Bayesian multi-source regression and monocyte-associated gene expression predict BCL-2 inhibitor resistance in acute myeloid leukemia

Brian S. White [ID][1,12,13✉], Suleiman A. Khan[2,13], Mike J. Mason[1], Muhammad Ammad-ud-din[2], Swapnil Potdar[2], Disha Malani [ID][2], Heikki Kuusanmäki[2,3], Brian J. Druker [ID][4,5], Caroline Heckman [ID][2], Olli Kallioniemi[2,6], Stephen E. Kurtz[5], Kimmo Porkka[7], Cristina E. Tognon [ID][4,5], Jeffrey W. Tyner [ID][5], Tero Aittokallio [ID][2,8,9,10,14], Krister Wennerberg[2,3,14] and Justin Guinney[1,11,14]

The FDA recently approved eight targeted therapies for acute myeloid leukemia (AML), including the BCL-2 inhibitor venetoclax. Maximizing efficacy of these treatments requires refining patient selection. To this end, we analyzed two recent AML studies profiling the gene expression and ex vivo drug response of primary patient samples. We find that ex vivo samples often exhibit a general sensitivity to (any) drug exposure, independent of drug target. We observe that this "general response across drugs" (GRD) is associated with *FLT3*-ITD mutations, clinical response to standard induction chemotherapy, and overall survival. Further, incorporating GRD into expression-based regression models trained on one of the studies improved their performance in predicting ex vivo response in the second study, thus signifying its relevance to precision oncology efforts. We find that venetoclax response is independent of GRD but instead show that it is linked to expression of monocyte-associated genes by developing and applying a multi-source Bayesian regression approach. The method shares information across studies to robustly identify biomarkers of drug response and is broadly applicable in integrative analyses.

INTRODUCTION

Acute myeloid leukemia (AML) is genetically, epigenetically, and transcriptionally heterogeneous. Nevertheless, patient treatment has been uniform and for decades standard first-line treatment has been "7 + 3" induction chemotherapy with cytarabine and an anthracycline. In addition, a significant portion of AML patients are not considered fit enough to tolerate induction chemotherapy and are instead typically treated with low-dose cytarabine (LDAC) and hypomethylating agents (HMAs), e.g., azacitidine or decitabine, that extend survival but are rarely curative. Together, outcomes for adult AML patients (aged ≥20 years) remain very poor, with the current 5-year relative survival rate at 24%[1]. The US Food and Drug Administration (FDA) has recently shifted this therapeutic landscape by approving eight drugs (enasidenib, gemtuzumab ozogamicin, glasdegib, ivosidenib, midostaurin, gilteritinib, venetoclax, and a liposome-encapsulated combination of daunorubicin and cytarabine) for newly diagnosed or relapse refractory patients, alone or in combination with LDAC or HMAs[2,3]. Maximizing clinical benefit from these treatments will require biomarkers for optimized patient selection and rational drug combination strategies.

We used data from two independent studies that comprehensively profiled AML patient cohorts to address this need. In both studies, ex vivo functional drug testing was performed on freshly isolated mononuclear cells from AML patients and cell viability was assessed following drug exposure across a concentration range. The multi-center Beat AML initiative, led by Oregon Health & Science University (OHSU), profiled 562 patient samples, including 411 with RNA sequencing (RNA-seq), 531 with exome sequencing, and 363 across a panel of 122 small-molecule inhibitors[4]. The AML Individualized Systems Medicine program at the Institute for Molecular Medicine Finland (FIMM) profiled 37 patients using RNA-seq, exome sequencing, and across a panel of 470 inhibitors[5]. We performed a comparative analysis, training regression models in one study to predict ex vivo drug response in the other. In both studies, we observed a tendency of a patient-derived sample to respond consistently across all drugs. This "general response across drugs" (GRD) correlated with complete response to standard induction therapy and with patient leukemia-free survival. Including GRD in the models significantly improved their performance.

Several drugs did not conform to this trend, including the BCL-2 inhibitor venetoclax[6], which was recently approved in combination with HMAs or LDAC for the treatment of AML in newly diagnosed elderly patients or in those unfit for intensive chemotherapy. Venetoclax in combination with azacitidine or decitabine showed a favorable overall response rate [complete

---

[1]Computational Oncology, Sage Bionetworks, Seattle, WA, USA. [2]Institute for Molecular Medicine Finland (FIMM), Helsinki Institute of Life Science (HiLIFE), University of Helsinki, Helsinki, Finland. [3]Biotech Research & Innovation Centre (BRIC) and Novo Nordisk Foundation Center for Stem Cell Biology (DanStem), University of Copenhagen, Copenhagen, Denmark. [4]Howard Hughes Medical Institute, Portland, OR, USA. [5]Division of Hematology and Medical Oncology, Knight Cancer Institute, Oregon Health & Science University, Portland, OR, USA. [6]Scilifelab, Karolinska Institute, Solna, Sweden. [7]HUS Comprehensive Cancer Center, Hematology Research Unit Helsinki and iCAN Digital Precision Cancer Center Medicine Flagship, University of Helsinki, Helsinki, Finland. [8]Department of Mathematics and Statistics, University of Turku, Turku, Finland. [9]Department of Cancer Genetics, Institute for Cancer Research, Oslo University Hospital, Oslo, Norway. [10]Centre for Biostatistics and Epidemiology (OCBE), University of Oslo, Oslo, Norway. [11]Department of Biomedical Informatics and Medical Education, University of Washington, Seattle, WA, USA. [12]Present address: The Jackson Laboratory for Genomic Medicine, Farmington, CT, USA. [13]These authors contributed equally: Brian S. White, Suleiman A. Khan. [14]These authors jointly supervised this work: Tero Aittokallio, Krister Wennerberg, Justin Guinney. ✉email: bstephenwhite@gmail.com
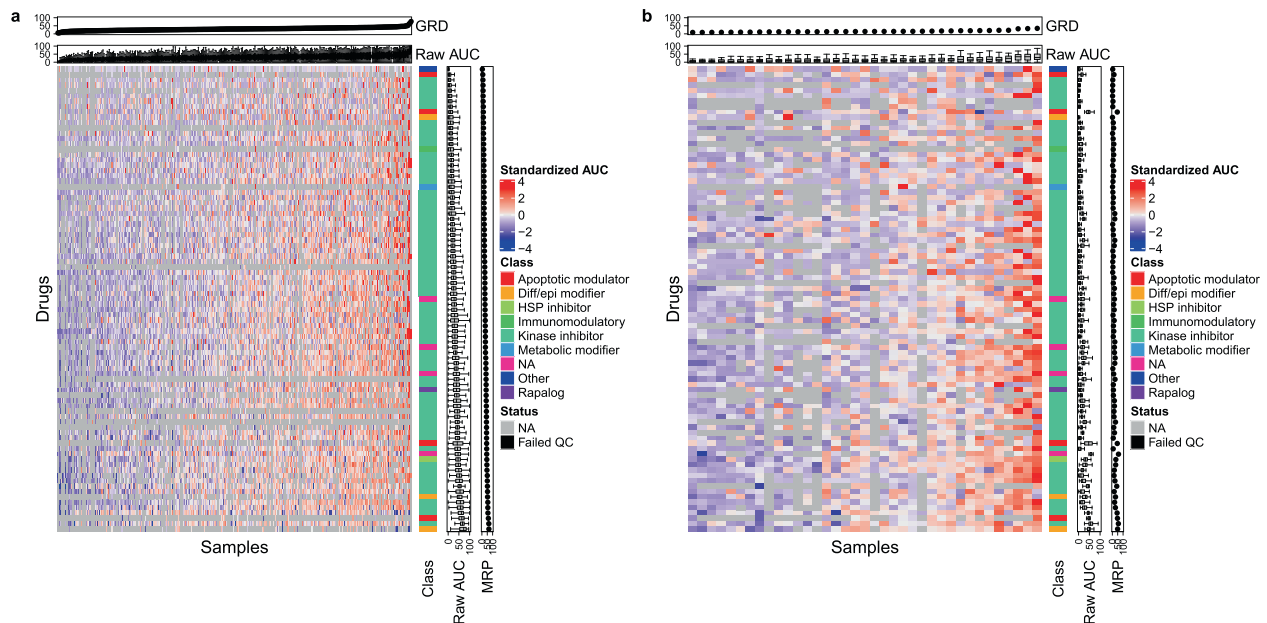
Fig. 1 A patient's ex vivo responses are similar across common drugs. AUCs calculated across patient-derived ex vivo samples (columns) and 87 drugs (rows) common to OHSU (**a**; $n = 338$) and FIMM (**b**; $n = 37$) datasets. Red values correspond to higher AUC or more sensitive samples, blue are less sensitive, black are filtered, and gray are missing. Standardized AUCs (i.e., with mean zero and standard deviation one across patients) displayed in heatmap. Raw AUCs displayed in top and side panels. General response across drugs (GRD) is mean of raw AUCs for an individual patient over drugs; mean response across patients (MRP) is mean of raw AUCs for an individual drug over patients. Samples ordered by GRD in each dataset. Drugs ordered by MRP in OHSU dataset. One sample displayed per patient, with sample assayed across the highest number of drugs displayed in cases with multiple samples per patient. Class drug class, Diff/epi differentiation/epigenetic, HSP heat shock protein.

remission (CR) + CR with incomplete count recovery (CRi)] of 67% in a phase 1b study of 145 patients aged ≥65 years with newly diagnosed AML[7], whereas an overall response rate of 28% has been observed for a similar patient population treated with azacitidine alone[8]. Nevertheless, a large minority of patients do not respond to venetoclax, while the majority who do ultimately relapse[7].

