. The pharmaceutical compounds were screened at eight concentrations (Log concentration M = -9.205031395, -8.699881554, -8.130127907, -7.70415926, -7.209309101, -6.70415926, -6.125806332, -5.699837685, -5.204987526, -4.699837685) using a Log2-fold dilution series with matched DMSO concentration vehicle controls.

Bulk proportion for bulk RNAseq and drug sensitivity

Six bulks were created by mixing the four cell lines with the following proportions. Pure and mixed cell populations were subjected to RNA extraction and drug sensitivity immediately after mixing.

Cell viability assay

Cells were plated in the 96-well plates containing chemical compounds previously equilibrated at 37C, at a density of 1×104 cells per well in 100µl culture medium. Cell viability was assessed after drug treatment for 24, 48, 72 or 96 in duplicates hours by incubating the cells with PrestoBlue™ Cell Viability Reagent, (Thermo

CatalogID	Name	Supplier
A3656	5-Aza-2′-deoxycytidine	Sigma
C3350000	Cytarabine	Sigma
CDS022173	Imatinib	Sigma
CDS023093	Nilotinib	Sigma
HY-15142	Doxorubicin (hydrochloride)	MedChemExpress
O9512	Oxaliplatin	Sigma
R0395	Rapamycin	Sigma
Y0000378	Mitomycin	Sigma

Table 1. Table of eight coumpounds used in the *in-vitro* drug sensitivity experiment.

Cell line	Mixture1	Mixture2	Mixture3	Mixture4	Mixture5	Mixture6
SUDHL4	0.33	0.07	0.00	0.61	0.188	0.20
K562	0.09	0.11	0.222	0.04	0.317	0.26
HL60	0.26	0.57	0.333	0.05	0.0198	0.12
THP1	0.32	0.25	0.444	0.30	0.475	0.42

Table 2. Table of cell line proportions added in tubes manually to generate the mixtures of cell lines.

Fisher Scientific - Life Technologies A13261). PrestoBlue was at room temperature before adding to the cells in each well with. The cells were incubated with 10µl PrestoBlue (10% of cell culture volume) for two hours at 37°C. The fluorescence was measured by the bottom with a 560nm excitation filter and a 590nm emission filter in a BioTek Neo HTS microplate. The acquisitions were performed with the following parameters: Gain: 54, Light Source Xenon Flash, Lamp Energy: High, Read Speed: Normal, Delay 0 msec, Mesurements/Data Point: 10, Read Height 5.25mm.

RNA extraction and sequencing

RNA-seq Total RNA from cell lines was isolated with NucleoSpin™ RNA Plus kit (Machery-Nagel). cDNA was prepared with Maxima Reverse Transcriptase (Thermo Scientific). Sequencing libraries were performed with Illumina Truseq Stranded mRNA LT kit. Reads were mapped to the human (hg19) genome using Hisat2. Counts on genes were generated using featureCounts. Normalization for sequencing depth has been done using the counts on genes as library size using the TMM method as implemented in the limma package of Bioconductor.

Single Cell RNAseq (MULTIome-Seq)

Single cell RNAseq was performed using 10x Genomics Chromium Single Cell Controller following the 10x Genomics Multiome (scRNA/scATAC-seq) demonstrated protocol: Chromium Next GEM Single Cell Multiome ATAC + Gene Expression User Guide (CG000338). To note, experimental parameters specific to the experiment were 3,000 for the Targeted Nuclei Recovery and 5 minutes for the lysis time.

Single-cell RNA-seq analysis

Cellranger-arc was used to obtain counts on genes using default parameters. Counts were obtained using cellranger using hg19. Only uniquely mapped reads on genes that were expressed in at least 1% of the samples were kept. Seurat's SCTransform was used to normalize the data and correct for mitochondrial percentage and total number of reads biases.

DATA AVAILABILITY

The data produced for the in-vitro experiment shown in Fig. 2 and Fig. 4 can be found on a Gene Expression Omnibus repository (GEO): GSE217924 (superserie). The bulk RNA-seq of the mixtures is available on the GSE217919 repository and the single-cell RNA-seq data on the GSE217922 repository. The rest of the study was relying on re-analysis of existing data available at the following GEO locations: Van Galen et al. (GSE116256), Naldini et al. (GSE185834), Lee et al. (GSE144735), Khaliq et al. (GSE200997), Gray et al. (GSE180878), Wu et al. (GSE176078), Jerby et al. (GSE115978), Tirosh et al. (GSE115978), Neftel et al. (SmartSeq2 and 10X data in GSE131928). BeatAML dataset is available via dbGaP study ID 30641 and accession ID phs001657.v1.p1.

CODE AVAILABILITY

The code is freely accessible on GitHub. CLIMB and CLIFF code can be accessed via an R library, 'ClimbTheCliff', available via the devtools command devtools::install_github('alexdray86/ClimbTheCliff'). Documentation are available on https://alexdray86.github.io/CLIMB/build/index.html and https://alexdray86.github.io/CLIFF/build/index.html website.

ACKNOWLEDGMENTS

This work was supported by grants from the Swiss National Science Foundation and the European Research Council (KRABnKAP, no. 268721; Transpos-X, no. 694658), the Personalized Health and Related Technologies (PHRT) and Swiss Data Science Center (SDSC) strategic focus areas of the Swiss Federal Institutes of Technology, the Swiss Personalized Health Network (SPHN) initiative of the Swiss Academy of Medical Sciences (project no. 2017-407), The Swiss Cancer League and the Aclon Foundation to D.T. We also extend our gratitude to Daria Korotkova, Cyril Pulver, and Paola Malsot, for their valuable scientific discussions. Additionally, I acknowledge my wonderful son, Matisse Coudray, for the inspiration he has provided.

