

Multi-omics factor analysis reveals a RAS mutant-like metabolically distinct AML subgroup resistant to venetoclax

Jan Jacob Schuringa

j.j.schuringa@umcg.nl

UMCG

Dominique Sternadt

UMCG

Shanna Hogeling

UMCG <https://orcid.org/0000-0003-2124-5333>

Lieve Oudejans

UMCG

Nikita la Rose

UMCG

Gerwin Huls

University Medical Center Groningen

Diego Peraira-Martins

UMCG <https://orcid.org/0000-0002-3302-4311>

Article

Keywords:

Posted Date: June 16th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9859908/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: There is **NO** conflict of interest to disclose.

Abstract

Acute myeloid leukemia (AML) exhibits extensive clonal and functional heterogeneity, limiting the ability of current genotype-based classifications to predict treatment response. We applied multi-omics factor analysis (MOFA) in a cohort of AML cases, integrating mutational status, proteomics, metabolomics, and two *ex vivo* drug screens comprising 115 agents targeting key metabolic and epigenetic pathways. A five-factor model captured interpretable, cross-modal axes of variation. One axis, Factor 2, defined a monocytic/inflammatory program with acylcarnitine accumulation and broad drug insensitivity. Projection of factor scores into four independent AML cohorts showed that Factor 2 reproducibly associated with reduced venetoclax sensitivity and enrichment of RAS-related signatures. This RAS-like phenotype extended beyond canonical RAS-mutant cases to include genotypically wild-type samples with high Factor 2 scores, indicating that proteomic and metabolomic integration reveals pathway activation not captured by genotype alone. Leveraging microenvironmental features of Factor 2^{high} disease, we further identified selective sensitivities to FABP4 inhibition and the POLG inhibitor alovudine, revealing candidate strategies for venetoclax-resistant AML.

Introduction

Acute myeloid leukemia (AML) is characterized by marked inter- and intra-patient heterogeneity, reflected by diverse mutational landscapes that define disease outcomes and therapeutic response. Large-scale genomic studies have distinguished multiple cytogenetic and molecular subtypes and identified more than 40 recurrently mutated genes implicated in AML pathogenesis and prognosis.^{1–4} Although these discoveries have transformed disease classification and risk stratification, genetic alterations alone do not fully explain the phenotypic diversity observed in AML, as patients with similar mutational profiles can exhibit distinct cellular states and therapeutic vulnerabilities.

Emerging evidence suggests that AML heterogeneity reflects coordinated alterations across multiple molecular layers, including transcriptional regulation, protein signaling networks, and cellular metabolism. While large cohorts integrating genomic and transcriptomic data have substantially improved our understanding of AML biology,^{5–11} the majority of studies have analyzed individual molecular and functional modalities such as transcriptomes, proteomes, metabolomes and drug screenings data independently, limiting the ability to uncover cross-modal programs that link genotype to functional cellular states. Because proteins and metabolites reflect pathway activity and cellular physiology more directly, integrating these molecular layers with functional phenotypes such as drug response may reveal clinically actionable disease programs that are not captured by single-omic analyses.

To uncover cross-modal disease programs linking genotype to functional cellular states in AML, we applied multi-omics factor analysis (MOFA), an unsupervised framework that identifies latent factors representing major sources of variation across datasets.^{12,13} We integrated mutational profiles, proteomic and metabolomic measurements obtained from stem/progenitor-enriched AML cell populations, together with two *ex vivo* metabolic and epigenetic drug response profiles generated from matched primary mononuclear cells. This integrative analysis aimed to identify molecular programs that link mutations and cellular phenotypes to therapeutic vulnerabilities. We identified a distinct AML subgroup associated with a RAS-mutant-like phenotype and reduced sensitivity to venetoclax, for which we propose possible therapeutic alternatives.

Results

Multi-omics factor analysis in primary AML samples.

To uncover novel targetable vulnerabilities in AML, we leveraged a unique cohort of primary AML samples in which the CD34⁺ or CD117⁺ compartments had been extensively profiled at the mutational, proteomic and metabolomics levels. While each of these molecular layers have been individually analyzed in our prior studies,^{14,15} we asked whether integrating these measurements could reveal coordinated sources of heterogeneity spanning genotype, protein networks, metabolism, and drug response that would be missed in single-view analyses. To provide functional context, we complemented the existing

omics layers with two additional *ex vivo* drug screens focused on metabolic (n = 42) and epigenetic (n = 73) dependencies (for experimental details see Materials and Methods section, Supplemental Tables 1 and 2). Thus, we used the five datasets to train a multi-omics factor analysis (MOFA) model to identify latent factors that capture major patterns of variation shared across the data. Each factor is characterized by a set of genes, each assigned a weight reflecting how strongly it contributes to that factor. The resulting weights from the proteome layer were then used to assign factor scores to independent AML cohorts with partially overlapping data types (Fig. 1A). Altogether, we obtained data from 30 patients, with 17 individuals having complete measurements across all five omic views (Fig. 1B-C).

To determine the dimensionality of the model, we trained MOFA with three to six latent factors (K) and evaluated variance explained, leave-one-out robustness, cross-K factor correspondence, and seed stability. Total variance explained increased with K across views, confirming that additional factors captured further structure in the data and suggesting that models with three or four factors may have insufficient capacity to represent the major axes of heterogeneity (Supplemental Fig. 1A). We next assessed robustness by leave-one-out refitting, retraining the model after removing one sample at a time and comparing recovered factor scores to the reference model. The five-factor model was generally robust to sample removal, with the exception of Factor 5 which showed a marked drop in worst-case correlation upon removal of a single sample (Supplemental Fig. 1B). However, cross-K correspondence analysis confirmed that this factor is consistently recovered in the six-factor model (Supplemental Fig. 1C), supporting its interpretation as reproducible biological variation rather than a sample-specific artefact. In contrast, the six-factor model did not improve upon this robustness, as a greater number of factors showed sensitivity to sample removal (Supplemental Fig. 1B). This suggests that increasing K may lead factors to become progressively more dependent on individual samples. The additional factor introduced in the six-factor model showed limited correspondence to any factor in the five-factor model, indicating a potentially new biological signal, although its weak association with Factor 3 suggests that it may further dissect an existing axis of variation. Given the limited cohort size, with only 17 samples with complete overlap across all views, we could not reliably determine whether this additional factor represents a true biological axis or over-parameterization. Altogether, although the six-factor model remained plausible, we selected the model with five factors as it may represent a most conservative solution for downstream analysis, as it captured the dominant axes of multi-omic heterogeneity without sacrificing generalizability across samples. Finally, we evaluated robustness across 10 random seeds. After aligning factors across runs, factor scores were nearly identical, confirming stable inference (Supplemental Fig. 1D).