As such, recent efforts have attempted to refine patient selection for venetoclax treatment. An ex vivo study[9] suggested that patients with monocytic AML have reduced sensitivity to venetoclax. An in vivo study[10] additionally demonstrated that intra-patient heterogeneity arising from monocytic subclones contributes to venetoclax resistance. Here we show that the degree of patient monocyticity accounts for the majority of inter-patient variation in resistance. We quantify this effect through a robust signature comprised of monocyte-associated genes identified via a Bayesian multi-source regression (BMSR) method. BMSR nominates drug biomarkers by performing a joint multi-source analysis across the OHSU and FIMM datasets. The method is broadly applicable in integrating multiple expression datasets to overcome technical variation, biological heterogeneity, and small sample size that contribute to the low reproducibility of biomarker studies[11].

We extend BMSR to additionally perform simultaneous multi-task analysis across multiple drugs to identify their common biomarkers. Applying it to three mitogen-activated protein kinase kinase (MEK) inhibitors (trametinib, PD184352, and selumetinib) reveals that their response is correlated with the monocytic venetoclax resistance signature. This, coupled with the observation that venetoclax treatment selects for pre-existing monocytic subclones[10], provides independent rationale for combination therapies targeting the BCL-2 and MEK pathways[12].

RESULTS

GRD is associated with improved patient outcome

The 122 OHSU drug panel and the 470 FIMM drug panel shared 87 drugs (71 of which are kinase inhibitors). We used these common drugs to assess the consistency of drug response across the two studies. We first quantified response as area under the dose–response curve (AUC; Supplementary Figs. 1 and 2, Supplementary Table 1, and Fig. 1). Responses were positively correlated across drugs in each dataset [OHSU: 84% of Pearson correlation rs are positive, mean $r = 0.22$, 95% confidence interval (CI) $= -0.22$ to 0.59; FIMM: 85% positive r, mean $r = 0.33$, 95% CI $= -0.29$ to 0.83; Supplementary Fig. 3]. We measured the cross-study correlation of the intra-study drug–drug correlations, i.e., the "correlation of correlations," a general measure of interstudy consistency ($r = 0.35$, $p < 10^{-10}$; Supplementary Fig. 3). As expected, correlations were higher when restricted to drugs with a common target, including for 5 cyclin-dependent kinase inhibitors ($r = 0.76$), 5 epidermal growth factor receptor inhibitors ($r = 0.51$), 7 vascular endothelial growth factor receptor inhibitors ($r = 0.61$), 6 fms-like tyrosine kinase 3 (FLT3) inhibitors ($r = 0.84$), and 4 mammalian target of rapamycin/phosphoinositide 3-kinase inhibitors ($r = 0.95$). We also calculated each drug's mean response across patients and found these to be highly concordant between the datasets ($r = 0.67$; $p = 2.04 \times 10^{-12}$; Supplementary Fig. 4). Finally, we assessed the consistency of each individual drug's response across the two datasets. To do so, we represented each drug in each dataset by a vector of its correlations to all other drugs. For each drug, we then calculated the correlation between these dataset- and drug-specific correlation vectors (Supplementary Fig. 5). This cross-dataset drug correlation was positively associated with the drug's range of response [i.e., interquartile range (IQR) of unnormalized AUCs] in both the OHSU ($r = 0.45$; $p = 1.02 \times 10^{-5}$) and FIMM ($r = 0.47$; $p = 4.55 \times 10^{-6}$) datasets. We
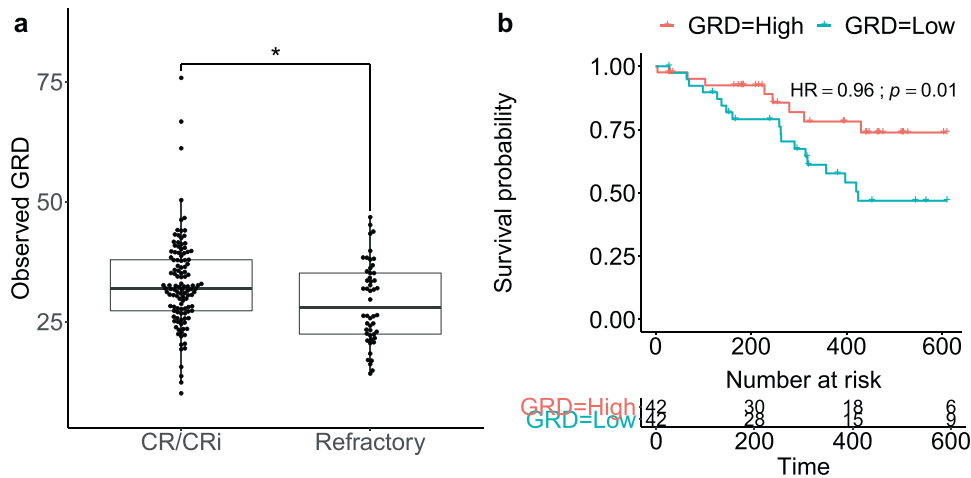
Fig. 2 Ex vivo general response across drugs is associated with clinical response and improved patient outcome. a GRD in patients who achieve complete remission (CR) or complete remission with incomplete hematologic recovery (CRi) to standard induction chemotherapy ($n = 118$) versus those refractory to induction ($n = 50$) in OHSU dataset. *Wilcoxon rank-sum test $p < 0.05$. **b** Kaplan–Meier survival curves of patients in OHSU dataset with GRD above the upper quartile (red; "responders"; $n = 42$) and of those with GRD below the lower quartile (blue; "non-responders"; $n = 42$). Data are right censored at 610 days. HR: Cox proportional hazard ratio. **a, b** GRD is computed across drugs common to OHSU and FIMM datasets.

found no evidence that drug correlation was associated with its class (analysis of variance $p = 0.37$).

Each patient sample tended to respond uniformly across all drugs in the panel—e.g., many responded relatively poorly to all drugs (blue; leftmost samples, Fig. 1), while a smaller number responded relatively strongly to most drugs (red; rightmost samples, Fig. 1). We describe this phenomenon as a sample's GRD, i.e., its mean AUC across all drugs. Notably, this trend held also across the full drug panels of each dataset (OHSU: 122 drugs; FIMM: 470 drugs; Supplementary Fig. 6), which are less biased toward tyrosine kinase inhibitors (TKIs) than the shared set of drugs. More specifically, we found that GRD computed across all drugs in a dataset was strongly correlated ($r > 0.9$) with that computed from the common set of drugs, the set of drugs that excludes class III TKIs (of which *FLT3* inhibitors are members), and the set of drugs that excludes all TKIs (Supplementary Fig. 7). Consistently, we found that GRD calculated from a random selection of drugs was highly concordant with GRD calculated from all drugs in the respective dataset for both the OHSU (mean r across 100 bootstrap samples $= 0.97$; 95% CI $= 0.95–0.98$) and FIMM (mean $r = 0.99$; 95% CI $= 0.98–1.00$) datasets. An observation similar to GRD, General Levels of Drug Sensitivity (GLDS), has previously been reported across cell lines representing various cancer types[13].

GRD was higher in samples from patients who achieved a CR or a CRi to standard induction chemotherapy relative to those refractory to induction (two-sided Wilcoxon rank-sum test $p = 0.02$; Fig. 2a; GRD computed across common drugs). Consistently, CR/CRi patients were enriched among those with high GRD (enrichment $p = 7.5 \times 10^{-3}$; Supplementary Fig. 8). Notably, patients in the top quartile of GRD showed improved overall survival relative to those in the bottom quartile [Cox proportional hazard ratio (HR) $= 0.96$, log-rank $p = 0.01$; Fig. 2b and Supplementary Fig. 9]. These findings held when GRD was instead computed across all drugs (two-sided Wilcoxon rank-sum test $p = 8.2 \times 10^{-4}$; HR $= 0.95$; log-rank $p = 0.01$; Supplementary Fig. 10).

Despite these associations, refractory patients were not statistically enriched in extreme GRD values (enrichment $p = 0.15$; Supplementary Fig. 8). To understand this heterogeneity in response, we examined clinical features, genes, and pathways that differentiated refractory from CR/CRi patients among those with low (bottom quartile) GRD. Increased age was associated with refractory response among GRD-low patients (two-sided Wilcoxon

rank-sum test $p = 9.0 \times 10^{-3}$), though the trend was not significant after correcting for testing of the multiple clinical variables [Benjamini–Hochberg (BH)-adjusted $p = 0.26$; Supplementary Table 2]. One hundred and twenty-two genes were differentially expressed (false discovery rate (FDR) $< 20\%$; Supplementary Tables 3 and 4), with lymphocyte costimulation, T cell receptor signaling, antigen binding, and antigen receptor-mediated signaling (all significant at an FDR $< 20\%$) upregulated in refractory patients and among the strongest enrichments (Supplementary Table 5). Conversely, in examining GRD-high patients, we found that higher creatinine levels trended with CR/CRi (two-sided Wilcoxon rank-sum test $p = 0.02$; BH-adjusted $p = 0.62$; Supplementary Table 6). Forty-six differentially expressed genes (FDR $< 20\%$; Supplementary Table 7) were most strongly enriched within extracellular pathways upregulated in CR/CRi patients (Supplementary Table 8).

GRD is associated with *FLT3*-ITD

Presence of internal tandem duplication in *FLT3* (*FLT3*-ITD) was positively associated with GRD in the OHSU dataset (BH-adjusted two-sided Wilcoxon rank-sum test $p = 6.86 \times 10^{-8}$; Supplementary Table 9). This positive association remained significant and independent of *NPM1* mutation status, ethnicity, age, and sex in a multivariate analysis [$R^2 = 0.26$; F-statistic $p = 6.97 \times 10^{-10}$; two-sided t test $p = 4.95 \times 10^{-8}$; Supplementary Fig. 11 and Supplementary Table 10]. We validated the *FLT3*-ITD/GRD association in an independent AML dataset profiling expression and drug response of ex vivo samples published by Tavor and colleagues[14] (one-sided Wilcoxon rank-sum test $p < 0.01$; Supplementary Fig. 12). We also observed a consistent trend in the FIMM dataset (one-sided Wilcoxon rank-sum test $p = 0.08$). Significantly, we found that *FLT3*-ITD status remained associated with GRD even when the latter was computed from a subset of drugs that excluded *FLT3* inhibitors (Supplementary Fig. 13).