REFERENCES

1. L. Kohtamäki, M. Arjama, S. Mäkelä, P. Ianevski, K. Välimäki, S. Juteau, S. Ilmonen, D. Ungureanu, O. Kallioniemi, A. Murumägi, and M. Hernberg, "High-throughput ex vivo drug testing identifies potential drugs and drug combinations for NRAS-positive malignant melanoma," Transl Oncol **15**, 101290 (2022).
2. S. de Mel, M. B. M. Rashid, X. Y. Zhang, J. Goh, C. T. Lee, L. M. Poon, E. H. L. Chan, X. Liu, W. J. Chng, Y. L. Chee, J. Lee, Y. C. Yuen, J. Q. Lim, B. K. H. Chia, Y. Laurensia, D. Huang, W. L. Pang, D. M. Z. Cheah, E. K. Y. Wong, C. K. Ong, T. Tang, S. T. Lim, S. B. Ng, S. Y. Tan, H. Y. Loi, L. K. Tan, E. K. Chow, and A. D. Jeyasekharan, "Application of an ex-vivo drug sensitivity platform towards achieving complete remission in a refractory T-cell lymphoma," Blood Cancer J **10**, 9 (2020).
3. M. A. Spinner, A. Aleshin, M. T. Santaguida, S. A. Schaffert, J. L. Zehnder, A. S. Patterson, C. Gekas, D. Heiser, and P. L. Greenberg, "Ex vivo drug screening defines novel drug sensitivity patterns for informing personalized therapy in myeloid neoplasms," Blood Adv **4**, 2768–2778 (2020).
4. P. Hu, W. Zhang, H. Xin, and G. Deng, "Single Cell Isolation and Analysis," Front Cell Dev Biol **4**, 116 (2016).
5. J. N. Weinstein, E. A. Collisson, G. B. Mills, K. R. Shaw, B. A. Ozenberger, K. Ellrott, I. Shmulevich, C. Sander, J. M. Stuart, K. Chang, C. J. Creighton, C. Davis, L. Donehower, J. Drummond, D. Wheeler, A. Ally, M. Balasundaram, I. Birol, S. N. Butterfield, A. Chu, E. Chuah, H. J. Chun, N. Dhalla, R. Guin, M. Hirst, C. Hirst, R. A. Holt, S. J. Jones, D. Lee, H. I. Li, M. A. Marra, M. Mayo, R. A. Moore, A. J. Mungall, A. G. Robertson, J. E. Schein, P. Sipahimalani, A. Tam, N. Thiessen, R. J. Varhol, R. Beroukhim, A. S. Bhatt, A. N. Brooks, A. D. Cherniack, S. S. Freeman, S. B. Gabriel, E. Helman, J. Jung, M. Meyerson, A. I. Ojesina, C. S. Pedamallu, G. Saksena, S. E. Schumacher, B. Tabak, T. Zack, E. S. Lander, C. A. Bristow, A. Hadjipanayis, P. Haseley, R. Kucherlapati, S. Lee, E. Lee, L. J. Luquette, H. S. Mahadeshwar, A. Pantazi, M. Parfenov, P. J. Park, A. Protopopov, X. Ren, N. Santoso, J. Seidman, S. Seth, X. Song, J. Tang, R. Xi, A. W. Xu, L. Yang, D. Zeng, J. T. Auman, S. Balu, E. Buda, C. Fan, K. A. Hoadley, C. D. Jones, S. Meng, P. A. Mieczkowski, J. S. Parker, C. M. Perou, J. Roach, Y. Shi, G. O. Silva, D. Tan, U. Veluvolu, S. Waring, M. D. Wilkerson, J. Wu, W. Zhao, T. Bodenheimer, D. N. Hayes, A. P. Hoyle, S. R. Jeffreys, L. E. Mose, J. V. Simons, M. G. Soloway, S. B. Baylin, B. P. Berman, M. S. Bootwalla, L. Danilova, J. G. Herman, T. Hinoue, P. W. Laird, S. K. Rhie, H. Shen, T. Triche, D. J. Weisenberger, S. L. Carter, K. Cibulskis, L. Chin, J. Zhang, G. Getz, C. Sougnez, M. Wang, G. Saksena, S. L. Carter, K. Cibulskis, L. Chin, J. Zhang, G. Getz, H. Dinh, H. V. Doddapaneni, R. Gibbs, P. Gunaratne, Y. Han, D. Kalra, C. Kovar, L. Lewis, M. Morgan, D. Morton, D. Muzny, J. Reid, L. Xi, J. Cho, D. DiCara, S. Frazer, N. Gehlenborg, D. I. Heiman, J. Kim, M. S. Lawrence, P. Lin, Y. Liu, M. S. Noble, P. Stojanov, D. Voet, H. Zhang, L. Zou, C. Stewart, B. Bernard, R. Bressler, A. Eakin, L. Iype, T. Knijnenburg, R. Kramer, R. Kreisberg, K. Leinonen, J. Lin, Y. Liu, M. Miller, S. M. Reynolds, H. Rovira, I. Shmulevich, V. Thorsson, D. Yang, W. Zhang, S. Amin, C. J. Wu, C. C. Wu, R. Akbani, K. Aldape, K. A. Baggerly, B. Broom, T. D. Casasent, J. Cleland, C. Creighton, D. Dodda, M. Edgerton, L. Han, S. M. Herbrich, Z. Ju, H. Kim, S. Lerner, J. Li, H. Liang, W. Liu, P. L. Lorenzi, Y. Lu, J. Melott, G. B. Mills, L. Nguyen, X. Su, R. Verhaak, W. Wang, J. N. Weinstein, A. Wong, Y. Yang, J. Yao, R. Yao, K. Yoshihara, Y. Yuan, A. K. Yung, N. Zhang, S. Zheng, M. Ryan, D. W. Kane, B. A. Aksoy, G. Ciriello, G. Dresdner, J. Gao, B. Gross, A. Jacobsen, A. Kahles, M. Ladanyi, W. Lee, K. V. Lehmann, M. L. Miller, R. Ramirez, G. tsch, B. Reva, C. Sander, N. Schultz, Y. Senbabaoglu, R. Shen, R. Sinha, S. O. Sumer, Y. Sun, B. S. Taylor, N. Weinhold, S. Fei, P. Spellman, C. Benz, D. Carlin, M. Cline, B. Craft, K. Ellrott, M. Goldman, D. Haussler, S. Ma, S. Ng, E. Paull, A. Radenbaugh, S. Salama, A. Sokolov, J. M. Stuart, T. Swatloski, V. Uzunangelov, P. Waltman, C. Yau, J. Zhu, S. R. Hamilton, G. Getz, C. Sougnez, S. Abbott, R. Abbott, N. D. Dees, K. Delehaunty, L. Ding, D. J. Dooling, J. M. Eldred, C. C. Fronick, R. Fulton, L. L. Fulton, J. Kalicki-Veizer, K. L. Kanchi, C. Kandoth, D. C. Koboldt, D. E. Larson, T. J. Ley, L. Lin, C. Lu, V. J. Magrini, E. R. Mardis, M. D. McLellan, J. F. McMichael, C. A. Miller, M. O'Laughlin, C. Pohl, H. Schmidt, S. M. Smith, J. Walker, J. W. Wallis, M. C. Wendl, R. K. Wilson, T. Wylie, Q. Zhang, R. Burton, M. A. Jensen, A. Kahn, T. Pihl, D. Pot, Y. Wan, D. A. Levine, A. D. Black, J. Bowen, J. Frick, J. M. Gastier-Foster, H. A. Harper, C. Helsel, K. M. Leraas, T. M. Lichtenberg, C. McAllister, N. C. Ramirez, S. Sharpe, L. Wise, E. Zmuda, S. J. Chanock, T. Davidsen, J. A. Demchok, G. Eley, I. Felau, B. A. Ozenberger, M. Sheth, H. Sofia, L. Staudt, R. Tarnuzzer, Z. Wang, L. Yang, J. Zhang, L. Omberg, A. Margolin, B. J. Raphael, F. Vandin, H. T. Wu, M. D. Leiserson, S. C. Benz, C. J. Vaske, H. Noushmehr, T. Knijnenburg, D. Wolf, L. Van 't Veer, E. A. Collisson, D. Anastassiou, T. H. Ou Yang, N. Lopez-Bigas, A. Gonzalez-Perez, D. Tamborero, Z. Xia, W. Li, D. Y. Cho, T. Przytycka, M. Hamilton, S. McGuire, S. Nelander, P. Johansson, R. rnsten, T. Kling, and J. Sanchez, "The Cancer Genome Atlas Pan-Cancer analysis project," Nat Genet. **45**, 1113–1120 (2013).
6. E. Papaemmanuil, M. Gerstung, L. Bullinger, V. I. Gaidzik, P. Paschka, N. D. Roberts, N. E. Potter, M. Heuser, F. Thol, N. Bolli, G. Gundem, P. Van Loo, I. Martincorena, P. Ganly, L. Mudie, S. McLaren, S. O'Meara, K. Raine, D. R. Jones, J. W. Teague, A. P. Butler, M. F. Greaves, A. Ganser, K. hner, R. F. Schlenk, H. hner, and P. J. Campbell, "Genomic Classification and Prognosis in Acute Myeloid Leukemia," N Engl J Med **374**, 2209–2221 (2016).
7. J. W. Tyner, C. E. Tognon, D. Bottomly, B. Wilmot, S. E. Kurtz, S. L. Savage, N. Long, A. R. Schultz, E. Traer, M. Abel, A. Agarwal, A. Blucher, U. Borate, J. Bryant, R. Burke, A. Carlos, R. Carpenter, J. Carroll, B. H. Chang, C. Coblentz, A. d'Almeida, R. Cook, A. Danilov, K. T. Dao, M. Degnin, D. Devine, J. Dibb, D. K. Edwards, C. A. Eide, I. English, J. Glover, R. Henson, H. Ho, A. Jemal, K. Johnson, R. Johnson, B. Junio, A. Kaempf, J. Leonard, C. Lin, S. Q. Liu, P. Lo, M. M. Loriaux, S. Luty, T. Macey, J. MacManiman, J. Martinez, M. Mori, D. Nelson, C. Nichols, J. Peters, J. Ramsdill, A. Rofelty, R. Schuff, R. Searles, E. Segerdell, R. L. Smith, S. E. Spurgeon, T. Sweeney, A. Thapa, C. Visser, J. Wagner, K. Watanabe-Smith, K. Werth, J. Wolf, L. White,

A. Yates, H. Zhang, C. R. Cogle, R. H. Collins, D. C. Connolly, M. W. Deininger, L. Drusbosky, C. S. Hourigan, C. T. Jordan, P. Kropf, T. L. Lin, M. E. Martinez, B. C. Medeiros, R. R. Pallapati, D. A. Pollyea, R. T. Swords, J. M. Watts, S. J. Weir, D. L. Wiest, R. M. Winters, S. K. McWeeney, and B. J. Druker, "Functional genomic landscape of acute myeloid leukaemia," Nature **562**, 526–531 (2018).