Across the five omic layers, proteomics captured the largest proportion of variance, in line with being the dataset with the largest number of variables (Fig. 1D). Factor 1 explained most of total variance and was predominantly proteomics-derived, representing a dominant biological axis with minimal contribution from the drug screens. Factor 2 reflected shared contribution from the different layers, especially from proteome and the metabolome, suggesting a strong coordination between protein and metabolic states, reflected in drug sensitivities and independent of mutational status. Factors 3 and 4 were primarily shaped by the metabolic drug screen and metabolome, respectively, while Factor 5 showed weaker and more diffuse contributions, mainly from the proteomics and the epigenetic drug screen. Cumulatively, the five factors explained 65.96% of the variance in the proteome, 36.5% in the metabolome, 18.08% in the metabolic drug screen, 10.08% in the epigenetic drug screen, and 5.98% in the mutation data (Fig. 1E).

MOFA identifies proteome-driven AML programs with distinct drug response profiles.

Since MOFA learns latent factors that capture major axis of variation across multiple data modalities, each factor often represents a continuous molecular gradient or subgroups of samples along that axis.^{12,13} We first examined the distribution of factor scores per sample (Fig. 2A). Factor 1 showed the most pronounced spread among samples and was associated with *NPM1*, *FLT3* and *DNTM3A* mutations (Fig. 2A-C). A similar mutational enrichment pattern was observed for Factor 3, while Factor 4 primarily associated with *IDH1/2* mutations (Fig. 2B-C). In contrast, Factor 2 showed limited association with common driver mutations, in line with the limited contribution of this layer to the factor definition (Fig. 1D), with weaker enrichments for *TET2* and splicing (*SRSF2/SF3B1*) mutations (Fig. 2B-C).

To further understand the underlying biology of each factor, we focused on the modalities contributing most strongly to that factor and examined the top feature loadings within each layer (Fig. 2D-G, Supplemental Fig. 2A-D). Factor 1 and 3 shared several features, consistent with their partial overlap in mutational composition (Fig. 2C). Both factors were enriched for oxidative phosphorylation (OXPHOS)-related signatures, aligning with previous studies associating *FLT3*-ITD mutations to increased mitochondrial reliance.^{15–18} Notably, Factor 3 also showed a stronger lipid metabolism component, supported by concordant signals in both proteomic and metabolomic layers (Figs. 2D-E). This pattern was reflected in the metabolic drug screen, where higher Factor 3 scores were associated with increased sensitivity to multiple compounds targeting lipid metabolic pathways, and higher Factor 1 scores to OXPHOS inhibitors (Fig. 2G, Supplemental Fig. 2C).

Factor 2 captured a distinct profile, characterized by enrichment of inflammatory and mTOR signatures in the proteome, together with accumulation of acylcarnitine species in the metabolome and relative insensitivity to many drugs, both to epigenetic and metabolic inhibitors (Supplemental Fig. 2C-D, Fig. 2D-G). Furthermore, Factor 4 was dominated by *IDH1/2* mutations, consistent with the established link between these mutations and accumulation of the oncometabolite 2-hydroxyglutarate (2-HG) (Fig. 2B-E).^{19–21} In addition, Factor 4 showed signatures related to heme metabolism, consistent with reports that *IDH1/2* mutations can perturb heme biosynthesis through metabolic rewiring.²² This factor also aligned with heightened sensitivity to multiple inhibitors in the epigenetic drug screen (Fig. 2F, Supplemental Fig. 2D), consistent with the broader concept that *IDH1/2*-mutant AML is characterized by epigenetic dysregulation driven by 2-HG. Even though Factor 5 explained less variance compared to other factors, we still observed an enrichment in features associated with MECOM/EVI1-related biology, a high-risk AML subgroup with poor outcomes.^{23,24} Together, these factors define biologically interpretable axes of AML heterogeneity that integrate molecular state with functional vulnerabilities, motivating their evaluation in independent AML cohorts.

External validation in independent cohorts reveals association of Factor 2 with venetoclax resistance.

To assess whether the latent programs learned by MOFA reflect broader AML biology, we transferred the factor structure to independent datasets by computing factor-derived scores from the original MOFA loadings. This approach was motivated by the strong contribution of the proteomic view to the variance explained in our model and by the fact that most publicly available independent cohorts provide either gene expression or protein abundance measurements that can be mapped to out proteomic features.

Before applying the approach to independent datasets, we validated it within our own cohort. We first imputed missing values by using k-nearest neighbor and computed weighted sum of expression (WSE), a weighted sum of standardized feature values multiplied by the corresponding score of each MOFA loading. The resulting WSE scores reproduced the original factor score distributions (Fig. 2A, Supplemental Fig. 3A) and correlated almost perfectly with the MOFA-inferred factor scores across all factors (Supplemental Fig. 3B). These results indicate that loadings-based scoring provides a faithful approximation of the factor space learned by MOFA in our dataset.

We next applied this scoring strategy to four independent AML cohorts with clinical and molecular annotations and, for some subsets, drug sensitivity data, namely BeatAML,⁶ TCGA-LAML,⁴ FPMTB,⁷ and a large proteomics AML cohort from Cheng et al.⁸ Across datasets, substantial overlap between our proteomic feature set and each cohort's gene/protein identifiers enabled computation of factor-derived scores, with the lowest overlap being of 89% in the TCGA (Supplemental Fig. 3C). Factor scores per patient were all standardized by cohort per factor to verify the distribution of scores across patients (Fig. 3A, Supplemental Fig. 3D).

The contribution of mutations to Factors was broadly reproduced across cohorts. The triple-mutant pattern (*DNMT3A/NPM1/FLT3*) associated with Factor 1 and 3 was again identified, although the relative weighting shifted toward Factor 3 compared with the discovery cohort (Fig. 3B, Supplemental Fig. 3E). These analyses also revealed a new association between *CEBPA* mutations and Factor 1, not covered in our initial mutational panel, and Factor 4 remained consistently and strongly defined by *IDH1/2* mutations across all cohorts, indicating a stable and transferable *IDH1/2*-linked

program. For Factor 5, the association with *MECOM/EVI1* was reproduced in the BeatAML cohort, and an additional association with *TP53* mutations was observed, reinforcing the poor-outcome phenotype captured by this factor. Interestingly, while Factor 2 in our cohort showed a mild enrichment for *TET2* and splicing-related mutations, external cohorts additionally revealed associations with *NRAS/KRAS* mutations (RAS^{mut}) (Fig. 3B), which was not revealed in our discovery set, potentially due to the limited representation of RAS^{mut} AMLs (Fig. 3B).