To determine whether GRD could be modeled using gene expression data, we trained an expression-based ridge regression model of GRD using the OHSU dataset (Supplementary Figs. 14–19; Supplementary Table 3; see "Methods"). The model was validated in the FIMM dataset, demonstrating good predictive performance ($r = 0.67$; $p = 5.6 \times 10^{-6}$; Fig. 3a). We hypothesized that robust biomarkers of GRD should be consistent between the OHSU and FIMM datasets (e.g., having large positive model coefficients in both datasets relative to other genes). To test this,
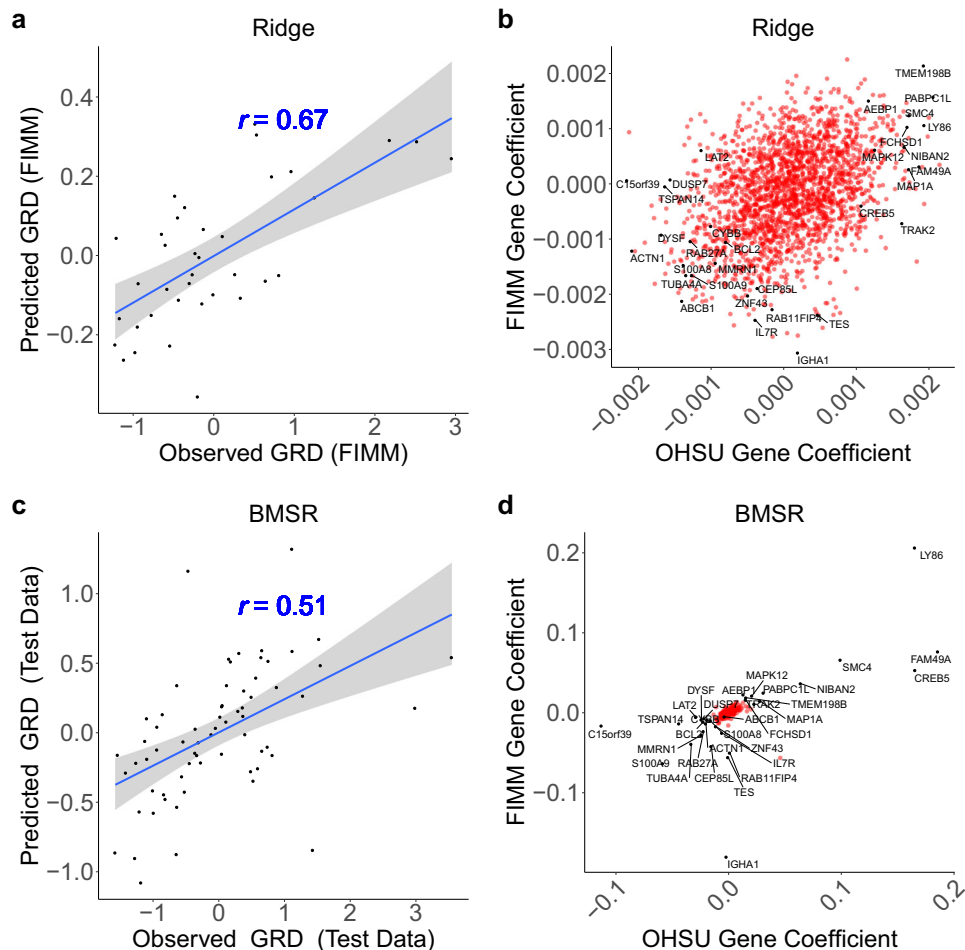
Fig. 3 Integrative analysis reveals growth-, apoptosis-, and drug efflux-associated biomarkers of general response across drugs. **a**, **c** Observed (*x* axis) versus model-predicted (*y* axis) GRD. **a** Expression-based ridge regression model trained on OHSU samples ($n = 292$) and tested on FIMM samples ($n = 37$). **c** Expression-based Bayesian multi-source regression (BMSR) model trained using fivefold cross-validation on combined OHSU and FIMM datasets ($n = 263$) and tested on held-out fold yielding median performance across the fivefolds ($n = 66$). **b**, **d** Coefficients of genes ($n = 2132$) in OHSU (*x* axis) or FIMM (*y* axis) datasets following **b** training of ridge regression model independently on both datasets or **d** training of BMSR model simultaneously on entirety of both datasets ($n = 329$). GRD is computed across drugs common to OHSU and FIMM datasets. *r*: Pearson correlation; dashed line: identity line; blue line: linear regression fit; gray shading: 95% confidence interval. Labeled genes were those having extremal (Stouffer's $p < 0.01$) combined coefficients across both datasets, as well as *ABCB1* (i.e., *MDR1*).

we trained a GRD model on the FIMM dataset and compared the model coefficients associated with each gene between the OHSU- and FIMM-trained models. Unexpectedly, this did not reveal candidate biomarkers with outlying coefficients in both datasets (Fig. 3b). Nevertheless, we did confirm that the *ABCB1* gene [i.e., Multidrug Resistance Protein 1 (*MDR1*)], encoding a drug efflux pump and previously observed to be associated with GLDS[13], was negatively correlated with GRD in both datasets (Fig. 3b). Further, the ABC transporter family was among the gene sets and pathways having the strongest negative association with GRD (Supplementary Tables 11–15).

BMSR identifies biomarkers jointly across studies
We reasoned that our inability to detect consistent biomarkers through independent analysis of the two datasets resulted from the highly correlated expression of subsets of genes (e.g., within a pathway), thereby dampening the effect of any single gene and hindering the identification of genes significantly impacting response. This is further complicated by study site-specific technical artifacts, biological variation, and small sample sizes, all of which compound noise. These factors could be ameliorated by

performing regularization across multiple datasets simultaneously, which would shrink the coefficients of all but one or a few of the highly correlated genes toward zero.

To do so, we developed an integrative BMSR method that detects consistent and robust biomarkers through joint analysis of the two datasets. BMSR assumes that the biomarker expression/drug response relationship is similar across multiple datasets but allows for relatively small differences due to technical noise, biological variation, or clinical heterogeneity of different patient populations. It achieves this by modeling the contribution to the response by gene *g* in dataset *d* (i.e., the regression coefficient $\beta_g^{(d)}$) as arising from a mean contribution (i.e., β_g) for that gene that is shared across datasets (see "Methods"). BMSR effectively performs regularization on the shared β_g rather than the dataset-specific coefficient $\beta_g^{(d)}$ through a prior distribution, which acts as a penalty term in a non-Bayesian/frequentist formulation. For two genes g_1 and g_2 that are correlated with response, this approach encourages (but does not guarantee) that the shared coefficient of one of them be shrunk to zero by exploiting correlations in the data rather than explicit biological annotations (e.g., pathways).
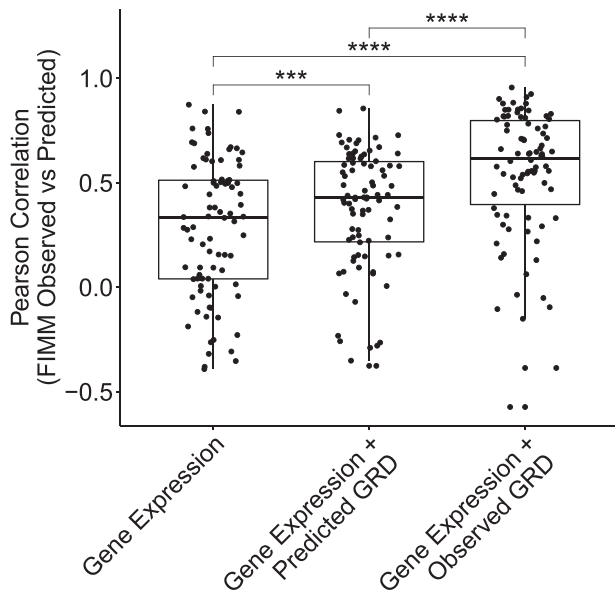
Fig. 4 Expression-based predictions of drug response indicate concordance of independent ex vivo datasets and may be improved by incorporating general response across drugs. Performance (Pearson correlation between observed and model-predicted drug response; y axis) of ridge regression models trained on OHSU data and tested on FIMM data using genes as predictors (Gene Expression), genes and GRD predicted by applying ridge regression to gene expression (Gene Expression + Predicted GRD), or genes and GRD calculated from drug response data (Gene Expression + Observed GRD). Each point corresponds to a drug (n = 87). Drug d is excluded from observed and predicted GRD in modeling d's response. ****One-sided paired Wilcoxon signed rank p < 0.0001; ***p < 0.001.

Modeling GRD with the BMSR approach (Fig. 3c) revealed candidate biomarkers (Fig. 3d and Supplementary Fig. 20) that were distinctly separated from the bulk of non-contributing genes (i.e., those with coefficients near zero in both datasets). Candidate biomarkers included genes involved in cell proliferation [IL7R[15] and NIBAN2[16]], cell cycle [MAPK12[17]], cell growth [BCL2[18] and S100A8/S100A9[19]], apoptosis [BCL2[20], NIBAN2[21], S100A8/S100A9[22]], and drug response [BCL2[23]]). Notably, these genes were near the periphery of the ridge coefficient distribution (Fig. 3b), thus demonstrating the consistency of the two methods. As expected, candidate GRD biomarkers were also consistently correlated with response to *individual* drugs (Supplementary Fig. 21).