8. D. Pulte, F. A. Castro, L. Jansen, S. Luttmann, B. Holleczek, A. Nennecke, M. Ressing, A. Katalinic, and H. Brenner, "Trends in survival of chronic lymphocytic leukemia patients in Germany and the USA in the first decade of the twenty-first century," J Hematol Oncol **9**, 28 (2016).
9. Z. Song, W. Wu, W. Wei, W. Xiao, M. Lei, K. Cai, D. W. Huang, S. Jeong, J. P. Zhang, H. Wang, M. E. Kadin, T. A. Waldmann, L. M. Staudt, M. Nakagawa, and Y. Yang, "Analysis and Therapeutic Targeting of the IL-1R Pathway in Anaplastic Large Cell Lymphoma," Blood (2023).
10. K. A. Romine, T. Nechiporuk, D. Bottomly, S. Jeng, S. K. McWeeney, A. Kaempf, M. R. Corces, R. Majeti, and J. W. Tyner, "Monocytic differentiation and AHR signaling as Primary Nodes of BET Inhibitor Response in Acute Myeloid Leukemia," Blood Cancer Discov **2**, 518–531 (2021).
11. S. E. Kurtz, C. A. Eide, A. Kaempf, N. Long, D. Bottomly, O. Nikolova, B. J. Druker, S. K. McWeeney, B. H. Chang, J. W. Tyner, and A. Agarwal, "Associating drug sensitivity with differentiation status identifies effective combinations for acute myeloid leukemia," Blood Adv **6**, 3062–3067 (2022).
12. H. ki, A. M. ä, P. nen, M. Kontro, O. Dufva, D. Deb, B. Yadav, O. ck, A. Kumar, H. Everaus, B. T. Gjertsen, M. niemi, K. Porkka, S. Mustjoki, and C. A. Heckman, "Phenotype-based drug screening reveals association between venetoclax response and differentiation stage in acute myeloid leukemia," Haematologica **105**, 708–720 (2020).
13. S. Pei, D. A. Pollyea, A. Gustafson, B. M. Stevens, M. Minhajuddin, R. Fu, K. A. Riemondy, A. E. Gillen, R. M. Sheridan, J. Kim, J. C. Costello, M. L. Amaya, A. Inguva, A. Winters, H. Ye, A. Krug, C. L. Jones, B. Adane, N. Khan, J. Ponder, J. Schowinsky, D. Abbott, A. Hammes, J. R. Myers, J. M. Ashton, T. Nemkov, A. D'Alessandro, J. A. Gutman, H. E. Ramsey, M. R. Savona, C. A. Smith, and C. T. Jordan, "Monocytic Subclones Confer Resistance to Venetoclax-Based Therapy in Patients with Acute Myeloid Leukemia," Cancer Discov **10**, 536–551 (2020).
14. H. Zhang, Y. Nakauchi, T. hnke, M. Stafford, D. Bottomly, R. Thomas, B. Wilmot, S. K. McWeeney, R. Majeti, and J. W. Tyner, "Integrated analysis of patient samples identifies biomarkers for venetoclax efficacy and combination strategies in acute myeloid leukemia," Nat Cancer **1**, 826–839 (2020).
15. B. S. White, S. A. Khan, M. J. Mason, M. Ammad-Ud-Din, S. Potdar, D. Malani, H. ki, B. J. Druker, C. Heckman, O. Kallioniemi, S. E. Kurtz, K. Porkka, C. E. Tognon, J. W. Tyner, T. Aittokallio, K. Wennerberg, and J. Guinney, "Bayesian multi-source regression and monocyte-associated gene expression predict BCL-2 inhibitor resistance in acute myeloid leukemia," NPJ Precis Oncol **5**, 71 (2021).
16. E. Lainey, M. bert, S. pot, M. Scoazec, C. Bouteloup, C. Leroy, S. De Botton, L. Galluzzi, P. Fenaux, and G. Kroemer, "Erlotinib antagonizes ABC transporters in acute myeloid leukemia," Cell Cycle **11**, 4079–4092 (2012).
17. L. Xu, Y. Zhang, M. Gao, G. Wang, and Y. Fu, "Concurrent targeting Akt and sphingosine kinase 1 by A-674563 in acute myeloid leukemia cells," Biochem. Biophys Res Commun **472**, 662–668 (2016).
18. A. M. Cook, L. Li, Y. Ho, A. Lin, L. Li, A. Stein, S. Forman, D. Perrotti, R. Jove, and R. Bhatia, "Role of altered growth factor receptor-mediated JAK2 signaling in growth and maintenance of human acute myeloid leukemia stem cells," Blood **123**, 2826–2837 (2014).
19. C. Bogani, N. Bartalucci, S. Martinelli, L. Tozzi, P. Guglielmelli, A. Bosi, and A. M. Vannucchi, "mTOR inhibitors alone and in combination with JAK2 inhibitors effectively inhibit cells of myeloproliferative neoplasms," PLoS One **8**, e54826 (2013).
20. A. K. Garz, S. Wolf, S. Grath, V. Gaidzik, S. Habringer, B. Vick, M. Rudelius, C. Ziegenhain, S. Herold, M. T. Weickert, M. Smets, C. Peschel, R. A. J. Oostendorp, S. Bultmann, I. Jeremias, C. Thiede, K. hner, U. Keller, and K. S. tze, "-ITD AML with concurrent epigenetic mutations," Oncotarget **8**, 108738–108759 (2017).
21. S. Sucha, A. Sorf, M. Svoren, D. Vagiannis, F. Ahmed, B. Visek, and M. Ceckova, "ABCB1 as a potential beneficial target of midostaurin in acute myeloid leukemia," Biomed Pharmacother **150**, 112962 (2022).
22. J. Etchin, J. Montero, A. Berezovskaya, B. T. Le, A. Kentsis, A. L. Christie, A. S. Conway, W. C. Chen, C. Reed, M. R. Mansour, C. E. Ng, S. Adamia, S. J. Rodig, I. A. Galinsky, R. M. Stone, B. Klebanov, Y. Landesman, M. Kauffman, S. Shacham, A. L. Kung, J. C. Wang, A. Letai, and A. T. Look, "Activity of a selective inhibitor of nuclear export, selinexor (KPT-330), against AML-initiating cells engrafted into immunosuppressed NSG mice," Leukemia **30**, 190–199 (2016).
23. S. Park, N. Chapuis, V. Bardet, J. Tamburini, N. Gallay, L. Willems, Z. A. Knight, K. M. Shokat, N. Azar, F. é, N. Ifrah, F. Dreyfus, P. Mayeux, C. Lacombe, and D. Bouscary, "PI-103, a dual inhibitor of Class IA phosphatidylinositide 3-kinase and mTOR, has antileukemic activity in AML," Leukemia **22**, 1698–1706 (2008).
24. P. van Galen, V. Hovestadt, M. H. Wadsworth Ii, T. K. Hughes, G. K. Griffin, S. Battaglia, J. A. Verga, J. Stephansky, T. J. Pastika, J. Lombardi Story, G. S. Pinkus, O. Pozdnyakova, I. Galinsky, R. M. Stone, T. A. Graubert, A. K. Shalek, J. C. Aster, A. A. Lane, and B. E. Bernstein, "Single-Cell RNA-Seq Reveals AML Hierarchies Relevant to Disease Progression and Immunity," Cell **176**, 1265–1281 (2019).
25. L. C. Stetson, D. Balasubramanian, S. P. Ribeiro, T. Stefan, K. Gupta, X. Xu, S. Fourati, A. Roe, Z. Jackson, R. Schauner, A. Sharma, B. Tamilselvan, S. Li, M. de Lima, T. H. Hwang, R. Balderas, Y. Saunthararajah, J. Maciejewski, T. LaFramboise, J. S. Barnholtz-Sloan, R. P. Sekaly, and D. N. Wald, "Single cell RNA sequencing of AML initiating cells reveals RNA-based evolution during disease progression," Leukemia **35**, 2799–2812 (2021).