To link factor programs to therapeutic response, we correlated factor-derived scores with available drug sensitivity measurements. Across datasets, we observed that Factor 2 was reproducibly associated with reduced venetoclax sensitivity (Fig. 3C-F, Supplemental Fig. 3F-G), consistent with our discovery cohort, although to a lesser extent (Supplemental Fig. 2C-D). This finding is consistent with prior work showing that RAS/MAPK pathway activation and related mutations can mediate resistance to venetoclax-based regimens. In line, Factor 2 scores showed positive associations with MEK inhibitors, such as selumetinib and trametinib, supporting a model in which this factor captures a RAS-driven venetoclax resistance (Fig. 3C).^{25,26} Additionally, venetoclax sensitivity correlated negatively with Factor 4 in both BeatAML and Cheng et al. cohorts, consistent with evidence that *IDH1/2* mutant AMLs can exhibit heightened BCL-2 dependence and sensitivity to venetoclax.²⁷⁻²⁹

Factor 2 uncovers a RAS mutant-like phenotype linked to venetoclax resistance.

Given the association of Factor 2 with venetoclax resistance in external cohorts enriched for RAS^{mut} , and that our discovery cohort showed a similar pattern despite the lack of RAS^{mut} , we questioned whether Factor 2 would capture an AML state that was only possible due to the integration of multiple layers, thereby revealing a new AML subgroup we could term RAS mutant-like patients. This factor was equally defined both by proteomic and metabolomic variations (Fig. 1D), a layer that is often not included in many AML studies.

In our discovery cohort, Factor 2 scores were correlated with FAB classification (Supplemental Fig. 4A). Because monocytic AMLs have been often linked to reduced sensitivity to venetoclax-based regimens,^{30,31} we examined Factor 2 scores across FAB subtypes. Samples classified as FAB M4/M5 showed the highest Factor 2 scores (Supplemental Fig. 4B), supporting the notion that Factor 2 captures a monocytic/myelomonocytic program. Considering that only a single RAS^{mut} case was present in our discovery cohort (Fig. 2B), we tested the relationship between Factor 2 and RAS status using external cohorts. In two out of four datasets, RAS^{mut} samples showed significantly higher Factor 2 scores (Fig. 4A). Often, *PTPN11* mutations can also recapitulate RAS^{mut} phenotypes due to activation of the same pathways, and even though *PTPN11*-mutant samples showed higher Factor 2 scores in the BeatAML cohort, this was not significant in the other cohorts (Supplemental Fig. 4C). Considering that additional lesions have been associated with inferior responses to venetoclax-based therapy, including *TP53* and *FLT3* mutations,³² we fitted multivariable linear regression models to determine whether Factor 2 was associated with venetoclax response independently of these covariates (Fig. 4B). In these models, the regression coefficient β represents the expected change in venetoclax AUC per 1-standard-deviation increase in Factor 2, after adjustment for RAS status, FAB morphology, *FLT3*, and *TP53* mutations. Because higher AUC reflects reduced ex vivo venetoclax sensitivity, positive β values indicate an association with venetoclax resistance. Factor 2 remained positively associated with venetoclax AUC in the full cohort and showed a stronger association in the RAS wild-type (RAS^{wt}) subset, suggesting Factor 2 captures resistance-associate variance beyond these established clinical correlates.³²

We therefore examined transcriptional programs associated with Factor 2 using single-sample gene set enrichment analysis (ssGSEA). High Factor 2 scores were associated with enrichment of inflammatory programs, fatty acid metabolism and more committed progenitor signatures programs, with concomitant depletion of stem-cell signatures (Supplemental Fig. 4D). We dichotomized our samples based on Factor 2 scores, defining Factor 2 high (Factor 2^{high}) samples as those with a z-score above 0, and focused on RAS biology. Across all datasets, we found that RAS signature enrichment in RAS^{wt} Factor 2^{high} was equally high to that observed in RAS^{mut} Factor 2^{high} samples (Fig. 4C, Supplemental Fig. 4E). This supports

the notion that Factor 2 is uncovering a RAS^{mut} -like phenotype in samples absent of RAS mutations. To check whether other genetic drivers could be playing a role in the RAS^{wt} group, we compared Factor 2^{high} versus Factor 2^{low} (Fig. 4D, Supplemental Fig. 4F-G). In the BeatAML, *ASXL1* mutations were enriched among samples with higher Factor 2 scores, whereas *CEBPA*, *FLT3*, *NPM1* and *WT1* were more frequent in samples with lower scores (Fig. 4D). No other recurrent mutations were identified specifically in the RAS^{wt} Factor 2^{high} group, but an enrichment for *CEBPA*, *FLT3* and *NPM1* mutations was observed in the RAS^{wt} Factor 2^{low} group (Supplemental Fig. 4F). By analyzing the mutational differences in only Factor 2^{high} cases, the *ASXL1* co-mutation was recurrent among RAS^{mut} cases (Supplemental Fig. 4G).

To explore potential differences and similarities between RAS^{mut} and RAS^{wt} Factor 2^{high} samples, we performed differential gene expression analysis between Factor 2^{high} versus Factor 2^{low} subgroups. These analyses revealed that Factor 2 identifies a transcriptionally distinct subgroup regardless of RAS mutation status (Fig. 4E). In line with this, PCA indicated that Factor 2^{high} samples overlap better than samples defined by RAS mutation status (Supplemental Fig. 4H). We intersected genes significantly upregulated in Factor 2^{high} samples ($\log_2FC > 1$; adj p. value < 0.05) with positive Factor 2 loadings. This yielded 237 overlapping genes, and gene ontology (GO) analysis highlighted metabolic processes with prominent fatty acid-related terms (Fig. 4F-G). Indeed, we have previously noted that acylcarnitines were part of the top metabolite loadings from our discovery cohort in defining Factor 2 (Fig. 2B, Supplemental Fig. 4I), consistent with altered fatty acid transport and beta-oxidation flux. Clearly, these signatures are equally present in the RAS^{wt} group but must be driven by unknown mechanisms that remain to be explored.