Drug response is robustly predicted by gene expression

Responses could be significantly predicted (p < 0.01) using expression-based ridge regression for 31 of the 87 drugs (Fig. 4 and Supplementary Table 16; median r = 0.33; 25th–75th percentile = 0.04–0.51; all significant correlations were positive; see "Methods"). Significantly predicted drugs included the heat shock protein inhibitor tanespimycin, the immunomodulatory agent lenalidomide, the bromodomain inhibitor JQ1, two apoptotic modulators (nutlin-3 and venetoclax), and 26 kinase inhibitors. Prediction performance was correlated with the spread of response (i.e., IQR of unnormalized AUCs) in the FIMM validation dataset (Supplementary Fig. 22; r = 0.36; p = 6.19 × 10⁻⁴). These results reinforce the consistency of the two ex vivo studies demonstrated by the drug–drug correlations above.

Using predicted GRD as a model variable in addition to gene expression improved response-modeling performance (median r = 0.43; 25th–75th percentile = 0.17–0.60) relative to modeling based on gene expression variables alone (one-sided paired

Wilcoxon signed rank p = 3.45 × 10⁻⁴; Fig. 4). Using observed, rather than predicted, GRD as a variable in addition to gene expression further improved performance (median r = 0.63; 25th–75th percentile = 0.39–0.75) relative to using gene expression alone (one-sided paired Wilcoxon signed rank p = 4.71 × 10⁻¹¹; Fig. 4).

These results were largely independent of whether GRD was computed across common drugs or all drugs in a dataset (correlation r between common drug and all drug-based models of observed versus predicted drug response correlations ≥0.96; Supplementary Figs. 23 and 24). This was true despite decreased performance in predicting GRD across all drugs (r = 0.52; Supplementary Fig. 25) relative to across common drugs (r = 0.67; Fig. 3).

A previous study proposed a filtering strategy to combat experimental and technical noise anticipated in large-scale drug screens[24]. To ensure that our findings were robust to such noise, we developed and applied a related outlier-removal approach (Supplementary Figs. 26–31; see "Methods"). Since prediction performance was not significantly different after removing outliers (Supplementary Fig. 32; two-sided Wilcoxon rank-sum test p = 0.89; Supplementary Tables 17 and 18), subsequent analysis does not exclude outliers.

Monocytic signature predicts venetoclax resistance

Having determined potential biomarkers of GRD above, we next asked whether responses of individual drugs were driven solely by GRD. To isolate drug-specific responses from general effects, we compared prediction performance of models using both gene expression and GRD as variables to that of models using only GRD (i.e., effectively drug–GRD correlations). Drugs showing the greatest specificity were the MEK1/2 inhibitors trametinib, PD184352, and selumetinib and the BCL-2 inhibitor venetoclax (Supplementary Fig. 33). Of these, gene-based prediction performance was highest for venetoclax (r = 0.84; p = 7.50 × 10⁻⁸; Supplementary Fig. 34A).

BMSR analysis revealed that venetoclax response had a positive association with the gene BCL2 encoding the drug target, as well as strong negative correlations with CD14 and SLC15A3 (Fig. 5 and Supplementary Figs. 34 and 35). SLC15A3 (Solute Carrier Family 15 Member 3) is highly expressed in monocytes at the protein[25] and mRNA[26,27] levels, while CD14 encodes a canonical (classical) monocyte cell surface marker[28]. As such, BMSR analysis indicated that monocyte-associated genes are correlated with venetoclax resistance. We confirmed that genes having expression correlated with venetoclax resistance are enriched for a monocyte signature (p = 2.0 × 10⁻⁴; Supplementary Fig. 36 and Supplementary Table 19). Both results are consistent with findings from Kuusanmäki and colleagues that the myeloid differentiation stage of AML cells impacts venetoclax response, with monocytic cells exhibiting resistance to BCL-2 inhibition[9].

As with GRD above, these candidate biomarkers contribute consistently (i.e., with the same sign) to the ridge regression model (Supplementary Fig. 34), but BMSR prioritizes a smaller set of features with large coefficients. As a comparison, we also applied minimum redundancy maximum relevance (mRMR), a method that performs feature selection within a single dataset to identify parsimonious feature sets[29,30]. mRMR identified a subset of the BMSR-prioritized, monocyte-associated genes across both datasets, but the individual genes selected differed. In particular, across 200 random (bootstrap) samples of the datasets, mRMR selected SLC15A3 across all random samples in the OHSU dataset but only once in the FIMM dataset and, conversely, selected PSAP across all random samples in the FIMM dataset but only once in the OHSU dataset. mRMR identified all BMSR-prioritized, monocyte-associated genes across one or more random samples; however, no gene was selected more than twice across both
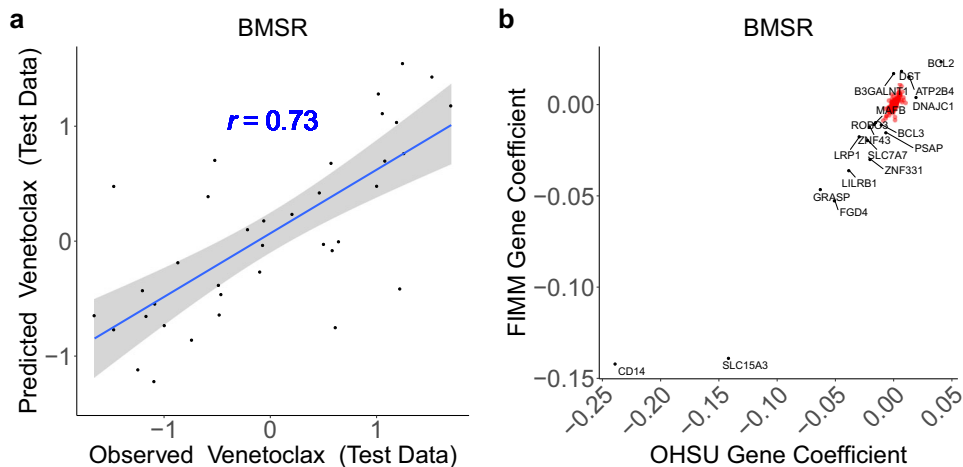
Fig. 5 Integrative analysis reveals monocyte-associated biomarkers predictive of venetoclax resistance. a Observed (*x* axis) versus BMSR-predicted (*y* axis) venetoclax response. Expression-based Bayesian regression model trained using fivefold cross-validation on combined OHSU and FIMM datasets ($n = 159$) and tested on held-out fold yielding median performance across the fivefolds ($n = 37$). **b** Coefficients of genes ($n = 2132$) in OHSU (*x* axis) or FIMM (*y* axis) datasets following training of Bayesian regression modeling simultaneously on both datasets ($n = 196$). *r*: Pearson correlation; dashed line: identity line; blue line: linear regression fit; gray shading: 95% confidence interval. Labeled genes were those having extremal (Stouffer's $p < 0.01$) combined coefficients across both datasets.

datasets (Supplementary Table 20). Collectively, these results demonstrate the biological consistency of the BMSR, ridge regression, and mRMR analyses, while highlighting BMSR's motivating contribution of identifying a small set of features concordantly across datasets.

Intriguingly, genes associated with venetoclax resistance were also enriched for T and/or B cell-mediated pathways (Supplementary Tables 21–25). Further, we observed that venetoclax response was positively associated with blast percentage of both peripheral blood ($r = 0.37$; two-sided Wilcoxon rank-sum test $p = 4.6 \times 10^{-4}$) and bone marrow ($r = 0.43$; $p = 5.9 \times 10^{-5}$; Supplementary Fig. 37). Nevertheless, several lines of evidence suggest that our results are unlikely to be compromised by impure leukemic samples and/or lymphocyte contamination. First, predicted levels of lymphocytes are lower and less variable than those of monocytes in both datasets (Supplementary Fig. 38). Second, genes associated with venetoclax resistance are more enriched for markers of monocytes and other cell types within the monocytic lineage (i.e., macrophages) than for markers of lymphocytes and are also strongly enriched in myeloid dendritic cells (Supplementary Table 19). Finally, we controlled for potential confounding effects by including lymphocyte levels as covariates in our ridge regression models. Our findings were consistent with the original ridge regression models: prediction results across drugs were highly correlated between the two models ($r > 0.99$; Supplementary Fig. 39) and genes contributing most to the lymphocyte-controlled model continued to show strong enrichments for markers of the monocytic lineage (Supplementary Table 26).
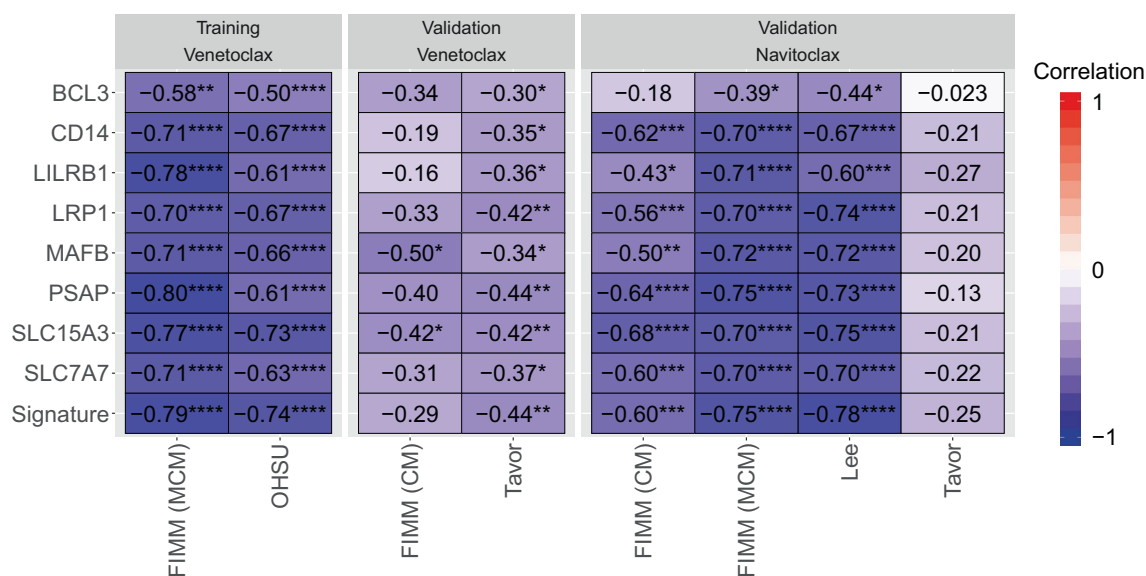
Because technical or biological variation may contribute noise to individual genes, we next sought to develop a robust signature of venetoclax resistance that would mitigate these fluctuations. To do so, we focused attention on monocyte-associated genes prioritized by BMSR (see "Methods"), which, in addition to *CD14* and *SLC15A3*, included *BCL3*, *LILRB1*, *LRP1*, *MAFB*, *PSAP*, and *SLC7A7* (Fig. 5b). We compressed the expression of these genes into a single enrichment-based signature of venetoclax resistance (see "Methods").