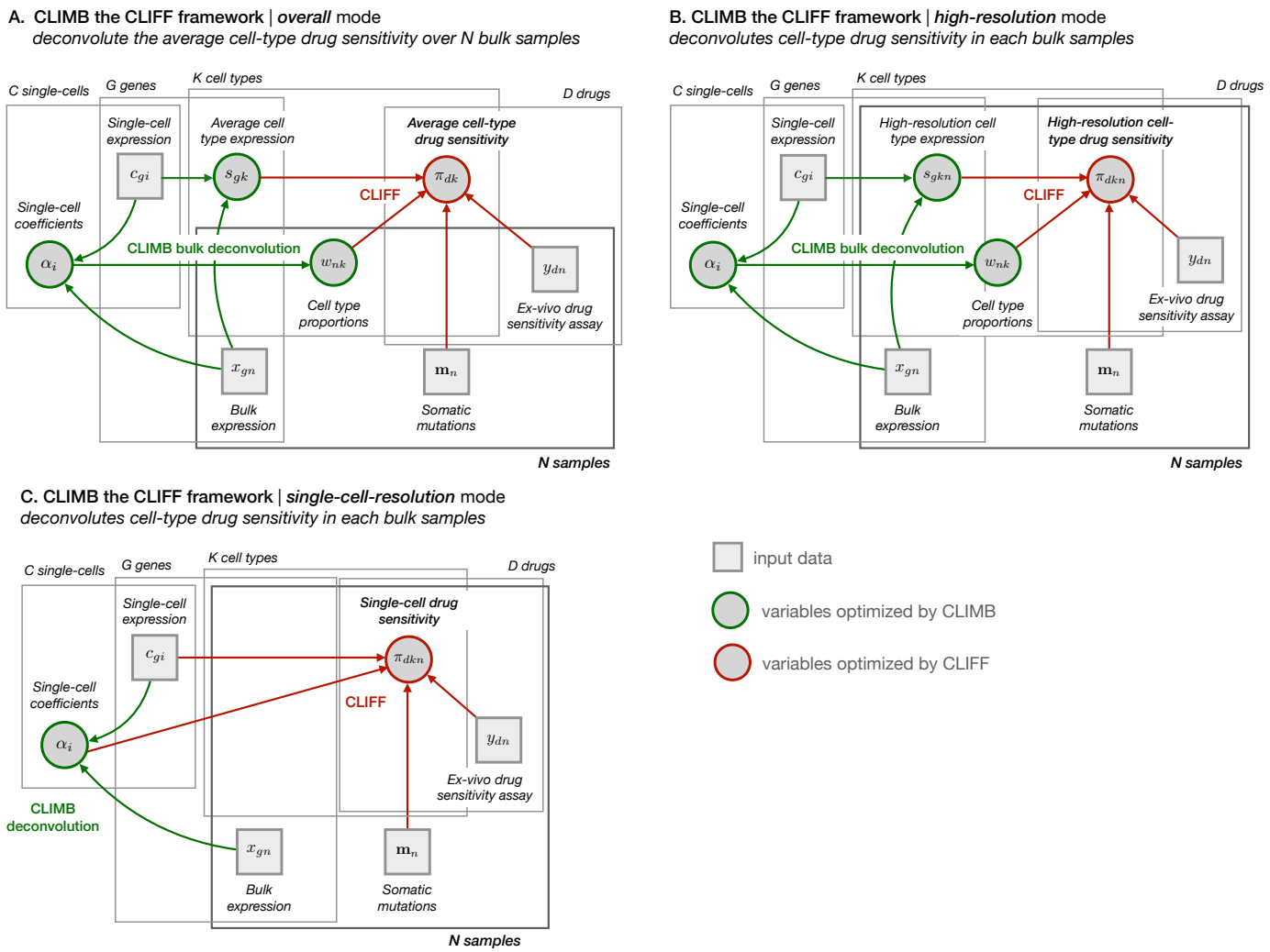
26. D. Vetrie, G. V. Helgason, and M. Copland, "Nat Rev CancerThe leukaemia stem cell: similarities, differences and clinical prospects in CML and AML," Nat Rev Cancer **20**, 158–173 (2020).
27. Y. Tan, Q. Wu, and F. Zhou, "Crit Rev Oncol HematolTargeting acute myeloid leukemia stem cells: Current therapies in development and potential strategies with new dimensions," Crit Rev Oncol Hematol **152**, 102993 (2020).
28. L. Velten, B. A. Story, P. Hernández-Malmierca, S. Raffel, D. R. Leonce, J. Milbank, M. Paulsen, A. Demir, C. Szu-Tu, R. Frömel, C. Lutz, D. Nowak, J. C. Jann, C. Pabst, T. Boch, W. K. Hofmann, C. Müller-Tidow, A. Trumpp, S. Haas, and L. M. Steinmetz, "Nat CommunIdentification of leukemic and pre-leukemic stem cells by clonal tracking from single-cell transcriptomics," Nat Commun **12**, 1366 (2021).
29. A. H. Elsayed, R. Rafiee, X. Cao, S. Raimondi, J. R. Downing, R. Ribeiro, Y. Fan, T. A. Gruber, S. Baker, J. Klco, J. E. Rubnitz, S. Pounds, and J. K. Lamba, "A six-gene leukemic stem cell score identifies high risk pediatric acute myeloid leukemia," Leukemia **34**, 735–745 (2020).
30. A. J. Gentles, S. K. Plevritis, R. Majeti, and A. A. Alizadeh, "Association of a leukemic stem cell gene expression signature with clinical outcomes in acute myeloid leukemia," JAMA **304**, 2706–2715 (2010).
31. S. W. Ng, A. Mitchell, J. A. Kennedy, W. C. Chen, J. McLeod, N. Ibrahimova, A. Arruda, A. Popescu, V. Gupta, A. D. Schimmer, A. C. Schuh, K. W. Yee, L. Bullinger, T. Herold, D. rlich, T. chner, W. Hiddemann, W. E. Berdel, B. rmann, M. Cheok, C. Preudhomme, H. Dombret, K. Metzeler, C. Buske, B. wenberg, P. J. Valk, P. W. Zandstra, M. D. Minden, J. E. Dick, and J. C. Wang, "A 17-gene stemness score for rapid determination of risk in acute leukaemia," Nature **540**, 433–437 (2016).
32. D. Bottomly, N. Long, A. R. Schultz, S. E. Kurtz, C. E. Tognon, K. Johnson, M. Abel, A. Agarwal, S. Avaylon, E. Benton, A. Blucher, U. Borate, T. P. Braun, J. Brown, J. Bryant, R. Burke, A. Carlos, B. H. Chang, H. J. Cho, S. Christy, C. Coblentz, A. M. Cohen, A. d'Almeida, R. Cook, A. Danilov, K. T. Dao, M. Degnin, J. Dibb, C. A. Eide, I. English, S. Hagler, H. Harrelson, R. Henson, H. Ho, S. K. Joshi, B. Junio, A. Kaempf, Y. Kosaka, T. Laderas, M. Lawhead, H. Lee, J. T. Leonard, C. Lin, E. F. Lind, S. Q. Liu, P. Lo, M. M. Loriaux, S. Luty, J. E. Maxson, T. Macey, J. Martinez, J. Minnier, A. Monteblanco, M. Mori, Q. Morrow, D. Nelson, J. Ramsdill, A. Rofelty, A. Rogers, K. A. Romine, P. Ryabinin, J. N. Saultz, D. A. Sampson, S. L. Savage, R. Schuff, R. Searles, R. L. Smith, S. E. Spurgeon, T. Sweeney, R. T. Swords, A. Thapa, K. Thiel-Klare, E. Traer, J. Wagner, B. Wilmot, J. Wolf, G. Wu, A. Yates, H. Zhang, C. R. Cogle, R. H. Collins, M. W. Deininger, C. S. Hourigan, C. T. Jordan, T. L. Lin, M. E. Martinez, R. R. Pallapati, D. A. Pollyea, A. D. Pomicter, J. M. Watts, S. J. Weir, B. J. Druker, S. K. McWeeney, and J. W. Tyner, "Integrative analysis of drug response and clinical outcome in acute myeloid leukemia," Cancer Cell **40**, 850–864 (2022).
33. A. G. X. Zeng, S. Bansal, L. Jin, A. Mitchell, W. C. Chen, H. A. Abbas, M. Chan-Seng-Yue, V. Voisin, P. van Galen, A. Tierens, M. Cheok, C. Preudhomme, H. Dombret, N. Daver, P. A. Futreal, M. D. Minden, J. A. Kennedy, J. C. Y. Wang, and J. E. Dick, "A cellular hierarchy framework for understanding heterogeneity and predicting drug response in acute myeloid leukemia," Nat Med **28**, 1212–1223 (2022).
34. E. S.-L. P. A. v. V. M. Z. J. J. v. D. A. M. O. C. J. H. P. v. B. H. V. M. J. R. M. G. E. B. v. d. A. E Onur Karakaslar, Jeppe Severens, "Single cell guided deconvolution of bulk transcriptomics recapitulates differentiation stages of acute myeloid leukemia and predicts drug response," Bioarxiv (2022).
35. T. Chu, Z. Wang, D. Pe'er, and C. G. Danko, "Cell type and gene expression deconvolution with BayesPrism enables Bayesian integrative analysis across bulk and single-cell RNA sequencing in oncology," Nat Cancer **3**, 505–517 (2022).
36. A. M. Newman, C. B. Steen, C. L. Liu, A. J. Gentles, A. A. Chaudhuri, F. Scherer, M. S. Khodadoust, M. S. Esfahani, B. A. Luca, D. Steiner, M. Diehn, and A. A. Alizadeh, "Determining cell type abundance and expression from bulk tissues with digital cytometry," Nat Biotechnol **37**, 773–782 (2019).
37. Y. Chen, Y. Wang, Y. Chen, Y. Cheng, Y. Wei, Y. Li, J. Wang, Y. Wei, T. F. Chan, and Y. Li, "Deep autoencoder for interpretable tissue-adaptive deconvolution and cell-type-specific gene analysis," Nat Commun **13**, 6735 (2022).
38. K. Menden, M. Marouf, S. Oller, A. Dalmia, D. S. Magruder, K. Kloiber, P. Heutink, and S. Bonn, "Deep learning-based cell composition analysis from tissue expression profiles," Sci Adv **6**, eaba2619 (2020).
39. X. Wang, J. Park, K. Susztak, N. R. Zhang, and M. Li, "Nat CommunBulk tissue cell type deconvolution with multi-subject single-cell expression reference," Nat Commun **10**, 380 (2019).
40. B. Jew, M. Alvarez, E. Rahmani, Z. Miao, A. Ko, K. M. Garske, J. H. Sul, K. H. Pietiläinen, P. Pajukanta, and E. Halperin, "Nat CommunAccurate estimation of cell composition in bulk expression through robust integration of single-cell information," Nat Commun **11**, 1971 (2020).
41. M. M. Naldini, G. Casirati, M. Barcella, P. M. V. Rancoita, A. Cosentino, C. Caserta, F. Pavesi, E. Zonari, G. Desantis, D. Gilioli, M. G. Carrabba, L. Vago, M. Bernardi, R. Di Micco, C. Di Serio, I. Merelli, M. Volpin, E. Montini, F. Ciceri, and B. Gentner, "Longitudinal single-cell profiling of chemotherapy response in acute myeloid leukemia," Nat Commun **14**, 1285 (2023).
42. H. O. Lee, Y. Hong, H. E. Etlioglu, Y. B. Cho, V. Pomella, B. Van den Bosch, J. Vanhecke, S. Verbandt, H. Hong, J. W. Min, N. Kim, H. H. Eum, J. Qian, B. Boeckx, D. Lambrechts, P. Tsantoulis, G. De Hertogh, W. Chung, T. Lee, M. An, H. T. Shin, J. G. Joung, M. H. Jung, G. Ko, P. Wirapati, S. H. Kim, H. C. Kim, S. H. Yun, I. B. H. Tan, B. Ranjan, W. Y. Lee, T. Y. Kim, J. K. Choi, Y. J. Kim, S. Prabhakar, S. Tejpar, and W. Y. Park, "Lineage-dependent gene expression programs influence the immune landscape of colorectal cancer," Nat Genet. **52**, 594–603 (2020).
43. A. M. Khaliq, C. Erdogan, Z. Kurt, S. S. Turgut, M. W. Grunvald, T. Rand, S. Khare, J. A. Borgia, D. M. Hayden, S. G. Pappas, H. R. Govekar, A. E. Kam,

J. Reiser, K. Turaga, M. Radovich, Y. Zang, Y. Qiu, Y. Liu, M. L. Fishel, A. Turk, V. Gupta, R. Al-Sabti, J. Subramanian, T. M. Kuzel, A. Sadanandam, L. Waldron, A. Hussain, M. Saleem, B. El-Rayes, A. A. Salahudeen, and A. Masood, "Refining colorectal cancer classification and clinical stratification through a single-cell atlas," Genome Biol **23**, 113 (2022).