Finally, in order to identify potential drugs that could be considered in the venetoclax-resistant Factor 2^{high} subgroup, we re-analyzed the drug screen data. In the metabolic screen performed in liquid cultures, Factor 2^{high} samples were generally associated with a broad insensitivity to most compounds, whereas only 3-Bromopyruvic acid (hexokinase 2 inhibitor), alovudine (POLG inhibitor³³), BMS-309403 (FABP4 inhibitor³⁴) and metformin (OXPHOS inhibitor¹⁵) demonstrated a trend towards increased sensitivity (Supplemental Fig. 2C). Metabolic dependencies and drug responses are often modulated by the bone marrow microenvironment whereas AML cells can promote remodeling of stromal cells towards an adipogenic, supportive niche to increase fatty acid availability and promote cell survival and proliferation.^{35–40} Therefore, we next evaluated efficacy of some of these compounds in the presence of bone marrow stromal cells. Under co-culture conditions, BMS-309403 and alovudine showed significantly increased activity in Factor 2^{high} samples, while metformin showed a similar trend that did not reach statistical significance (Fig. 4H). In the stromal context, FABP4 inhibition by BMS-309403 may disrupt intracellular fatty-acid trafficking and limit the use of stromal-derived fatty acids that can fuel FAO/OXPHOS-associated survival programs. In addition to fatty acid-related processes, GO analysis also revealed enrichment for processes related to redox regulation and nucleotide metabolism (Fig. 4G), consistent with increased demands on mitochondrial maintenance in this metabolic state. In line, alovudine-mediated inhibition of POLG may compromise mitochondrial DNA maintenance and respiratory-chain function, consistent with prior evidence that alovudine impairs OXPHOS and AML cell viability.³³ The weaker trend observed with metformin may similarly reflect the sensitivity to mitochondrial perturbation, although compensatory metabolic rewiring might attenuate its activity, as metformin might also promote ferroptosis in cells with high fatty acid program.¹⁵ Altogether, these observations suggest that Factor 2 is characterized by metabolic dependencies that may be therapeutically actionable, raising the possibility that targeting fatty acid or mitochondrial metabolism could represent a rational strategy for this subgroup, particularly in the context of venetoclax resistance.

Discussion

AML is a complex disease with pronounced inter- and intra-patient heterogeneity. Despite relatively high remission rates, relapses are common and lead to a more aggressive disease with dismal outcomes and limited treatment options. Over the past decade, large-scale studies have established genomic and transcriptomic subtype frameworks and have improved

biological understanding and risk stratification.^{6,10,41,42} However, resistance to targeted and cytotoxic therapies often reflects cell state and microenvironmental context in addition to genotype and therefore may not be fully captured by single-modality analyses. Here, we integrated mutational, proteomic, metabolomic, and functional drug response data from primary AML samples using MOFA to identify coordinated programs spanning molecular state and *ex vivo* drug sensitivity.

The identified latent factors recapitulated established links between genotype, downstream metabolic state, and drug sensitivities. Factor 1 and 3 aligned with AML samples enriched for *FLT3/NPM1/DNTM3A* mutations and mapped to mitochondrial/OXPHOS-associated programs. This is consistent with multiple reports connecting *FLT3*^{mut}-driven AML to mitochondrial metabolism.^{15,16,18} Notably, Factor 3 separated from Factor 1 by an additional lipid metabolism component, that was reflected both in the metabolome and drug sensitivities. These differences support the view that even within broadly similar mutational landscapes, AML can occupy distinct metabolic configurations that shape susceptibility to therapy. Such difference is consistent with work showing that co-mutation context and differentiation state modulate mitochondrial and lipid metabolic dependencies and can influence response to agents targeting these pathways.^{15–17,43,44} In this regard, our factor structure may provide a quantitative way to represent overlapping but separable metabolic programs that are not easily resolved by mutation calls alone.

Factor 4 emerged as the most consistent and transferable program across cohorts and was dominated by *IDH1/2* mutations, consistent with the known accumulation of the oncometabolite 2-hydroxyglutarate (2-HG) and its broad epigenetic consequences.^{19–21,45,46} Factor 4 also associated with heme metabolism-related signatures and increased sensitivity across multiple agents in the epigenetic-focused screen, supporting a metabolic-epigenetic axis driven by mutant *IDH1/2*. In independent validation datasets, Factor 4 also tended to correlate with greater venetoclax sensitivity, consistent with reported BCL-2 dependency in *IDH1/2*-mutant contexts.^{27–29}

In contrast to these mutation-driven programs, Factor 2 captured a distinct state defined jointly by proteomic and metabolomic variation, with limited association to common driver mutations in the discovery cohort. Factor 2 was characterized by inflammatory and mTOR-associated programs, accumulation of acylcarnitine species, and broad multi-drug resistance across both epigenetic and metabolic screens. When we projected MOFA loading scores to independent AML cohorts using a weighted-sum-of-expression approach, Factor 2 was consistently associated with reduced venetoclax sensitivity across datasets. Indeed, response to venetoclax is known to be influenced by both genetic alterations, including *RAS*, *TP53*, and *FLT3*,³² and differentiation state, whereas resistance has been linked to monocytic maturation programs and activation of pro-survival signaling pathways.^{26,31,47–49} Our results support a model in which Factor 2 represents a transferable resistance program that could complement mutation-based predictors.

Across independent validation cohorts, Factor 2 showed enrichment for *NRAS/KRAS* mutations in two of four datasets, an association that could not be found in our discovery cohort given the minimal representation of *RAS*^{mut} cases. This prompted us to evaluate other recurrent mutations that could explain the observed venetoclax-resistant phenotype. However, no additional mutations were found to be significantly associated with *RAS*^{wt} Factor 2^{high} group, whereas *FLT3*, *NPM1* and *CEBPA* mutations were frequent in samples with lower Factor 2 scores. We also assessed *PTPN11* separately given its functional convergence with *RAS* signaling,²⁵ but higher Factor 2 scores in *PTPN11*-mutant samples were only evident in one dataset, suggesting that no single mutation class uniformly dictates the Factor 2 program across cohorts. Additionally, Factor 2^{high} samples showed enrichment for *RAS*-focused transcriptional signatures, and *RAS*^{wt} Factor 2^{high} samples exhibited enrichment scores comparable to *RAS*^{mut} Factor 2^{high} cases. We therefore associate this axis as a *RAS*^{mut}-like program to reflect shared downstream transcriptional output beyond mutation status. In the BeatAML, the *ASXL1* co-mutation was recurrent among *RAS*^{mut} Factor 2^{high} samples, aligning with prior reports indicating that *ASXL1* can precede and facilitate *RAS*-driven leukemogenesis.⁵⁰ Importantly, Factor 2 remained associated with reduced venetoclax sensitivity after adjustment for major confounders, supporting the notion that factor-derived scoring captures additional,

clinically relevant variation beyond canonical mutation-based predictors and may help identify venetoclax-resistant biology in *RAS*^{wt} samples.