We sought to validate this signature (and its constituent genes) across several independent conditions, drugs, and/or ex vivo functional and transcriptomic profiles of AML (Fig. 6), including a study by Lee and colleagues of drug sensitivity that profiled the BCL-2/BCL-XL inhibitor navitoclax[31], the ex vivo study by Tavor

and colleagues that profiled both venetoclax and navitoclax[14], and a second FIMM dataset that profiled venetoclax and navitoclax in a stroma-derived conditioned medium (CM) that differed from the mononuclear cell medium (MCM) of the above FIMM dataset[32]. The signature was inversely correlated with venetoclax response (i.e., correlated with resistance) in the Tavor dataset ($r = -0.44$; $p = 3.2 \times 10^{-3}$). It also trended with resistance in the FIMM CM dataset ($r = -0.29$; $p = 0.17$), with a lack of significance consistent with the specific dampening of venetoclax response in CM relative to MCM[32]. Despite this, the signature smoothed out the weak anti-correlations of several genes (e.g., *CD14* and *LILRB1*) as intended. The monocytic signature was strongly anti-correlated with navitoclax response in three datasets: FIMM (CM) ($r = -0.60$; $p = 2.22 \times 10^{-4}$), FIMM (MCM) ($r = -0.75$; $p = 1.06 \times 10^{-7}$), and Lee ($r = -0.78$; $p = 2.83 \times 10^{-7}$). All genes in the signature, with the exception of *BCL3*, also validated against navitoclax in these three datasets, though their correlation was often not as strong as that of the signature itself. The signature trended with navitoclax resistance in the Tavor dataset ($r = -0.25$; $p = 0.11$), with *BCL3* again having the weakest association.

As a further validation of the robustness of the signature and of the generality of our Bayesian approach, we used BMSR to jointly analyze the Tavor and FIMM (CM) datasets. Following the approach above, we defined a signature from the monocyte-associated genes prioritized by the analysis. This revised signature was highly correlated with the original FIMM/OHSU-derived signature in the FIMM, FIMM (CM), OHSU, and Tavor datasets ($r \geq 0.75$; Supplementary Fig. 40). Finally, to demonstrate its applicability to more than two datasets, we similarly applied BMSR to three datasets, FIMM, OHSU, and Tavor, and again observed that the resulting signature was highly correlated with the original FIMM/OHSU-derived signature ($r \geq 0.95$; Supplementary Fig. 41).

Several biomarkers of venetoclax response in AML have been proposed at the protein [BCL-XL, MCL-1, and BCL-2[33]] and mRNA [*BCL2*, *BCL2/MCL1* ratio, *BCL2A1*, *CD11b*, *CD14*, *CD68*, *CD86*, *CLEC7A* (*CD369*), *HOX* gene family members, *MCL1*, *S100A8*, and *S100A9*[9,34–37]] levels. Of these, the monocytic signature, *BCL2A1*, *CD68*, *CD86*, and *CLEC7A* were robust predictors across both BCL-2-inhibitors (venetoclax or navitoclax), different cell culture conditions, and datasets, with none performing best across all conditions (Supplementary Fig. 42). Strikingly, all five were

	Training Venetoclax		Validation Venetoclax		Validation Navitoclax			
BCL3	−0.58**	−0.50****	−0.34	−0.30*	−0.18	−0.39*	−0.44*	−0.023
CD14	−0.71****	−0.67****	−0.19	−0.35*	−0.62***	−0.70****	−0.67****	−0.21
LILRB1	−0.78****	−0.61****	−0.16	−0.36*	−0.43*	−0.71****	−0.60***	−0.27
LRP1	−0.70****	−0.67****	−0.33	−0.42**	−0.56***	−0.70****	−0.74****	−0.21
MAFB	−0.71****	−0.66****	−0.50*	−0.34*	−0.50**	−0.72****	−0.72****	−0.20
PSAP	−0.80****	−0.61****	−0.40	−0.44**	−0.64****	−0.75****	−0.73****	−0.13
SLC15A3	−0.77****	−0.73****	−0.42*	−0.42**	−0.68****	−0.70****	−0.75****	−0.21
SLC7A7	−0.71****	−0.63****	−0.31	−0.37*	−0.60***	−0.70****	−0.70****	−0.22
Signature	−0.79****	−0.74****	−0.29	−0.44**	−0.60***	−0.75****	−0.78****	−0.25
	FIMM (MCM)	OHSU	FIMM (CM)	Tavor	FIMM (CM)	FIMM (MCM)	Lee	Tavor

Fig. 6 A monocyte expression signature robustly predicts resistance to BCL-2 inhibitors. Pearson correlation of response of the indicated drug (venetoclax or navitoclax; top) versus expression of the indicated gene or venetoclax monocyte signature (Signature) across FIMM (MCM), OHSU, FIMM (CM), Lee, or Tavor datasets. Dataset/drug combinations are indicated as "Training" if they were used to derive the signature and biomarkers and "Validation" otherwise. ****$p < 0.0001$; ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.

strongly correlated with one another (Supplementary Fig. 43) and are highly expressed in monocytes[27,38]. These findings validate our own monocytic signature and previously proposed, monocyte-associated biomarkers as predictors of resistance to BCL-2 inhibition in AML.

Monocytic signature predicts MEK inhibitor response
Similar to venetoclax, gene expression data improved the prediction performance of the three MEK inhibitors, trametinib, PD184352, and selumetinib, beyond that provided by GRD alone (Supplementary Fig. 33). We therefore next sought to determine biomarkers of their response. As the response of the MEK inhibitors are strongly correlated with one another ($r > 0.48$; Supplementary Fig. 44), we developed a Bayesian multi-source multi-task regression (BMSMTR) approach that simultaneously analyzes multiple datasets (i.e., multi-source, as above) and also multiple drugs (i.e., multi-task; see "Methods"). Joint analysis of trametinib, PD184352, and selumetinib with BMSMTR identified monocytic genes as candidate biomarkers, including *LRP1* and *CD300E* (Supplementary Fig. 45), that are *positively* associated with MEK inhibitor response. Hence, we hypothesized that the monocytic signature defined above to be correlated with venetoclax *resistance* would be correlated with MEK inhibitor *response*. Indeed, we observed a positive correlation between the response of each of the three MEK inhibitors and the monocytic signature in both the OHSU and FIMM datasets, where BCL-2 inhibitor response is inversely correlated with MEK inhibitor response (Supplementary Fig. 44). The association did not hold in the FIMM CM dataset, where BCL-2 inhibitor response is not significantly (inversely) (Supplementary Fig. 44) correlated with MEK inhibitor response and where our hypothesis would not be expected to hold.

DISCUSSION
We demonstrated consistency between two large-scale AML studies that profiled ex vivo drug sensitivity and gene expression. In both datasets, we observed that patient-derived samples exhibited a GRD, i.e., a sample often responded uniformly across drugs independent of target or mechanism of action—relatively

strongly to all drugs or relatively weakly to all drugs. GRD was further associated with clinical endpoints. Finally, we developed a BMSR method for biomarker discovery and applied it to reveal a robust monocytic signature of BCL-2 inhibition in AML.

We demonstrated that the two studies were consistent: first, drug–drug correlations were conserved across studies (Supplementary Fig. 3); second, the mean response of each drug across patients relative to that of the other drugs was also conserved across studies (Supplementary Fig. 4); and, finally, regression models trained on gene expression data in one study predicted response in the second, independent study for 31 of the 87 drugs (Fig. 4). Similar efforts comparing[39] in vitro drug screens[40,41] reported discordance in drug responses across the datasets. This finding spurred considerable activity in the research community[24,42–47], which ultimately resolved certain discrepancies between the Cancer Cell Line Encyclopedia (CCLE) and the Genomics of Drug Sensitivity in Cancer (GDSC) studies, in part, by harmonizing curve fitting methods applied across datasets and by quantifying drug response via a modified AUC[44,48]. We also employed AUC since it both intuitively and empirically (Supplementary Figs. 1 and 2) provides a more robust summary of a multi-parameter curve fit than does a single parameter such as EC_{50} or IC_{50}. A retrospective re-analysis of these in vitro studies applied quality control filters to exclude curves that grossly violated the assumptions that sensitivities range between 0 and 100% and that they increase monotonically with drug concentration[24]. We developed and applied related quality control measures (Supplementary Figs. 26–31) and found they had little impact on prediction performance of most drugs (Supplementary Fig. 32). Nevertheless, even low-frequency technical noise is expected to be observed in large-scale studies, and hence, we concur with Safikhani and colleagues[24] that its impact should be carefully investigated, as we have done here.