44. G. K. Gray, C. M. Li, J. M. Rosenbluth, L. M. Selfors, N. Girnius, J. R. Lin, R. C. J. Schackmann, W. L. Goh, K. Moore, H. K. Shapiro, S. Mei, K. D'Andrea, K. L. Nathanson, P. K. Sorger, S. Santagata, A. Regev, J. E. Garber, D. A. Dillon, and J. S. Brugge, "A human breast atlas integrating single-cell proteomics and transcriptomics," Dev Cell **57**, 1400–1420 (2022).
45. S. Z. Wu, G. Al-Eryani, D. L. Roden, S. Junankar, K. Harvey, A. Andersson, A. Thennavan, C. Wang, J. R. Torpy, N. Bartonicek, T. Wang, L. Larsson, D. Kaczorowski, N. I. Weisenfeld, C. R. Uytingco, J. G. Chew, Z. W. Bent, C. L. Chan, V. Gnanasambandapillai, C. A. Dutertre, L. Gluch, M. N. Hui, J. Beith, A. Parker, E. Robbins, D. Segara, C. Cooper, C. Mak, B. Chan, S. Warrier, F. Ginhoux, E. Millar, J. E. Powell, S. R. Williams, X. S. Liu, S. O'Toole, E. Lim, J. Lundeberg, C. M. Perou, and A. Swarbrick, "A single-cell and spatially resolved atlas of human breast cancers," Nat Genet. **53**, 1334–1347 (2021).
46. C. Neftel, J. Laffy, M. G. Filbin, T. Hara, M. E. Shore, G. J. Rahme, A. R. Richman, D. Silverbush, M. L. Shaw, C. M. Hebert, J. Dewitt, S. Gritsch, E. M. Perez, L. N. Gonzalez Castro, X. Lan, N. Druck, C. Rodman, D. Dionne, A. Kaplan, M. S. Bertalan, J. Small, K. Pelton, S. Becker, D. Bonal, Q. D. Nguyen, R. L. Servis, J. M. Fung, R. Mylvaganam, L. Mayr, J. Gojo, C. Haberler, R. Geyeregger, T. Czech, I. Slavc, B. V. Nahed, W. T. Curry, B. S. Carter, H. Wakimoto, P. K. Brastianos, T. T. Batchelor, A. Stemmer-Rachamimov, M. Martinez-Lage, M. P. Frosch, I. Stamenkovic, N. Riggi, E. Rheinbay, M. Monje, O. Rozenblatt-Rosen, D. P. Cahill, A. P. Patel, T. Hunter, I. M. Verma, K. L. Ligon, D. N. Louis, A. Regev, B. E. Bernstein, I. Tirosh, and M. L. à, "An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma," Cell. **178**, 835–849 (2019).
47. L. Jerby-Arnon, P. Shah, M. S. Cuoco, C. Rodman, M. J. Su, J. C. Melms, R. Leeson, A. Kanodia, S. Mei, J. R. Lin, S. Wang, B. Rabasha, D. Liu, G. Zhang, C. Margolais, O. Ashenberg, P. A. Ott, E. I. Buchbinder, R. Haq, F. S. Hodi, G. M. Boland, R. J. Sullivan, D. T. Frederick, B. Miao, T. Moll, K. T. Flaherty, M. Herlyn, R. W. Jenkins, R. Thummalapalli, M. S. Kowalczyk, I. adas, B. Schilling, A. N. R. Cartwright, A. M. Luoma, S. Malu, P. Hwu, C. Bernatchez, M. A. Forget, D. A. Barbie, A. K. Shalek, I. Tirosh, P. K. Sorger, K. Wucherpfennig, E. M. Van Allen, D. Schadendorf, B. E. Johnson, A. Rotem, O. Rozenblatt-Rosen, L. A. Garraway, C. H. Yoon, B. Izar, and A. Regev, "A Cancer Cell Program Promotes T Cell Exclusion and Resistance to Checkpoint Blockade," Cell. **175**, 984–997 (2018).
48. I. Tirosh, B. Izar, S. M. Prakadan, M. H. Wadsworth, D. Treacy, J. J. Trombetta, A. Rotem, C. Rodman, C. Lian, G. Murphy, M. Fallahi-Sichani, K. Dutton-Regester, J. R. Lin, O. Cohen, P. Shah, D. Lu, A. S. Genshaft, T. K. Hughes, C. G. Ziegler, S. W. Kazer, A. Gaillard, K. E. Kolb, A. C. Villani, C. M. Johannessen, A. Y. Andreev, E. M. Van Allen, M. Bertagnolli, P. K. Sorger, R. J. Sullivan, K. T. Flaherty, D. T. Frederick, J. Valbuena, C. H. Yoon, O. Rozenblatt-Rosen, A. K. Shalek, A. Regev, and L. A. Garraway, "Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq," Science. **352**, 189–196 (2016).
49. ?, https://data.leucegene.iricca/ .
50. T. J. Ley, C. Miller, L. Ding, B. J. Raphael, A. J. Mungall, A. Robertson, K. Hoadley, T. J. Triche, P. W. Laird, J. D. Baty, L. L. Fulton, R. Fulton, S. E. Heath, J. Kalicki-Veizer, C. Kandoth, J. M. Klco, D. C. Koboldt, K. L. Kanchi, S. Kulkarni, T. L. Lamprecht, D. E. Larson, L. Lin, C. Lu, M. D. McLellan, J. F. McMichael, J. Payton, H. Schmidt, D. H. Spencer, M. H. Tomasson, J. W. Wallis, L. D. Wartman, M. A. Watson, J. Welch, M. C. Wendl, A. Ally, M. Balasundaram, I. Birol, Y. Butterfield, R. Chiu, A. Chu, E. Chuah, H. J. Chun, R. Corbett, N. Dhalla, R. Guin, A. He, C. Hirst, M. Hirst, R. A. Holt, S. Jones, A. Karsan, D. Lee, H. I. Li, M. A. Marra, M. Mayo, R. A. Moore, K. Mungall, J. Parker, E. Pleasance, P. Plettner, J. Schein, D. Stoll, L. Swanson, A. Tam, N. Thiessen, R. Varhol, N. Wye, Y. Zhao, S. Gabriel, G. Getz, C. Sougnez, L. Zou, M. D. Leiserson, F. Vandin, H. T. Wu, F. Applebaum, S. B. Baylin, R. Akbani, B. M. Broom, K. Chen, T. C. Motter, K. Nguyen, J. N. Weinstein, N. Zhang, M. L. Ferguson, C. Adams, A. Black, J. Bowen, J. Gastier-Foster, T. Grossman, T. Lichtenberg, L. Wise, T. Davidsen, J. A. Demchok, K. R. Shaw, M. Sheth, H. J. Sofia, L. Yang, J. R. Downing, G. Eley, S. Alonso, B. Ayala, J. Baboud, M. Backus, S. P. Barletta, D. L. Berton, A. L. Chu, S. Girshik, M. A. Jensen, A. Kahn, P. Kothiyal, M. C. Nicholls, T. D. Pihl, D. A. Pot, R. Raman, R. N. Sanbhadti, E. E. Snyder, D. Srinivasan, J. Walton, Y. Wan, Z. Wang, J. P. Issa, M. Le Beau, M. Carroll, H. Kantarjian, S. Kornblau, M. S. Bootwalla, P. H. Lai, H. Shen, D. J. Van Den Berg, D. J. Weisenberger, D. C. Link, M. J. Walter, B. A. Ozenberger, E. R. Mardis, P. Westervelt, T. A. Graubert, J. F. DiPersio, and R. K. Wilson, "Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia," N Engl J Med **368**, 2059–2074 (2013).
51. J. M. Bennett, D. Catovsky, M. T. Daniel, G. Flandrin, D. A. Galton, H. R. Gralnick, and C. Sultan, "Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group," Br J Haematol **33**, 451–458 (1976).
52. J. Friedman, T. Hastie, and R. Tibshirani, "Regularization Paths for Generalized Linear Models via Coordinate Descent," J Stat Softw **33**, 1–22 (2010).

Drug	Cell subtype	Type	CLIFF's prediction (rank)	Methodology	Authors (Journal)
erlotinib	HSC-like, Prog-like	sensitive	3/19 (HSC-like), 1/19 (Prog-like)	CD34+ population	Elodie Lainey et al. (Cell Cycle)
sorafenib	HSC-like, Prog-like	sensitive	6/19 (HSC-like), 5/19 (Prog-like)	LSCs were isolated by FACS and the blast population enriched to >90% using immunomagnetic beads	Frances Linzee Mabrey et al. (Blood)
selumetinib	Monocyte-like	sensitive	1/19 (Mono-like)	Comparison of all blasts versus LSCs, showed that total blasts was having higher sensitivity than LSCs	Frances Linzee Mabrey et al. (Blood)
doramapimod	Monocyte-like	sensitive	2/19 (Mono-like)	M4/M5 French-American-British classification and CD14+ immunophenotype	Stephen E Kurtz et al. (Blood Advances)
venetoclax	Monocyte-like	resistant	1/19 (Mono-like)	M4/M5 French-American-British classification and CD14+ immunophenotype	Stephen E Kurtz et al. (Blood Advances)
				Resistance of M4/M5 patients to venetoclax	Heikki Kuusanmäki et al. (Haematologica)
				Responses to venetoclax +azacitidine in patients with AML correlate closely with developmental stage: phenotypically primitive AML is sensitive but monocytic AML is more resistant.	Shanshan Pei et al. (Cancer Discovery)
				Significant association between venetoclax resistance and high monocyte and neutrophil counts.	Zhang H. et al. (Nature Cancer)
venetoclax	HSC-like, Prog-like	sensitive	11/19 (HSC-like), 1/19 (Prog-like)	Sensitivity of M0/M1 patients to venetoclax	Heikki Kuusanmäki et al. (Haematologica)
				Responses to venetoclax +azacitidine in patients with AML correlate closely with developmental stage: phenotypically primitive AML is sensitive but monocytic AML is more resistant.	Shanshan Pei et al. (Cancer Discovery)
				Venetoclax response is linked to expression of monocyte-associated genes, shown by a multi-source Bayesian regression approach from bulk RNA-seq	White BS et al. (npj Precision Oncology)
midostaurin	HSC-like, Prog-like	sensitive	1/19 (HSC-like), 19/19 (Prog-like)	Sensitivity of CD34+ patients	Simona Sucha et al. (Biomed Pharmacother.)
trametinib	Monocyte-like	sensitive	1/19 (Mono-like)	Resistance of M4/M5 patients to trametinib	Heikki Kuusanmäki et al. (Haematologica)
ruxolitinib	Monocyte-like	sensitive	18/19 (Mono-like)	Resistance of M4/M5 patients to ruxolitinib	Heikki Kuusanmäki et al. (Haematologica)
JQ1	Monocyte-like	sensitive	1/19 (Mono-like)	Performed genome-wide CRISPR screens on BETi- sensitive and BETi-resistant AML cells. Both strategies revealed regulators of monocytic differentiation as determinants of BETi response (JQ1 is a BETi).	Romine KA et al. (Blood Cancer Discov.)
selinexor	HSC-like, Prog-like	sensitive	4/19 (HSC-like), 16/19 (Prog-like)	selinexor (KPT-330) was effective against AML-initiating cells engrafted into immunosuppressed NSG mice	J. Etchin et al. (Leukemia)
A-674563	HSC-like, Prog-like	sensitive	2/19 (HSC-like), 4/19 (Prog-like)	A-674563 is cytotoxic for AML progenitor cells. A-674563 activates caspase-3/9 and apoptosis in AML progenitor cells.	Lin Xu et al. (Biochemical and Biophysical Res. Comm.)
Crenolanib	HSC-like, Prog-like	sensitive	1/19 (HSC-like), 6/19 (Prog-like)	Azacitidine combined with the selective FLT3 kinase inhibitor crenolanib disrupts stromal protection and inhibits expansion of residual leukemia-initiating cells in FLT3-ITD AML	Anne-Kathrin Garz et al. (Oncotarget)
AZD1480	HSC-like, Prog-like	sensitive	4/19 (HSC-like), 3/19 (Prog-like)	AZD1480, a potent ATP competitive inhibitor of JAK1/2 kinases, resulted in rapid (30 minutes) and sustained (72 hours) inhibition of pJAK2 and pSTAT35 levels in AML CD34cells	Amy M. Cook et al. (Blood)
				Primary hematopoietic progenitors from MPN patients were challenged with AZD1480 and ruxolitinib.	Costanza Bogani (PLoS One)