Furthermore, Factor 2 was also linked to a specific metabolic profile. In the discovery cohort, Factor 2 was marked by accumulation of acylcarnitine species, consistent with altered lipid transport and fatty acid oxidation rates. Intersecting genes upregulated in Factor 2^{high} samples with positive Factor 2 loadings highlighted metabolic processes with prominent fatty acid-related terms, reinforcing the view that this factor reflects an important metabolic rewiring. Notably, certain metabolic vulnerabilities that were not reaching significance in the liquid-culture metabolic screen did emerge under stromal co-culture conditions where Factor 2 scores associated with response to BMS-309403, a FABP4 inhibitor, and alovudine, a POLG inhibitor. This context dependency is consistent with evidence that the bone marrow microenvironment can reshape nutrient availability, lipid handling, and drug response.^{38,39} Because these findings are based on single-agent associations, an important next step will be to test whether targeting these dependencies can restore venetoclax sensitivity and to define whether such effects are mediated by altered fatty acid handling, mitochondrial function, or microenvironment-driven signaling.

A practical outcome of this work is that factor-derived scores can be projected onto external cohorts using the MOFA loadings. This enables evaluation of the same latent programs in larger datasets with greater statistical power and with clinical or molecular annotations that were not available in our discovery cohort. In our data, weighted-sum scoring closely reproduced the original MOFA factor scores, and the high overlap between our proteome-informed features and external gene/protein identifiers supported robust factor projection across cohorts. More broadly, this strategy illustrates how deep multi-omic profiling in a smaller, intensively characterized cohort can be leveraged to generate interpretable programs that are then tested at scale in cohorts with partially overlapping datasets. Loadings-based scoring therefore provides a practical bridge between deep discovery studies and large transcriptome- or proteome-annotated AML resources.

This study also has limitations related to cohort size and composition. Although the dataset captures key biological signals, broader representation of patients and genetic backgrounds, together with increased overlap across molecular layers, would likely enable identification of additional latent factors and deeper resolution of AML heterogeneity. The mutational landscape of the cohort is enriched for *FLT3* and *NPM1* mutations, while many other recurrent AML mutations are represented by few individuals, precluding robust mutation-specific analyses. Moreover, a more balanced number of features across data modalities could improve model fitting and reduce the predominance of proteomic signals, allowing a more nuanced view of how different molecular layers contribute to the observed phenotypes.

In summary, our multi-omic factor analysis provides a unified view of AML heterogeneity that links molecular state to functional drug response. Beyond recovering known genotype-associated programs, this framework identified a *RAS*^{mut}-like, monocytic-skewed state associated with venetoclax resistance across cohorts. These findings support the value of multi-layer integration of proteomic and metabolomic layers with functional profiling to discover clinically relevant AML states and to generate testable hypotheses for better patient stratification and predicting treatment outcomes.

Methods

Resources Table

Reagents	Source	Identifier
Anti-CD45-PECy7	BioLegend	304016
Anti-CD117-PE	ImmunoTools	21271174X2
Anti-CD11b-FITC	ImmunoTools	21279113X2
Anti-CD38-FITC	BioLegend	303504
Anti-CD34-APC	BioLegend	343510
Annexin V-PE	BioLegend	640947
FcR blocking reagent	Mylteni Biotech	130-059-901
10x Annexin V binding buffer	BD Pharmingen	556454
4',6-diamidino-2-phenylindole (DAPI)	Sigma-Aldrich	28718-90-3
DNase I	Roche	11284932001
MgSO4	Sigma-Aldrich	M7506
Heparin	Sigma-Aldrich	60800-63-7
Human Interleukin 3	Sandoz	-
β -mercaptoethanol	Merck Sharp & Dohme BV	60-24-2
Human Granulocyte colony-stimulating factor	Amgen	-
Human Thrombopoietin	Amgen	-
2DG	MedChemExpress	HY-13966
3-Bromopyruvic acid	MedChemExpress	HY-19992
3-Deazaadenosine	Sigma-Aldrich	D8296
3K	SelleckChem	S8616
3-nitropropionic acid	Sigma-Aldrich	N22908
6MP mercaptopurine	MedChemExpress	HY-13677
Ac-CoA Synthase inhibitor	MedChemExpress	HY-104032
AG270	MedChemExpress	HY-138630
Alovudine	MedChemExpress	HY-B1516
AraC	MedChemExpress	HY-13605

Auranofin	MedChemExpress	HY-B1123
AZD3965	MedChemExpress	HY-12750
BMS-309403	MedChemExpress	HY-101903
Brequinar	MedChemExpress	HY-108325
Buthionine Sulfoximine (BSO)	MedChemExpress	HY-106376
CB-839	MedChemExpress	HY-12248
DAP	Sigma-Aldrich	D54850
DASA-58	SelleckChem	S7928
Dipyridamole	MedChemExpress	HY-B0312
Enasidenib	MedChemExpress	HY-18690
Eprenetapopt (APR-246)	SelleckChem	S7724
Etomoxir	MedChemExpress	HY-50202
Fatostatin	MedChemExpress	HY-14452
GW501516	Enzo	ALX-420-032
IACS-010759	MedChemExpress	HY-112037
Ivosidenib	MedChemExpress	HY-18767
KPT9274	SelleckChem	S8444
Metformin	MedChemExpress	HY-B0627
Midostaurin	MedChemExpress	HY-10230
Mubritinib	SelleckChem	S2216
OT-82	MedChemExpress	HY-136241
Rotenone	MedChemExpress	HY-B1756
RZ-2994 (SHIN1)	MedChemExpress	HY-112066
SHIN2	AOBIOUS	AOB31454
SSI-4	Dembitz et al. ⁵¹	-
SSO	Sigma-Aldrich	SML2148
ST1326	Sigma-Aldrich/Avanti	SML2148
Syrosingopine	MedChemExpress	HY-N4115
Triacsin C	MedChemExpress	HY-N6707
TTFA	Abcam	ab223880
UPGL00004	MedChemExpress	HY-119377
Venetoclax	MedChemExpress	HY-15531