Geeleher and colleagues[13] reported a phenomenon similar to GRD—GLDS—across cell lines spanning cancer types[40,41,49,50]. They showed that conditioning on GLDS eliminated spurious biomarker predictions while identifying evidence-supported biomarkers that otherwise went undetected. We similarly showed that including GRD in prediction models improved overall accuracy (Fig. 4), thus demonstrating that the general trend detected in vitro is active in patient-derived AML samples as well.

Thus, precision oncology studies should account for both target-specific and target-agnostic effects when correlating drug response with biomarkers.

Our findings directly relate this trend to clinical endpoints: increased GRD is associated with complete response to induction therapy and to improved overall survival (Fig. 2). This generalizes the previous observation that ex vivo response to individual drugs may be correlated with AML remission status[31]. Further, we found that patients with *FLT3*-ITD mutations have higher GRD. This result held even when GRD was calculated from subsets of drugs that excluded *FLT3* inhibitors. Relatedly, Tavor and colleagues found that patient-derived samples with high sensitivity across TKIs and several other drugs were enriched in *FLT3*-ITD mutations relative to resistant samples[14]. Collectively, these results suggest that *FLT3*-ITD mutations may confer a generalized drug sensitivity. One such mechanism for doing so may be through its association with reduced levels of the ABCB1 drug efflux pump[51]. Indeed, our expression-based analysis revealed that *ABCB1* expression was inversely correlated with GRD. Nevertheless, prior results relating *ABCB1* expression to ex vivo response to agents used in induction therapy have been inconsistent[51–53]. Additionally, the modest correlation of candidate biomarkers with GRD (Supplementary Fig. 20), particularly relative to the consistent correlation observed for venetoclax biomarkers (Supplementary Fig. 35), suggests that mechanisms underlying GRD may involve a complex interplay of multiple genes, possibly conditioned on *FLT3* status.

We developed BMSR to mitigate factors that limit reproducibility in prioritizing biomarkers in high-dimensional gene feature spaces, including small sample sizes, correlated expression of functionally related genes, technical variation across datasets, and heterogeneity of patient populations[54–57]. It does so by performing integrated analysis[11,58] across multiple heterogeneous datasets to increase cumulative sample size, both of which reduce the likelihood of overfitting. BMSR shares information across the datasets: gene expression values (or, more formally, the gene expression coefficients) in each dataset are modeled as arising from a shared prior distribution. BMSR differs from meta-analysis methods that first analyze datasets independently before combining effect sizes or *p* values across datasets. By regularizing the prior's hyperparameters, BMSR effectively selects features simultaneously across the two datasets to prioritize a sparse set of biomarkers. This is manifested in a small number of genes having coefficients that are well separated from the majority of gene coefficients in both datasets, in contrast to independent ridge regression analysis across the two datasets in which coefficients are evenly distributed with no clear separation indicating candidate biomarkers (Fig. 3 and Supplementary Fig. 34).

We demonstrated BMSR's generalizability by jointly analyzing different pairs of datasets [FIMM and OHSU; FIMM (CM) and Tavor] and by analyzing a dataset trio (FIMM, OHSU, and Tavor; Supplementary Figs. 40 and 41). As such, we anticipate BMSR will have broad applicability in integrative biomarker studies. However, such analyses are often plagued by the effort required to harmonize multiple datasets. To address this issue, we collaborated with the developers of ORCESTRA in the Haibe-Kains laboratory to make several of the datasets (the Tavor and OHSU datasets) more broadly and easily accessible[59]. ORCESTRA is a cloud-based platform that provides automated processing of pharmacogenomic profiles and packages them into a fully documented and DOI-indexed "PharmacoSet" (PSet). These PSets are compatible with PharmacoGx, an open-source computational framework that facilitates integrative studies of multiple pharmacogenomic datasets through routines for standardized access and analysis[60]. To provide a template for others in applying BMSR for integrative analysis, we included a demo in our BMSR GitHub repository (https://github.com/suleimank/bmsr) that downloads the Tavor (https://doi.org/10.5281/zenodo.4585705) and OHSU (https://doi.org/10.5281/zenodo.4582786) datasets from ORCESTRA and uses

the functions of PharmacoGx to predict biomarkers of venetoclax response. Through our ongoing collaboration, we will also make these datasets available through PharmacoDB (https://pharmacodb.ca), a web application that similarly assembles pharmacogenomic datasets into a single database for cross-study analyses[61]. The online, interactive capabilities of PharmacoDB will make these resources accessible beyond computational biologists and those with programming skills.

BMSR revealed that venetoclax resistance is correlated with expression of monocyte-associated genes (Fig. 5). We combined these genes into a single signature that is robust to the variations of its individual genes[62] across datasets (Fig. 6). Additionally, we showed that the venetoclax-derived signature strongly predicts resistance to the BCL-2/BCL-XL inhibitor navitoclax. Resistance to both venetoclax and navitoclax was significantly correlated with the signature in standard culture conditions (MCM), though the trends in a stroma-derived CM reached significance only for navitoclax and not for venetoclax. This may have resulted from the reduced sensitivity of AML cells to venetoclax observed in CM relative to MCM, mediated by a switch from BCL-2- to BCL-XL-dependent cell survival that has a less pronounced effect on navitoclax[32].

Our findings extend and support recent reports demonstrating that monocytic cells are resistant to BCL-2 inhibition, whereas myeloid progenitors exhibit sensitivity. Kuusanmäki and colleagues showed that differentiated cells from AML samples expressing monocytic markers are less sensitive to venetoclax than immature blasts[9]. They further showed that sensitivity to venetoclax decreased along the differentiation spectrum, from less differentiated AML samples [French-American-British subtype M1] to more differentiated samples with significant monocytic differentiation (M5 subtype). They concluded by associating increased expression of monocyte markers *CD14*, *CD11b*, *CD86*, and *CD68* with decreased ex vivo responsiveness to venetoclax. Among other identified biomarkers, Zhang and colleagues also implicated high expression of *CD14*, as well as the monocyte-associated *CLEC7A* (*CD369*), with reduced sensitivity to venetoclax[36]. We provide independent validation of these biomarkers, confirming that leukemic cells expressing high levels of *CD68*, *CD86*, and *CLEC7A*, in particular, are more resistant to BCL-2 inhibition via both venetoclax and navitoclax (Supplementary Fig. 42). Further, we identified a mostly non-overlapping, but strongly correlated (Supplementary Fig. 43), set of genes highly expressed in monocytes: *BCL3*, *CD14*, *LILRB1*, *LRP1*, *MAFB*, *PSAP*, *SLC15A3*, and *SLC7A7*. We combined these genes into a per-sample monocytic score using a straightforward approach that did not require additional training or parameter turning. This signature was competitive with *CD68*, *CD86*, and *CLEC7A* across BCL-2 inhibitors, culture conditions, and datasets. Since neither the signature nor the genes nominated by these two studies outperformed the others across all conditions, it may be beneficial to combine them for increased robustness, much as the signature itself smoothed out variation in its constituent genes. Such an approach should be validated in the future using independent data.

These results are likely to be of clinical relevance, as Pei and colleagues demonstrated that phenotypically primitive AML is sensitive to venetoclax in vivo, whereas monocytic AML is more resistant[10]. Further, they showed that venetoclax (in combination with azacitidine) selects for monocytic subclones present at diagnosis. This resistance is associated with a loss of BCL-2 expression and a concomitant shift to MCL1 for survival that is inherent in monocytic differentiation. MCL1 is itself stabilized by extracellular signal-regulated kinase (ERK)[63,64]. As such, targeting ERK via a MEK inhibitor has been shown to sensitize cells to the BCL-2/BCL-XL inhibitor ABT-737 in vitro and in xenograft models[65]. Targeting MCL1 has also been shown to forestall acquired resistance to venetoclax in vitro[66], while leukemic blasts resistant to venetoclax are sensitive to MEK inhibition via trametinib

ex vivo[9]. Finally, combination of venetoclax with the MEK inhibitor cobimetinib exhibited synergy in vitro, inhibited growth ex vivo, and reduced leukemia burden in xenografts[12]. Our work has implications for the rational selection of patient groups for combination therapies targeting these two pathways. Since the monocytic signature is correlated with BCL-2 inhibitor resistance and MEK inhibitor response, it may provide a means of prospectively identifying patients for combination treatment.

METHODS

Drug–response curve fitting and filtering

We fit 3- (LL3) and 4-parameter log-logistic (LL4) curves to the dose–response data using `PharmacoGx`[60] and `drc`[67] in R[68], respectively. We excluded non-AML patients or those exhibiting gross dissimilarities across replicates from analysis. We excluded any drug–sample pair having a concentration range outside the most common (dataset-specific) concentration range for that corresponding drug. We further excluded a drug–sample pair if it did not include all concentration points and only analyzed one sample per drug–patient pair (see Supplementary Methods). Additionally, we assessed the impact of an outlier-removal strategy that excluded drug–sample pairs: (1) whose fits were not monotonically increasing, (2) that had large differences between fits that did (LL3) and did not (LL4) constrain the curve to asymptote to zero response at low drug concentration, or (3) had a replicate screen (technical in OHSU and biological in FIMM) to which it strongly differed (Supplementary Figs. 26–31). However, we found that this outlier-removal strategy had little impact on prediction performance (Supplementary Fig. 32) and hence did not apply it in the analyses.