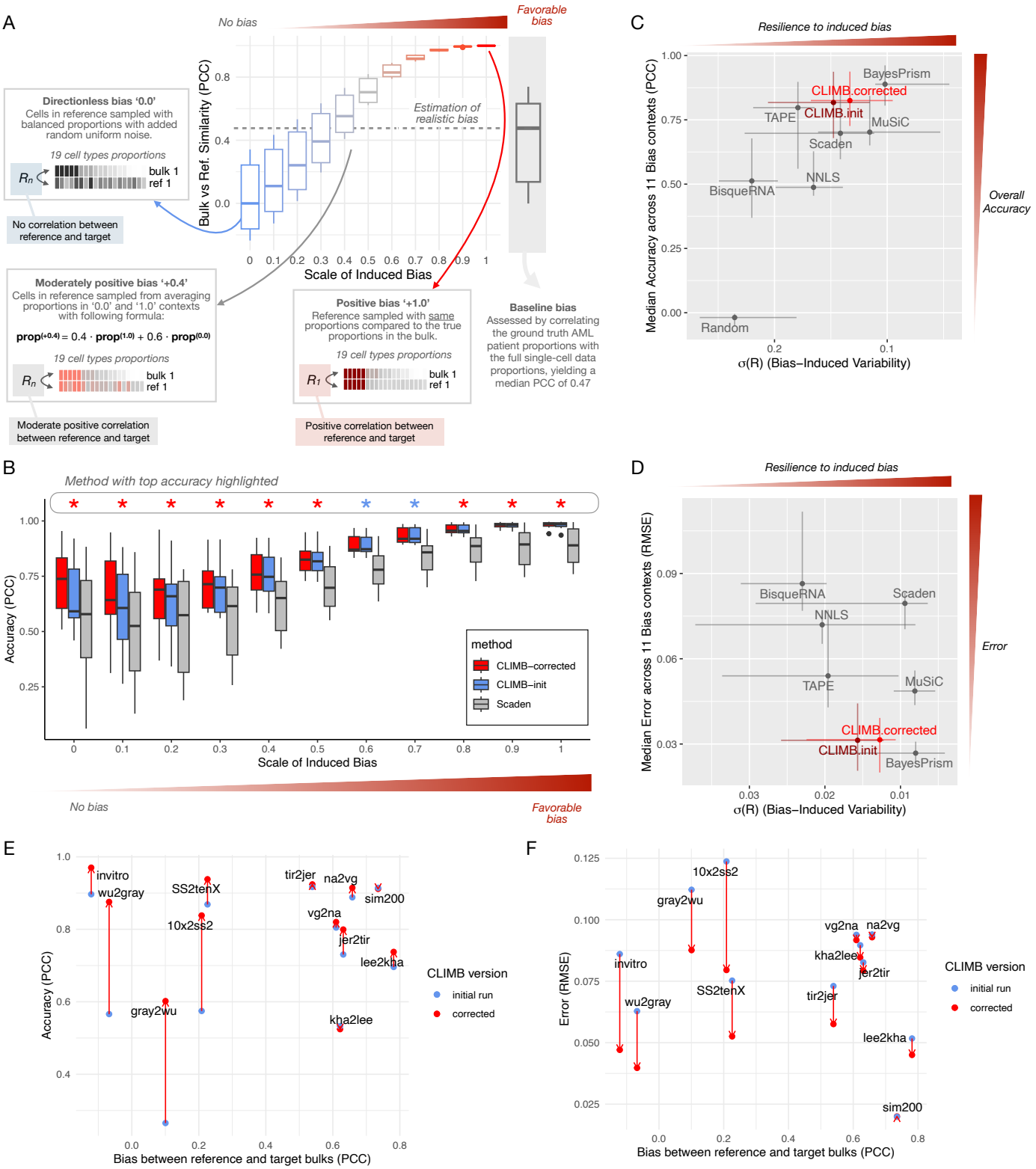
Supplementary Table 1. Documented cell-subtype drug sensitivity in Acute Myeloid Leukemia. Green signifies CLIFF's predictions aligned with the literature (25% most sensitive/resistant). Red denotes CLIFF's predictions conflicting with the literature (25% less sensitive/resistant). Orange represents uncertain connections, with values predicted between the 25% and 75% quartiles.



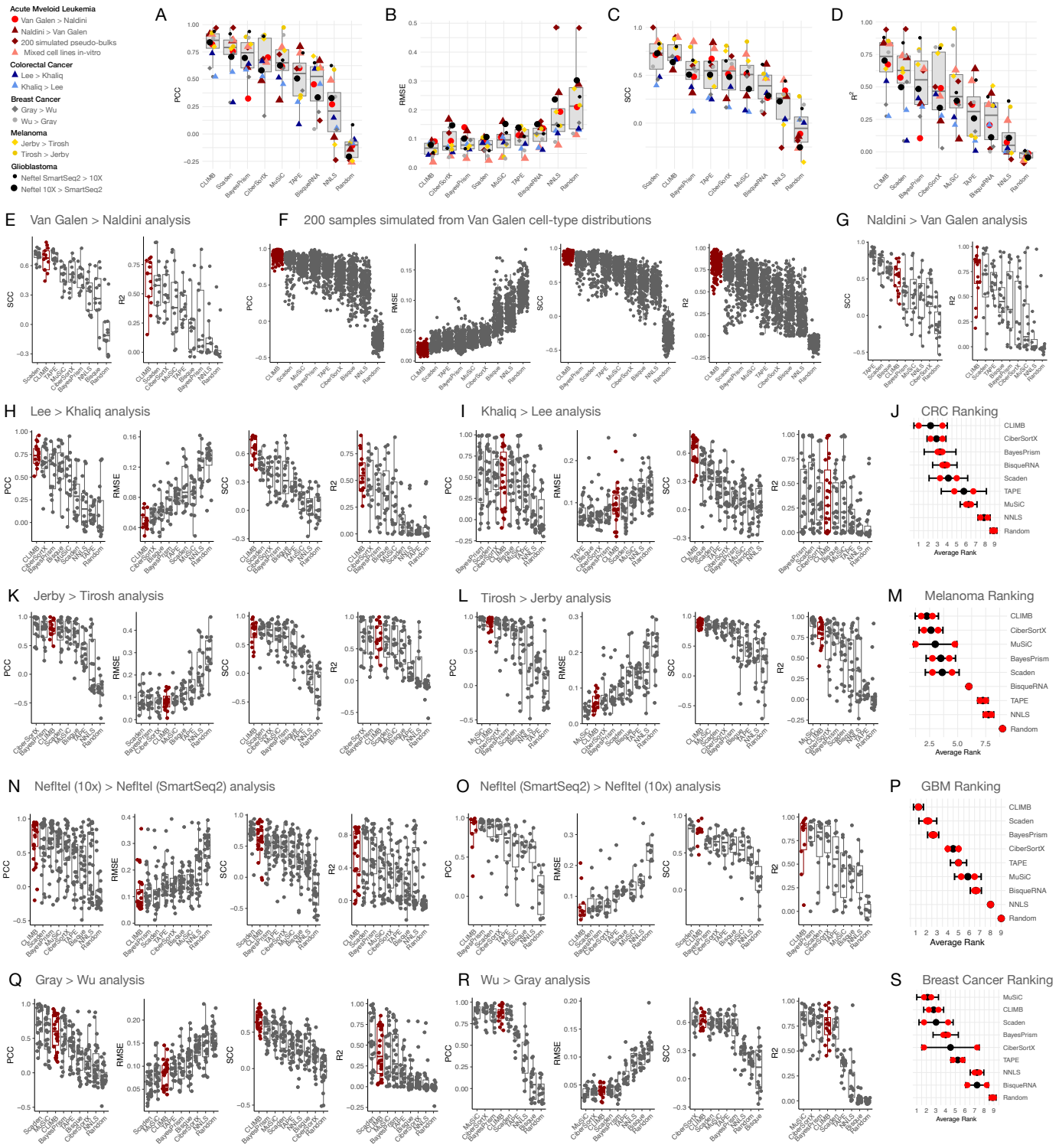
Supplementary Figure 1. (**A**) Diagram of CLIMB / CLIFF input / output variables in overall mode. (**B**) Same as (**A**) in high-resolution mode: cell-subtype expression and cell-subtype drug sensitivity is predicted in each bulk sample.

infer $\mathbf{W}^0$ from initial CLIMB run.
$\mathbf{W}^{prior} \leftarrow \mathbf{W}^0$
First pass
for $n \in N$ **do**
  Bulk-specific gene selection
  Sub-sample cells in $\mathbf{C}$ with $\mathbf{W}^{prior}_{n,-}$ proportions to compute bulk-specific reference $\mathbf{C}_n$
  Infer $\mathbf{W}^1_{n,-}$ by running CLIMB with $\mathbf{C}_n$ as a reference matrix
end
$\mathbf{W}^{prior} \leftarrow \mathbf{W}^1$
Second pass
for $n \in N$ **do**
  Bulk-specific gene selection
  Sub-sample cells in $\mathbf{C}$ with $\mathbf{W}^{prior}_{n,-}$ proportions to compute bulk-specific reference $\mathbf{C}_n$
  Infer $\mathbf{W}^2_{n,-}$ by running CLIMB with $\mathbf{C}_n$ as a reference matrix
end
Assess bias b_n for each bulk with $\mathrm{cor}(\mathbf{W}^0_{n,-}, \mathbf{w}^{(ref)})$
if $b_n \leq 0$ **then**
  $\mathbf{W}^{final} \leftarrow \mathbf{W}^2$
else
  $\mathbf{W}^{final} \leftarrow b_n\mathbf{W}^0 + (1 - b_n)\mathbf{W}^2$
end
Return $\mathbf{W}^{final}$ *as the final prediction.*