FlowJo v10.0.6	Treestar	http://www.flowjo.com/RRID:SCR_008520
Adobe Illustrator	Adobe	https://www.adobe.com/nl/RRID:SCR_010279
RStudio version 2026.01.1+403	Posit team	https://posit.co/download/rstudio-desktop
R version 4.5.2	CRAN	www.r-project.org
Single sample gene set enrichment analysis (ssGSEA)	Barbie et al. ⁵²	https://www.genepattern.org/modules/docs/ssGSEAProjection/4#gsc.tab=0
circlize v.0.4.18	Gu et al. ⁵³	10.32614/CRAN.package.circlize
clue v.0.3-68	Hornik K. ⁵⁴	10.32614/CRAN.package.clue
clusterProfiler v.4.16.0	Yu et al. ⁵⁵	10.18129/B9.bioc.clusterProfiler
ComplexHeatmap v.2.24.1	Gu et al. ⁵⁶	10.18129/B9.bioc.ComplexHeatmap
EnhancedVolcano v.1.26.0	Blighe et al.	10.18129/B9.bioc.EnhancedVolcano
ggplot2 v.4.0.2	Wickham, H. ⁵⁷	https://ggplot2.tidyverse.org/index.html
impute v.1.82.0	Hastie et al.	10.18129/B9.bioc.impute
limma v.3.66.0	Ritchie et al. ⁵⁸	10.18129/B9.bioc.limma
matrixStats v.1.5.0	Bengtsson et al.	10.32614/CRAN.package.matrixStats
MOFA2 v.1.18.0	Argelaguet et al. ¹²	https://biofam.github.io/MOFA2/
msigdb v.26.1.0	Dolgalev I.	10.32614/CRAN.package.msigdb
pheatmap v.1.0.13	Kolde R.	https://cran.r-project.org/web/packages/pheatmap/refman/pheatmap.html
randomForest v.4.7-1.2	Liaw A, Wiener M.	10.32614/CRAN.package.randomForest
rstatix v.0.7.3	Kassambara A.	10.32614/CRAN.package.rstatix
tidyverse v.2.0.0	Wickham W et al. ⁵⁹	https://tidyverse.tidyverse.org/

Method Details

Ex vivo metabolic and epigenetic drug screening.

Cryopreserved mononuclear (MNCs) of AML patients were thawed, resuspended in newborn calf serum (NCS) supplemented with DNase I (20 U/mL), 4 µM MgSO₄ and heparin (5 U/mL) and incubated for 15 min at 37°C. MNCs were either cultured in alpha-MEM with 25% fetal calf serum (FCS) and 1% penicillin/streptomycin (P/S, Gibco™) supplemented with G-CSF (Amgen), N-Plate (TPO)(Amgen) and IL-3 (Sandoz) (all 20 ng/mL) or co-cultured in Gartner's medium consisting of alpha-MEM (Lonza) supplemented with 12.5% FCS (Sigma-Aldrich), 12.5% heat-inactivated horse serum (Invitrogen), 1% P/S (Gibco™), 57.2 µM β-mercaptoethanol and 1 mM hydrocortisone (Sigma-Aldrich) and the above mentioned cytokines on

MS5, which were confluent plated on 0.1% gelatin-coated 96 wells plates and pre-treated with Mitomycin C. Primary AML samples were obtained after informed consent and approval by the Ethical Committee in accordance with the Declaration of Helsinki. The study was approved by the UMCG Medical Ethical Committee (protocol #NL43844.042.13). For this study, we used data from the OncoLifeS data-biobank for Oncology (Dutch Trial Register: NL7839). In the OncoLifeS data-biobank, all patients are included after written informed consent and the study protocol has been approved by the local Institutional Review Board (UMCG METc2010/109). MNCs were treated with DMSO or epigenetic (10 days) or metabolic inhibitors (4 days). Fresh medium and inhibitor were added at day 3/4 after demi-population of the cells for the epigenetic drug screen. For the metabolic screen, cells were stained on day 4 with CD45-PECy7, CD38-FITC, CD34-APC, AnnexinV-PE and DAPI; for the epigenetic screen, cells were stained on day 10 with CD45-PECy7, CD117-PE, CD11b-FITC, and DAPI. Both stainings were done in a 96-well plate, and incubated for 30 min at 4°C. Fluorescence measurements were taken using a MACSQuant® X Flow Cytometer (Miltenyi Biotec). The total list of 73 epigenetic inhibitors included in the screen is depicted in Supplemental Table 1 and end concentrations were 1 µM and 10 µM unless otherwise indicated. The total list of 42 metabolic inhibitors is listed in Supplemental Table 2 with respective end concentrations.

Multi-omics factor analysis (MOFA).

We trained a multi-omics factor analysis model in R using the MOFA2 package.¹³ Five data views were integrated, namely mutational status,^{14,15} proteomics (PXD030487),¹⁴ metabolomics (172 detected metabolites),¹⁵ and two independent drug screens (metabolic and epigenetic, 115 drugs). Label free quantitative proteome analysis table was filtered to maintain proteins that were called in at least 50% of the samples, and duplicates/ isoforms were removed based on the row with highest intensity across all samples. This led to a reduction in the 15612 detected proteins to 7992 proteins. No imputation method was used since MOFA is a probabilistic model and treats missing values as unobserved data points. Mutational status was modelled with a Bernoulli likelihood, while continuous layers were z-scored and modelled with default Gaussian likelihoods. Model training was performed using MOFA2 with a slow convergence mode for the final model. Factor interpretation relied on MOFA2 visualizations and ggplot2 package⁵⁷, and for proteomic weights, gene set enrichment analysis (GSEA) was performed by ranking genes by factor loadings and testing gene sets using the MSigDB and clusterProfiler packages.^{55,60}

Selection of number of factors and robustness analyses.

To select the number of latent factors (K), we trained models with 3 to 6 Ks and evaluated variance explained per view and factor, and robustness of factors. For robustness, we performed seed stability, by training multiple models per K across random seeds and aligning factors to reference models using correlations of factor scores. Additionally, we performed leave-one-out (LOO) stability, retraining the model after removing one sample at a time and quantifying the minimum absolute correlation of each factor's scores relative to the complete reference model.

Projection of MOFA factors to external independent cohorts.

We projected factor scores using a weighted sum of expression (WSE) approach onto multiple independent datasets, including three transcriptomes (BeatAML,⁶ TCGA-LAML,⁴ FPMTB⁷) and one proteome (Cheng et al.⁸). MOFA proteomics factor weights (loadings) were extracted from the original model and mapped to gene symbols. For each factor, weights were scaled to unit maximum absolute value. Expression matrices were harmonized to gene symbols by removing duplicate gene symbols (retaining the feature with the highest mean expression across samples), followed by gene-wise z-scoring across samples. For TCGA RNA-seq (RSEM) data, expression values were log-transformed ($\log_2(x+1)$) prior to gene-wise z-scoring to reduce skewness. Genes with zero variance across samples were assigned a z-score of 0 to ensure they contributed neutrally to the weighted sum projection. Projected factor scores were then computed for each sample and factor as the dot product between the scaled weight vector and the standardized expression vector over the overlapping genes. Finally, projected factor scores were z-scored across samples per factor to facilitate comparability across datasets.

Mutational analysis.