Bayesian regression analysis

BMSR models the gene coefficients in each dataset as a sample from a shared, underlying data generating distribution and by regularizing the mean of that distribution to provide parsimony. Formally, it performs joint regression across all datasets constrained to a set of N_G common genes as

$$\mathbf{y}^{(d)} \sim N(X^{(d)}\boldsymbol{\beta}^{(d)}, \sigma^{(d)}\mathbf{I})$$
$$\boldsymbol{\beta}^{(d)} \sim N(\boldsymbol{\beta}, 0.5) \tag{1}$$
$$\beta_g \sim N(0, \lambda_g^2 \tau^2),$$

where $\mathbf{y}^{(d)} \in \mathbb{R}^{N_d \times 1}$ is the response vector for a particular drug across the N_d patient samples in dataset $d \in \{\text{FIMM}, \text{OHSU}\}$, $X^{(d)} \in \mathbb{R}^{N_d \times N_G}$ is the corresponding expression matrix over N_G genes, $\boldsymbol{\beta}^{(d)} \in \mathbb{R}^{N_G \times 1}$ is the gene regression coefficient vector, $\mathbf{I}$ is the $N_G \times N_G$ identity matrix, and the standard deviation $\sigma^{(d)}$ has a non-informative noise prior

$$\sigma^{(d)} \sim \text{IG}(1, 1),$$

with $\text{IG}(\alpha, \beta)$ the Inverse Gamma distribution.

BMSR models coefficient vectors $\boldsymbol{\beta}^{(d)}$ using the joint hierarchical prior with the shared mean coefficient vector $\boldsymbol{\beta} \in \mathbb{R}^{N_G \times 1}$. It regularizes the scalar components β_g using the Finnish horse-shoe prior[69]

$$\lambda_g \sim C^+(0, 1)$$
$$\tau \sim C^+(0, \tau_0)$$
$$\tau_0 = \frac{p_0}{N_G - p_0} \frac{\sum_d \sigma^{(d)}}{\sqrt{N_d}},$$

where $C^+(\mu, \sigma)$ is the half-Cauchy distribution with location μ and scale σ, the scalar λ_g induces localized gene-wise regularization, and the scalar τ is the global regularization parameter that induces the number of active genes (p_0) a priori. Collectively, this formulation encourages the dataset-specific coefficients to either have large magnitude in both datasets (i.e., representing genes whose expression makes a large contribution to the response) or small magnitude in both datasets (i.e., genes with little or no contribution) and to have the same direction (i.e., sign) in both datasets.

BMSMTR simultaneously analyzes multiple datasets (i.e., multi-source, as in BMSR) and also multiple drugs (i.e., multi-task) in a set of drugs $\mathcal{I}$. It does so by generalizing Eq. (1) according to

$$\mathbf{y}^{(d,i)} \sim N\left(X^{(d)}\boldsymbol{\beta}^{(d)}w^{(i)}, \sigma^{(d)}\mathbf{I}\right),$$

with the response vector $\mathbf{y}^{(d,i)} \in \mathbb{R}^{N_d \times 1}$ for dataset d and drug $i \in \mathcal{I}$ distributed about a mean that is a product of a factor $X^{(d)}\boldsymbol{\beta}^{(d)}$ common to drugs in $\mathcal{I}$ and a factor $w^{(i)} \sim N(0.5, 0.5)$ specific to drug i.

The BMSR and BMSMTR models were implemented in STAN[70], a platform for statistical modeling that provides efficient automatic procedures for Bayesian statistical inference. We performed inference using MCMC sampling with 500 samples of the posterior and a burn-in of 500. BMSR and BMSMTR are available at https://github.com/suleimank/bmsr.

Statistical analysis

Statistical analyses were performed in R[68]. Wilcoxon rank-sum test was performed using `wilcox.test`. Fisher's exact test was performed with `fisher.test`. Linear regression was performed with `lm`. Forest plot for multivariate linear regression was generated using the `forestmodel` package.

Gene expression analysis

Ridge regression was performed using `glmnet`[71]. Genes used as variables in (ridge, BMSR, and BMSMTR) regression models were filtered to exclude those with low expression (Supplementary Figs. 14–17) and low variance (Supplementary Figs. 18–19). Gene set enrichment was performed using `fgsea`[72], including relative to the set of monocyte marker genes defined in CIBERSORT[73]. Genes differentially expressed in monocytes relative to other cell populations were determined by applying `limma`[74] to the expression dataset GSE24759[26]. Immune cell fractions were computed with CIBERSORT using non-log expression data and parameters QN=TRUE, absolute=TRUE, and abs_method="no.sumto1". Expression of the genes *BCL3*, *CD14*, *LILRB1*, *LRP1*, *MAFB*, *PSAP*, *SLC15A3*, and *SLC7A7* was compressed into a single enrichment score using GSVA[75]. This enrichment score is the monocytic signature. Additional details are provided in Supplementary Methods.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

DATA AVAILABILITY

We re-analyzed four datasets in this study. The OHSU/Beat AML dataset is available on the Synapse data-sharing platform (https://www.synapse.org/#!Synapse:syn2942337/wiki/390658), at dbGaP (study ID 30641 and accession ID phs001657.v1.p1), as supplementary data within the original manuscript[4], and via ORCESTRA (https://doi.org/10.5281/zenodo.4582786). Additionally, the OHSU dataset can be interactively browsed using Vizome (www.vizome.org). Our re-analysis of the OHSU/Beat AML drug response data are included in Supplementary Tables 1 and 17. The FIMM (MCM) and FIMM (CM) RNA-seq data are available for re-analysis upon request at a secure and GDPR-compliant data analysis data lake environment. Only aggregate data can be downloaded from the data lake. The Lee dataset is available as Supplementary Data within the original manuscript[31]. The Tavor dataset is available via ORCESTRA (https://doi.org/10.5281/zenodo.4585705)[14].

CODE AVAILABILITY

The BMSR and BMSMTR code is available on GitHub (https://github.com/suleimank/bmsr).

REFERENCES

1. American Cancer Society. Cancer facts & figures. https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2019/cancer-facts-and-figures-2019.pdf (2019).
2. Bohl, S. R., Bullinger, L. & Rucker, F. G. New targeted agents in acute myeloid leukemia: new hope on the rise. *Int. J. Mol. Sci.* **20**, 1983 (2019).
3. Short, N. et al. Advances in the treatment of acute myeloid leukemia: new drugs and new challenges. *Cancer Discov.* **10**, 506–525 (2020).
4. Tyner, J. W. et al. Functional genomic landscape of acute myeloid leukaemia. *Nature* **562**, 526–531 (2018).
5. Malani, D. et al. Identification and clinical exploration of individualized targeted therapeutic approaches in acute myeloid leukemia patients by integrating drug response and deep molecular profiles. *Blood* **130**, 854–854 (2017).

6. Sharma, P. & Pollyea, D. A. Shutting down acute myeloid leukemia and myelo-dysplastic syndrome with BCL-2 family protein inhibition. *Curr. Hematol. Malig. Rep.* **13**, 256–264 (2018).

7. DiNardo, C. et al. Venetoclax combined with decitabine or azacitidine in treatment-naive, elderly patients with acute myeloid leukemia. *Blood* **133**, 7–17 (2019).

8. Dombret, H. et al. International phase 3 study of azacitidine vs conventional care regimens in older patients with newly diagnosed AML with >30% blasts. *Blood* **126**, 291–299 (2015).

9. Kuusanmaki, H. et al. Phenotype-based drug screening reveals association between venetoclax response and differentiation stage in acute myeloid leukemia. *Haematologica* **105**, 708–720 (2020).

10. Pei, S. et al. Monocytic subclones confer resistance to venetoclax-based therapy in patients with acute myeloid leukemia. *Cancer Discov.* **10**, 536–551 (2020).

11. Hamid, J. S. et al. Data integration in genetics and genomics: methods and challenges. *Hum. Genomics Proteomics* **2009**, 869093 (2009).

12. Han, L. et al. Concomitant targeting of bcl2 with venetoclax and mapk signaling with cobimetinib in acute myeloid leukemia models. *Haematologica* **105**, 697–707 (2020).

13. Geeleher, P., Cox, N. J. & Huang, R. S. Cancer biomarker discovery is improved by accounting for variability in general levels of drug sensitivity in pre-clinical models. *Genome Biol.* **17**, 190 (2016).

14. Tavor, S. et al. Dasatinib response in acute myeloid leukemia is correlated with FLT3/ITD, PTPN11 mutations and a unique gene expression signature. *Haematologica* **105**, 2795–2804 (2020).

15. Reche, P. A. et al. Human thymic stromal lymphopoietin preferentially stimulates myeloid cells. *J. Immunol.* **167**, 336–343 (2001).

16. Zhang, X. et al. Maternally expressed gene 3, an imprinted noncoding RNA gene, is associated with meningioma pathogenesis and progression. *Cancer Res.* **70**, 2350–2358 (2010).

17. Wang, X. et al. Involvement of the MKK6-p38gamma cascade in gamma-radiation-induced cell cycle arrest. *Mol. Cell. Biol.* **20**, 4543–4552 (2000).

18. Lam, M. et al. Evidence that BCL-2 represses apoptosis by regulating endoplasmic reticulum-associated Ca2+ fluxes. *Proc. Natl Acad. Sci. USA* **91**, 6569–6573 (1994).

19. Vogl, T., Gharibyan, A. L. & Morozova-Roche, L. A. Pro-inflammatory S100A8 and S100A9 proteins: self-assembly into multifunctional native and amyloid complexes. *Int. J. Mol. Sci.* **13**, 2893–2917 (2012).

20. Chipuk, J. E., Moldoveanu, T., Llambi, F., Parsons, M. J. & Green, D. R. The BCL-2 family reunion. *Mol. Cell* **37**, 299–310 (2010).

21. Chen, S., Evans, H. G. & Evans, D. R. FAM129B/MINERVA, a novel adherens junction-associated protein, suppresses apoptosis in HeLa cells. *J. Biol. Chem.* **286**, 10201–10209 (2011).