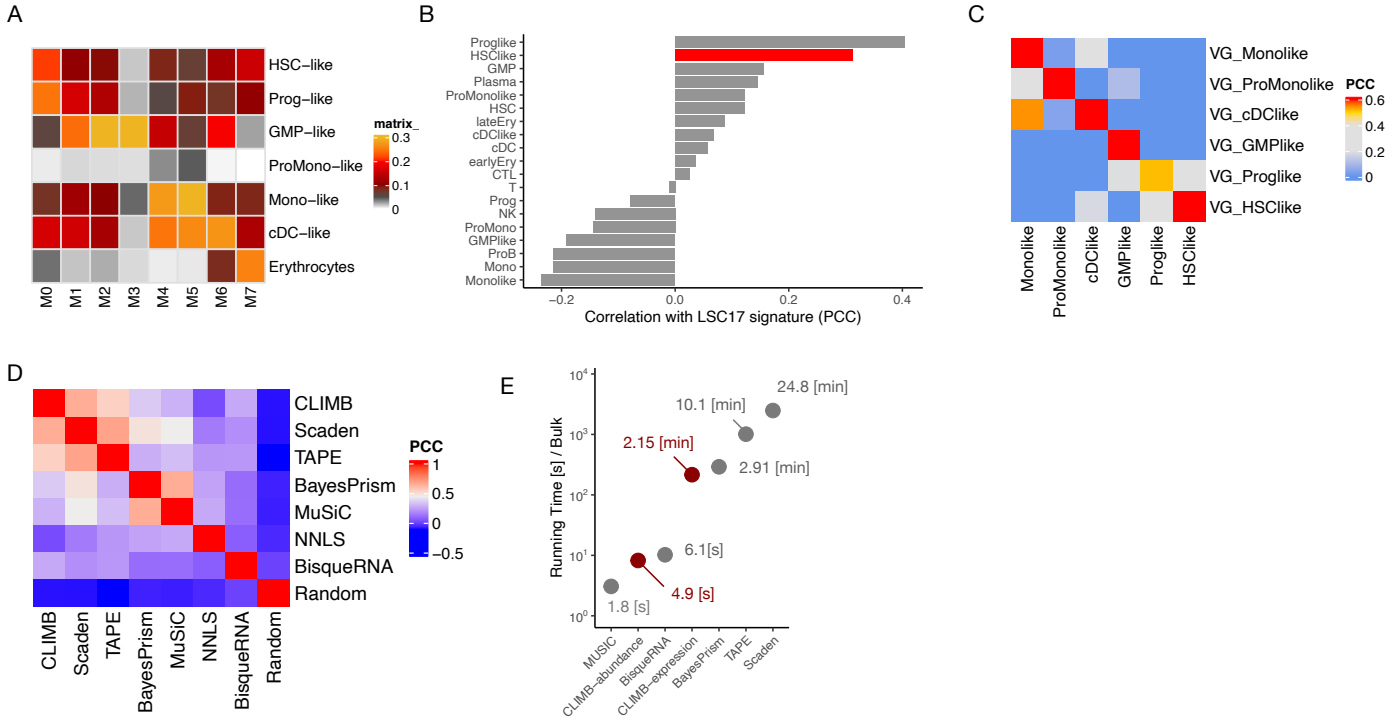
Supplementary Figure 2. CLIMB 2-pass algorithm. The algorithm subset cells in the reference single-cell matrix C two times. W represents the predicted cell-type proportions, of size NxK, N being the number of bulk sample and K the number of cell types. The first pass allows to subsample cells in the reference matrix, which is then used during the second pass to infer the cell subtype proportions. This process is iterated twice.



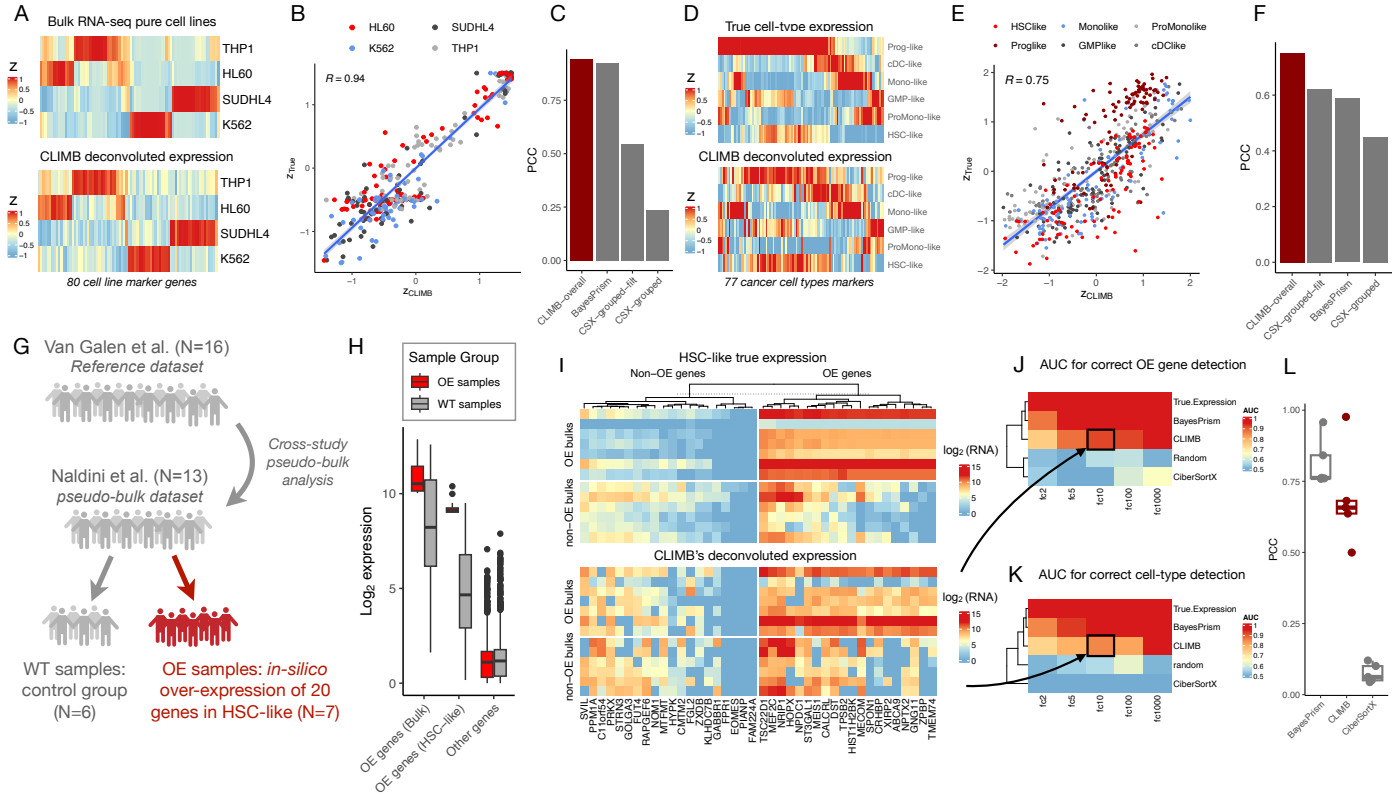
Supplementary Figure 3. (**A**) Illustration of the simulated pseudo-bulks with varying biases between reference and target proportions. On the right, we show the bias by comparing Van Galen et al. data with its own pseudo-bulks. (**B**) Accuracy, shown as a correlation coefficient (PCC), comparing true vs. predicted cell type fractions across 11 simulated scenarios for CLIMB-init, CLIMB-corrected, and Scaden methods. (**C**) Comparative accuracy and resilience to bias for all deconvolution methods. Resilience to bias is assumed to be equal to the variability in accuracy across different biases. (**D**) Same as (C), but showing the results as an error (RMSE). (**E**) Comparison between CLIMB-init and CLIMB-corrected deconvolution accuracy (PCC) across all analyses (in vitro and pseudo-bulks). (**F**) Same as (E), but showing the error instead.



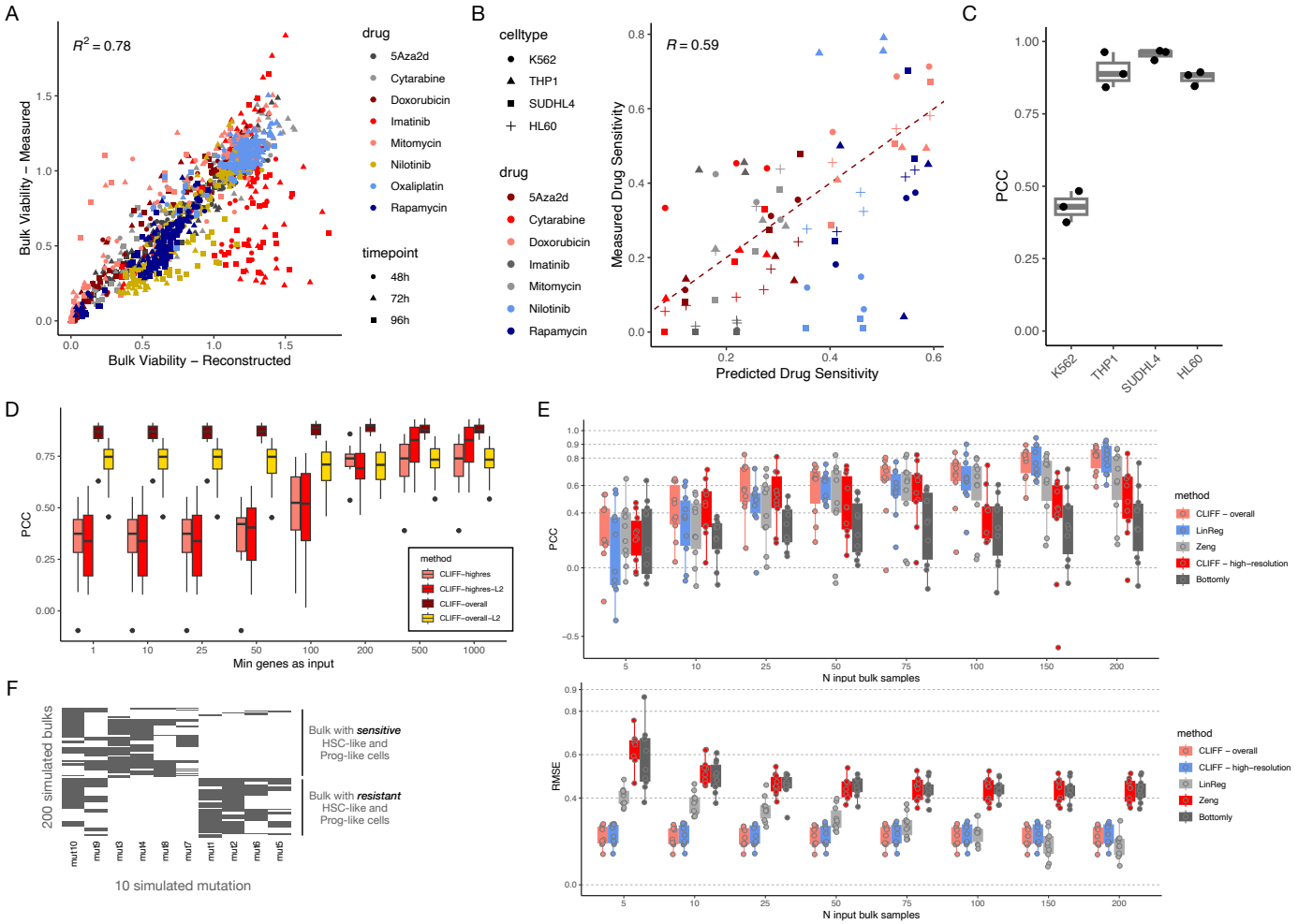
Supplementary Figure 4. (**A-D**) Combined pseudo-bulk deconvolution analysis across datasets showing four accuracy metrics. (**E**) Pseudo-bulk deconvolution analysis across datasets (Van Galen > Naldini) displaying accuracy as Spearman correlation coefficient / R-square value per sample (one data point per sample), comparing true and predicted cell subtype proportions. (**F**) Deconvolution method prediction accuracy for 200 simulated pseudo-bulk samples derived from the Van Galen et al. dataset. Results presented as boxplots, per sample level, for PCC, SCC, RMSE, and R-square metrics. (**G**) Similar to (**E**), but for Naldini > Van Galen cross-dataset deconvolution analysis. (**H-J**) Cross-dataset deconvolution analysis in colorectal cancer, depicted as boxplots for PCC, SCC, RMSE, and R-square metrics, covering Lee > Khaliq analysis (**H**) and Khaliq > Lee analysis (**I**). (**J**) Average rank over 4 metrics considering Lee > Khaliq and Khaliq > Lee cross-dataset analyses. (**K-M**) Melanoma cross-dataset deconvolution analysis, showcasing the same metrics as (**H-J**), for Jerby > Tirosh analysis (**K**) and Tirosh > Jerby analysis (**L**). (**N-P**) Glioblastoma cross-dataset deconvolution analysis, for Neftel (10X) > Neftel (SmartSeq2) analysis (**N**) and Neftel (SmartSeq2) > Neftel (10X) analysis (**O**). (**Q-S**) Cross-dataset deconvolution analysis in breast cancer, for Gray > Wu analysis (**Q**) and Wu > Gray analysis (**R**).