Factor-mutation associations were tested using Wilcoxon rank-sum tests, with multiple-testing correction by Benjamini-Hochberg FDR when applicable. To determine whether *NRAS/KRAS* wild-type (RAS^{wt}) or Factor 2-high (Factor 2^{high}) samples were enriched for additional genetic alterations beyond *NRAS/KRAS* mutations (RAS^{mut}), we performed a targeted enrichment analysis within BeatAML. Samples were dichotomized into either RAS^{wt} to evaluate differences between Factor 2^{high} vs Factor 2-low (Factor 2^{low}), using a threshold of 0 on the z-scored Factor 2 projection; or restricted to Factor 2^{high} samples to evaluate the differences between RAS^{mut} and RAS^{wt} . For each gene, mutations were compared between RAS^{wt} and RAS^{mut} groups using Fisher's exact test, reporting odds ratios (OR) and Benjamini-Hochberg adjusted q-values across genes. Results were summarized using a volcano-style plot of $\log_2(OR)$ versus $-\log_{10}(q)$. Where French American British (FAB) classification was available, factor score differences across FAB groups (M0 to M7) were assessed using a Kruskal-Wallis test.

Drug response association analysis.

To test whether projected factor scores were associated with drug response in external datasets, we correlated per-sample factor values with drug sensitivity metrics including area under the curve (AUC), inhibitory concentration at 50% (IC50), or selective drug sensitivity scores (sDSS). All drug response measures were harmonized so that higher values correspond to lower sensitivity prior to testing. For FPMTB's dataset that reported sDSS, scores were sign-adjusted to match the AUC/IC50 convention. For BeatAML, drug response data were filtered to include only converged dose-response curve fits with finite AUC values, and sample identifiers were matched between drug response and projected factor scores. Associations between Factor 2 and each drug were evaluated using Spearman correlation, requiring a minimum number of samples per drug prior to testing, and p-values were adjusted for multiple testing using the Benjamini-Hochberg procedure.

Validation of WSE-based factor projection.

To validate this approach, we applied the same procedure to the proteomics cohort used to train the MOFA model. Protein expression values were k-nearest neighbor (kNN) imputed, deduplicated by retaining the entry with the highest mean expression per protein, and protein-wise z-scored. WSE factor scores were computed for each sample, z-scored per factor and compared to the original MOFA factor scores by Spearman correlation across samples.

Single-sample gene set enrichment analysis.

Single-sample gene set enrichment analysis (ssGSEA) was performed using a curated set of gene sets especially focused on metabolism and cell differentiation state, producing per-sample enrichment scores that were subsequently z-scored across samples. ssGSEA z-scores were visualized in a heatmap, with samples annotated by Factor 2 projection scores (WSE-derived), *KRAS/NRAS* mutation status, and venetoclax AUC as continuous annotation. Samples were ordered by Factor 2 score to facilitate visualization of monotonic trends. On predefined *RAS*-related gene sets, we performed the analysis to compare *KRAS/NRAS* mutants and dichotomized by Factor 2 scores (high and low). Group comparisons were performed using non-parametric tests (Wilcoxon rank-sum; Kruskal-Wallis for multi-group comparisons), with multiple-testing correction.

Differential expression analysis.

To characterize gene expression differences associated with Factor 2 conditional on *RAS* status, samples were grouped into four groups defined by *KRAS/NRAS* status (RAS^{wt}/RAS^{mut}) and Factor 2 group (Factor 2^{high}/Factor 2^{low}). Differential expression was assessed using limma with a group-based design matrix and a set of contrasts within-status comparisons ($RAS^{wt}/\text{Factor 2}^{\text{high}}$ vs $RAS^{wt}/\text{Factor 2}^{\text{low}}$; $RAS^{mut}/\text{Factor 2}^{\text{high}}$ vs $RAS^{mut}/\text{Factor 2}^{\text{low}}$). Ranked gene lists from limma contrasts were used GSEA against MSigDB collections and gene ontology (GO) using PantherDB to summarize pathway-level changes associated with Factor 2 and/or *RAS* status.

Venetoclax response was modelled using linear regression with venetoclax AUC as the outcome. Projected Factor 2 scores (z-scored) derived from the WSE procedure were used as the primary predictor. Models were adjusted for *KRAS/NRAS* (*RAS*) status, *FLT3* and *TP53* mutation status (identified as poor responders to VEN in the ELN 2024 risk stratification⁶¹), and FAB morphology collapsed to a monocytic (M4/M5) vs non-monocytic (M0/M1/M2/M3/NOS) variable. Samples with missing key covariates (AUC or Factor 2) were excluded. Regression coefficients (β) with 95% confidence intervals and two-sided p-values were derived using bias-corrected heteroskedasticity-consistent standard errors (HC3). This approach was used because ex vivo AUC measurements often show substantial variability across samples, violating the assumption of constant error variance underlying standard linear regression. HC3 provides more reliable inference in moderate-sized, biologically heterogeneous datasets. Analyses were performed in the full cohort, in *RAS* wild-type samples only, and in *RAS*-mutant samples only; p-values across these three models were adjusted for multiple comparisons using the Benjamini-Hochberg procedure.

Graphical representation and statistical analysis. Graphs and statistical analyses were performed using R studio, using several packages including ggplot2,⁵⁷ ComplexHeatmap,⁶² and EnhancedVolcano. Statistical tests are specified in each figure legend.

Declarations

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Jan Jacob Schuringa (j.j.schuringa@umcg.nl).

Materials availability

This study did not generate new unique reagents.

Data and code availability

All data are available upon request to the lead contact, Jan Jacob Schuringa (j.j.schuringa@umcg.nl).

Acknowledgements

This work was supported by a Dutch Cancer Foundation grant (KWF-12910) and an EU PerMED/ZonMW grant (EPPERMED2024-048, MetaboTargetAML) awarded to J.J.S. We greatly appreciate the epigenetic drug library that was provided by dr. Ensar Halilovic (Novartis). We acknowledge dr. Paolo Gallipoli for providing the SSI-4 inhibitor.

Author contributions

D.S., S.M.H., L.L.O., N.D.L.R., D.A.P.M performed research; D.S. performed bioinformatic analysis; D.S., S.M.H., L.L.O., N.D.L.R., D.A.P.M, discussed and analyzed data; D.S. and J.J.S. conceptualized the study and wrote the manuscript; and all authors edited and approved the manuscript.

Declaration of Interests

The authors declare no competing financial interests.