22. Ghavami, S. et al. S100A8/A9 induces autophagy and apoptosis via ROS-mediated cross-talk between mitochondria and lysosomes that involves BNIP3. *Cell Res.* **20**, 314–331 (2010).

23. Joshi, A. D. et al. ATM, CTLA4, MNDA, and HEM1 in high versus low CD38 expressing B-cell chronic lymphocytic leukemia. *Clin. Cancer Res.* **13**, 5295–5304 (2007).

24. Safikhani, Z. et al. Revisiting inconsistency in large pharmacogenomic studies. *F1000Res* **5**, 2333 (2016).

25. Fishilevich, S. et al. Genic insights from integrated human proteomics in Gene-Cards. *Database* **2016**, baw030 (2016).

26. Novershtern, N. et al. Densely interconnected transcriptional circuits control cell states in human hematopoiesis. *Cell* **144**, 296–309 (2011).

27. Bagger, F. O., Kinalis, S. & Rapin, N. BloodSpot: a database of healthy and malignant haematopoiesis updated with purified and single cell mRNA sequencing profiles. *Nucleic Acids Res.* **47**, D881–D885 (2019).

28. Ziegler-Heitbrock, L. et al. Nomenclature of monocytes and dendritic cells in blood. *Blood* **116**, 74–80 (2010).

29. Ding, C. & Peng, H. Minimum redundancy feature selection from microarray gene expression data. *J. Bioinform. Comput. Biol.* **3**, 185–205 (2005).

30. De Jay, N. et al. mRMRe: an R package for parallelized mRMR ensemble feature selection. *Bioinformatics* **29**, 2365–2368 (2013).

31. Lee, S. I. et al. A machine learning approach to integrate big data for precision medicine in acute myeloid leukemia. *Nat. Commun.* **9**, 42 (2018).

32. Karjalainen, R. et al. JAK1/2 and BCL2 inhibitors synergize to counteract bone marrow stromal cell-induced protection of AML. *Blood* **130**, 789–802 (2017).

33. Pan, R. et al. Selective BCL-2 inhibition by ABT-199 causes on-target cell death in acute myeloid leukemia. *Cancer Discov.* **4**, 362–375 (2014).

34. Niu, X. et al. Acute myeloid leukemia cells harboring MLL fusion genes or with the acute promyelocytic leukemia phenotype are sensitive to the Bcl-2-selective inhibitor ABT-199. *Leukemia* **28**, 1557–1560 (2014).

35. Kontro, M. et al. HOX gene expression predicts response to BCL-2 inhibition in acute myeloid leukemia. *Leukemia* **31**, 301–309 (2017).

36. Zhang, H. et al. Integrated analysis of patient samples identifies biomarkers for venetoclax efficacy and combination strategies in acute myeloid leukemia. *Nat. Cancer* **8**, 826–839 (2020).

37. Karjalainen, R. et al. Elevated expression of S100A8 and S100A9 correlates with resistance to the BCL-2 inhibitor venetoclax in AML. *Leukemia* **33**, 2548–2553 (2019).

38. Rapin, N. et al. Comparing cancer vs normal gene expression profiles identifies new disease entities and common transcriptional programs in AML patients. *Blood* **123**, 894–904 (2014).

39. Haibe-Kains, B. et al. Inconsistency in large pharmacogenomic studies. *Nature* **504**, 389–393 (2013).

40. Garnett, M. J. et al. Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature* **483**, 570–575 (2012).

41. Barretina, J. et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).

42. Geeleher, P., Gamazon, E. R., Seoighe, C., Cox, N. J. & Huang, R. S. Consistency in large pharmacogenomic studies. *Nature* **540**, E1–E2 (2016).

43. Safikhani, Z. et al. Safikhani et al. reply. *Nature* **540**, E2–E4 (2016).

44. Mpindi, J. P. et al. Consistency in drug response profiling. *Nature* **540**, E5–E6 (2016).

45. Safikhani, Z. et al. Safikhani et al. reply. *Nature* **540**, E6–E8 (2016).

46. Bouhaddou, M. et al. Drug response consistency in CCLE and CGP. *Nature* **540**, E9–E10 (2016).

47. Safikhani, Z. et al. Safikhani et al. reply. *Nature* **540**, E11–E12 (2016).

48. Yadav, B. et al. Quantitative scoring of differential drug sensitivity for individually optimized anticancer therapies. *Sci. Rep.* **4**, 5193 (2014).

49. Seashore-Ludlow, B. et al. Harnessing connectivity in a large-scale small-molecule sensitivity dataset. *Cancer Discov.* **5**, 1210–1223 (2015).

50. Rees, M. G. et al. Correlating chemical sensitivity and basal gene expression reveals mechanism of action. *Nat. Chem. Biol.* **12**, 109–116 (2016).

51. Boyer, T. et al. Clinical significance of ABCB1 in acute myeloid leukemia: a comprehensive study. *Cancers* **11**, 1323 (2019).

52. Guerci, A. et al. Predictive value for treatment outcome in acute myeloid leukemia of cellular daunorubicin accumulation and P-glycoprotein expression simultaneously determined by flow cytometry. *Blood* **85**, 2147–2153 (1995).

53. Walter, R. B. et al. CD33 expression and P-glycoprotein-mediated drug efflux inversely correlate and predict clinical outcome in patients with acute myeloid leukemia treated with gemtuzumab ozogamicin monotherapy. *Blood* **109**, 4168–4170 (2007).

54. He, Z. & Yu, W. Stable feature selection for biomarker discovery. *Comput. Biol. Chem.* **34**, 215–225 (2010).

55. Buness, A., Ruschhaupt, M., Kuner, R. & Tresch, A. Classification across gene expression microarray studies. *BMC Bioinformatics* **10**, 453 (2009).

56. Di Camillo, B. et al. Effect of size and heterogeneity of samples on biomarker discovery: synthetic and real data assessment. *PLoS ONE* **7**, e32200 (2012).

57. Michiels, S., Koscielny, S. & Hill, C. Prediction of cancer outcome with microarrays: a multiple random validation strategy. *Lancet* **365**, 488–492 (2005).

58. Ma, S., Huang, J., Wei, F., Xie, Y. & Fang, K. Integrative analysis of multiple cancer prognosis studies with gene expression measurements. *Stat. Med.* **30**, 3361–3371 (2011).

59. Mammoliti, A. et al. Orchestrating and sharing large multimodal data for transparent and reproducible research. Preprint at *bioRxiv* https://doi.org/10.1101/2020.09.18.303842v2 (2021).

60. Smirnov, P. et al. PharmacoGx: an R package for analysis of large pharmacogenomic datasets. *Bioinformatics* **32**, 1244–1246 (2016).

61. Smirnov, P. et al. Pharmacodb: an integrative database for mining in vitro anticancer drug screening studies. *Nucleic Acids Res.* **46**, D994–D1002 (2018).

62. Loscalzo, S., Yu, L. & Ding, C. Consensus group stable feature selection. In *Proc. 15th ACM SIGKDD International Conference of Knowledge Discovery and Data Mining* 567–576 (ACM Press, 2009).

63. Townsend, K. J. et al. Regulation of mcl1 through a serum response factor/elk-1-mediated mechanism links expression of a viability-promoting member of the bcl2 family to the induction of hematopoietic cell differentiation. *J. Biol. Chem.* **274**, 1801–1813 (1999).

64. Domina, A. M., Vrana, J. A., Gregory, M. A., Hann, S. R. & Craig, R. W. MCL1 is phosphorylated in the pest region and stabilized upon ERK activation in viable cells, and at additional sites with cytotoxic okadaic acid or taxol. *Oncogene* **23**, 5301–5315 (2004).

65. Konopleva, M. et al. Mek inhibition enhances ABT-737-induced leukemia cell apoptosis via prevention of ERK-activated MCL-1 induction and modulation of MCL-1/BIM complex. *Leukemia* **26**, 778–787 (2012).

66. Lin, K. H. et al. Targeting MCL-1/BCL-XL forestalls the acquisition of resistance to ABT-199 in acute myeloid leukemia. *Sci. Rep* **6**, 27696 (2016).

67. Ritz, C., Baty, F., Streibig, J. C. & Gerhard, D. Dose-response analysis using R. *PLoS ONE* **10**, e0146021 (2015).

68. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2014).

69. Piironen, J. & Vehtari, A. Sparsity information and regularization in the horseshoe and other shrinkage priors. *Electron. J. Stat.* **11**, 5018–5051 (2017).

70. Carpenter, B. et al. Stan: a probabilistic programming language. *J. Stat. Softw.* **76**, 1–32 (2017).

71. Friedman, J., Hastie, T. & Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **33**, 1–22 (2010).

72. Korotkevich, G. et al. Fast gene set enrichment analysis. Preprint at bioRxiv. https://www.biorxiv.org/content/early/2021/02/01/060012 (2021).

73. Newman, A. M. et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **12**, 453–457 (2015).

74. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

75. Hanzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* **14**, 7 (2013).

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

T.A., K.W. and J.G. conceived and supervised the project. B.S.W., S.A.K., M.J.M., M.A., T.A., K.W. and J.G. designed computational experiments. B.S.W., S.A.K., M.J.M. M.A. and S.P. analyzed the data. S.A.K. and M.A. developed software methodology (BMSR and/or BMSMTR). D.M., H.K., B.J.D., C.A.H., O.K., S.E.K., K.P., C.E.T. and J.W.T. contributed data. B.S.W., S.A.K., D.M., H.K., C.A.H., C.E.T., J.W.T., T.A., K.W. and J.G. interpreted results. B.S.W., S.A.K., T.A., K.W. and J.G. wrote the paper. All authors read, edited, and approved the final manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41698-021-00209-9.

Correspondence and requests for materials should be addressed to B.S.W.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.