Supplementary Figure 5. (**A**) Average cell type proportions in different categories of FAB patients. (**B**) Correlation (PCC) between CLIMB-predicted cell subtype proportions and LSC17 signature. (**C**) Correlation (PCC) between Van Galen cell type signature and cell type proportions estimated by CLIMB. (**D**) Correlation (PCC) between cell subtype abundance estimated by different deconvolution methods in merged BeatAML, Leucegene, TCGA-AML datasets. (**E**) Running time of various deconvolution methods.

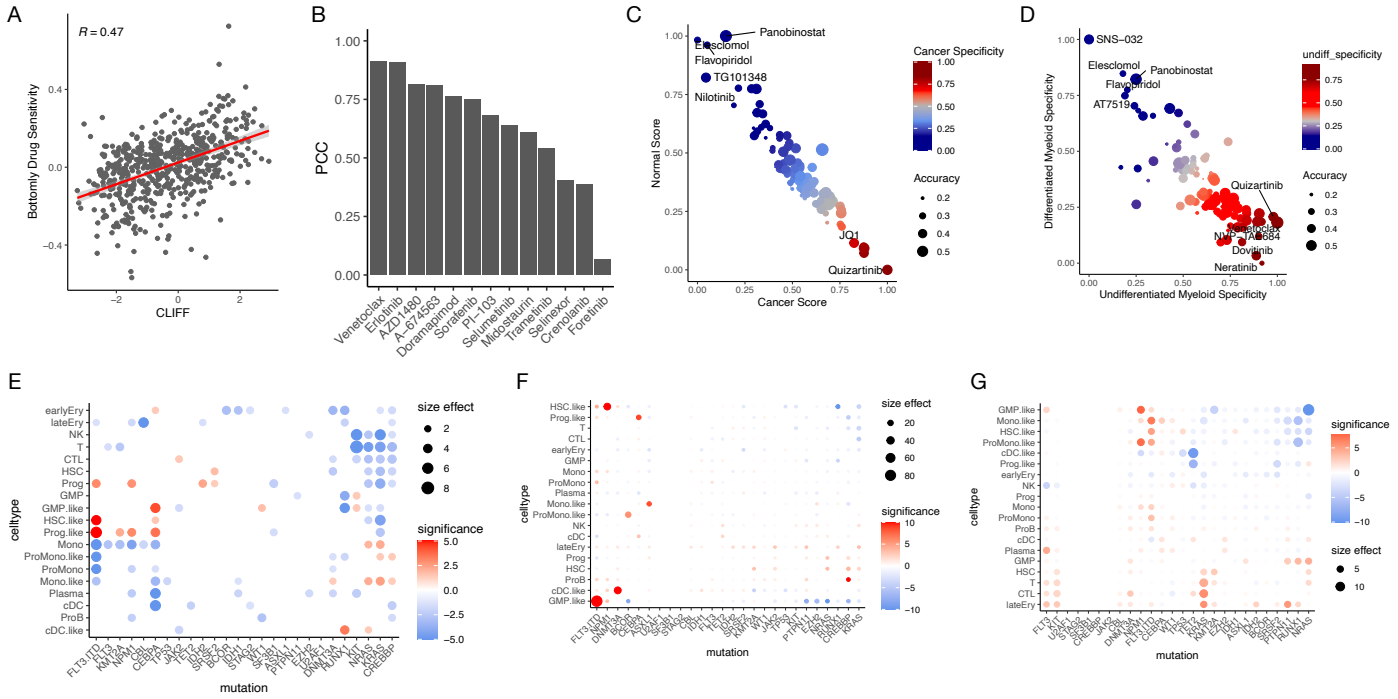


Supplementary Figure 6. Deconvolution of cell subtype Expression and Simulated Over-Expression Evaluation. (**A**) Heatmap comparing standardized true expression from bulk RNA-seq on pure cell lines (top) with deconvoluted expression from CLIMB based on bulk RNA-seq of cell line mixes (bottom). (**B**) Scatter plot comparing standardized true expression (y-axis) and CLIMB's deconvoluted expression (x-axis) from bulk RNA-seq. (**C**) Accuracy as PCC comparing true expression from bulk RNA-seq and deconvoluted expression from CLIMB, BayesPrism, CiberSortX (group mode, filtered and unfiltered). (**D**) Heatmap showing standardized true expression from pseudo-bulks (top) and CLIMB's deconvoluted expression of pseudo-bulks (bottom). (**E**) Scatter plot of deconvoluted expression (x-axis) versus true expression (y-axis) for pseudo-bulks. (**F**) Accuracy as PCC comparing true expression from bulk RNA-seq and deconvoluted expression from CLIMB, BayesPrism, CiberSortX (group mode, filtered and unfiltered). (**G**) Schematic of cross-dataset analysis with over-expression of 20 genes in 7/13 bulk samples and in HSC-like cell subtype. (**H**) Fold changes observed at bulk and cell-subtype levels when inducing single cells with a fold change of 10 before pseudo-bulk generation. Control genes are also compared. (**I**) Comparison between true and CLIMB's deconvoluted expression for HSC-like with over-expressed genes. Top heatmap displays true expression, with 20 over-expressed genes (right) and 20 random non-over-expressed genes (left). Overexpression samples are above, WT samples below. Bottom heatmap shows CLIMB's deconvoluted expression with the same gene and sample order. (**J**) Accuracy to detect correct over-expressed genes in HSC-like, using $-\log_{10}(\text{padj}) \cdot \text{FC}$ by DESeq2, presented as an average AUC-ROC over 20 genes. (**K**) Similar to (**J**), but showing accuracy to detect the correct cell subtype among cancer cell subtypes for each over-expressed gene, averaged over 20 genes. (**L**) Accuracy as PCC comparing true expression with CLIMB, BayesPrism, and CiberSortX deconvoluted expression.



]

Supplementary Figure 7. (**A**) Comparative analysis of true bulk viability (measured for 3 timepoints, 7 drugs, and 10 doses) against reconstructed bulk viability, using experimentally-assessed viability of pure cell lines comprising the bulks. (**B**) True cell line drug sensitivity for multiple timepoints, cell lines, and drugs, juxtaposed with CLIMB's high-resolution mode predictions (subsequently averaged at cell type level). (**C**) CLIFF's predictive accuracy presented as Pearson correlation coefficient (PCC), grouped by cell lines. (**D**) Accuracy comparison of CLIFF with varying numbers of input genes, based on the 200 simulated bulks. We evaluated four modes for CLIFF: CLIFF-overall, CLIFF-overall-L2 (CLIFF-overall with additional L2 regularization), CLIFF-highres (subsequent average at cell subtype level), and CLIFF-highres-L2 (CLIFF-highres with L2 regularization). (**E**) Predictive accuracy depicted as PCC for deconvolution of cell subtype proportions in 200 simulated pseudo-bulks, featuring CLIFF-overall, CLIFF-highres, Karakalsar's approach (linear regression using MuSiC-derived proportions as input matrix), and Bottomly's approach (correlation between cell-subtype score obtained from PC1 of cell subtype signature, and bulk drug sensitivity). (**F**) Simulated mutational landscape for 200 simulated bulks employed in high-resolution drug sensitivity deconvolution.



Supplementary Figure 8. (**A**) Scatter plot illustrating the Pearson correlation coefficient (PCC) between Bottomly's drug sensitivity predictions (for the 6 cancer cell subtypes defined by Van Galen et al.) and standardized CLIFF-overall predicted AUC (relative values of cell subtype drug sensitivity). Each data point represents a specific cell subtype-drug combination. (**B**) Corresponding to (**A**), this plot displays the overall PCC (averaging 6 data points for the 6 cancer cell subtypes) for the 14 drugs listed in Supplementary Table 1. (**C**) Predicted cancer cell specificity for drugs within the beatAML cohort, as projected by CLIFF-overall. (**D**) Forecasted drug specificity for undifferentiated versus differentiated blasts in the beatAML cohort, according to CLIFF-overall. (**E**) Overall correlations between mutations and cell subtype proportions using CLIMB's deconvoluted proportions on the beatAML cohort. A set of mutations with at least 5 positive samples is analyzed, with size-effect/significance depicted through t-test-derived measures. Positive correlations are represented in red, negative in blue. The size indicates the size effect (mean proportion difference). (**F**) Correlation between high-resolution deconvoluted drug sensitivity for each cell subtype and mutations in the beatAML cohort, focusing on the A-674562 compound. (**G**) Same as (**F**), but for the venetoclax drug.