References

1. Papaemmanuil, E. et al. Genomic classification and prognosis in acute myeloid leukemia. *New England Journal of Medicine* 374, 2209–2221 (2016).
2. Miles, L. A. et al. Single-cell mutation analysis of clonal evolution in myeloid malignancies. *Nature* 587, 477–482 (2020).
3. Morita, K. et al. Clonal evolution of acute myeloid leukemia revealed by high-throughput single-cell genomics. *Nat. Commun.* 11, (2020).
4. Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia. *New England Journal of Medicine* 368, 2059–2074 (2013).
5. Tyner, J. W. et al. Functional genomic landscape of acute myeloid leukaemia. *Nature* 562, 526–531 (2018).
6. Bottomly, D. et al. Integrative analysis of drug response and clinical outcome in acute myeloid leukemia. *Cancer Cell* 40, 850–864.e9 (2022).
7. Malani, D. et al. Implementing a Functional Precision Medicine Tumor Board for Acute Myeloid Leukemia. *Cancer Discov.* 12, 388–401 (2022).
8. Cheng, W.-Y. et al. Proteomic subtypes enrich current acute myeloid leukemia nomenclature and reflect intrinsic pathogenesis alongside aging. *Blood* 146, 2681–2695 (2025).
9. Kramer, M. H. et al. Proteomic and phosphoproteomic landscapes of acute myeloid leukemia. *Blood* 140, 1533–1548 (2022).
10. Jayavelu, A. K. et al. The proteogenomic subtypes of acute myeloid leukemia. *Cancer Cell* 40, 301–317.e12 (2022).
11. De Jonge, H. J. M. et al. Gene expression profiling in the leukemic stem cell-enriched CD34 fraction identifies target genes that predict prognosis in normal karyotype AML. *Leukemia* 25, 1825–1833 (2011).
12. Argelaguet, R. et al. Multi-Omics Factor Analysis—a framework for unsupervised integration of multi-omics data sets. *Mol. Syst. Biol.* 14, (2018).
13. Argelaguet, R. et al. MOFA+: A statistical framework for comprehensive integration of multi-modal single-cell data. *Genome Biol.* 21, (2020).
14. Hogeling, S. M. et al. Bleximenib, the novel menin-KMT2A inhibitor JNJ-75276617, impairs long-term proliferation and immune evasion in acute myeloid leukemia. *Haematologica* 110, 1278–1291 (2025).
15. Sternadt, D. et al. Metformin induces ferroptosis associated with lipidomic remodeling in AML. *Blood Adv.* 10, 733–747 (2026).
16. Erdem, A. et al. The Glycolytic Gatekeeper PDK1 defines different metabolic states between genetically distinct subtypes of human acute myeloid leukemia. *Nat. Commun.* 13, (2022).
17. Pereira-Martins, D. A. et al. High mtDNA content identifies oxidative phosphorylation-driven acute myeloid leukemias and represents a therapeutic vulnerability. *Signal Transduct. Target. Ther.* 10, (2025).
18. Sabatier, M. et al. C/EBP α Confers Dependence to Fatty Acid Anabolic Pathways and Vulnerability to Lipid Oxidative Stress–Induced Ferroptosis in FLT3-Mutant Leukemia. *Cancer Discov.* 13, 1720–1747 (2023).
19. Chowdhury, R. et al. The oncometabolite 2-hydroxyglutarate inhibits histone lysine demethylases. *EMBO Rep.* 12, 463–469 (2011).
20. Xu, W. et al. Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α -ketoglutarate-dependent dioxygenases. *Cancer Cell* 19, 17–30 (2011).
21. Ward, P. S. et al. The Common Feature of Leukemia-Associated IDH1 and IDH2 Mutations Is a Neomorphic Enzyme Activity Converting α -Ketoglutarate to 2-Hydroxyglutarate. *Cancer Cell* 17, 225–234 (2010).
22. Gu, Y. et al. IDH1 mutation contributes to myeloid dysplasia in mice by disturbing heme biosynthesis and erythropoiesis. *Blood* 137, 945–958 (2021).
23. Gröschel, S. et al. High EVI1 expression predicts outcome in younger adult patients with acute myeloid leukemia and is associated with distinct cytogenetic abnormalities. *Journal of Clinical Oncology* 28, 2101–2107 (2010).

24. Richard-Carpentier, G. *et al.* Characteristics and clinical outcomes of patients with acute myeloid leukemia with inv(3)(q21q26.2) or t(3;3)(q21;q26.2). *Haematologica* 108, 2331–2342 (2023).
25. Zhang, H. *et al.* Integrated analysis of patient samples identifies biomarkers for venetoclax efficacy and combination strategies in acute myeloid leukemia. *Nat. Cancer* 1, 826–839 (2020).
26. Sango, J. *et al.* RAS-mutant leukaemia stem cells drive clinical resistance to venetoclax. *Nature* 636, 241–250 (2024).
27. Fruchtman, H., Avigan, Z. M., Waksal, J. A., Brennan, N. & Mascarenhas, J. O. Management of isocitrate dehydrogenase 1/2 mutated acute myeloid leukemia. *Leukemia* vol. 38 927–935 Preprint at <https://doi.org/10.1038/s41375-024-02246-2> (2024).
28. Venugopal, S. *et al.* Decitabine and venetoclax for IDH1/2-mutated acute myeloid leukemia. *American Journal of Hematology* vol. 96 E154–E157 Preprint at <https://doi.org/10.1002/ajh.26122> (2021).
29. Pollyea, D. A. *et al.* Impact of Venetoclax and Azacitidine in Treatment-Naïve Patients with Acute Myeloid Leukemia and IDH1/2 Mutations. *Clinical Cancer Research* 28, 2753–2761 (2022).
30. Konopleva, M. *et al.* Analysis of Patients With Monocytic and Monocytic-Like Acute Myeloid Leukemia, Including AML-M4 and AML-M5, Treated With Venetoclax Plus Azacitidine. *American Journal of Hematology* Preprint at <https://doi.org/10.1002/ajh.70161> (2026).
31. Pei, S. *et al.* A Novel Type of Monocytic Leukemia Stem Cell Revealed by the Clinical Use of Venetoclax-Based Therapy. *Cancer Discov.* 13, 2032–2049 (2023).
32. Döhner, H. *et al.* Genetic risk stratification and outcomes among treatment-naïve patients with AML treated with venetoclax and azacitidine. *Blood* 144, 2211–2222 (2024).
33. Yehudai, D. *et al.* The thymidine dideoxynucleoside analog, alovudine, inhibits the mitochondrial DNA polymerase γ , impairs oxidative phosphorylation and promotes monocytic differentiation in acute myeloid leukemia. *Haematologica* 104, 963–972 (2019).
34. Gyamfi, J. *et al.* Interaction between CD36 and FABP4 modulates adipocyte-induced fatty acid import and metabolism in breast cancer. *NPJ Breast Cancer* 7, (2021).
35. Tabe, Y., Konopleva, M. & Andreeff, M. Fatty Acid Metabolism, Bone Marrow Adipocytes, and AML. *Frontiers in Oncology* vol. 10 Preprint at <https://doi.org/10.3389/fonc.2020.00155> (2020).
36. Forte, D. *et al.* Bone Marrow Mesenchymal Stem Cells Support Acute Myeloid Leukemia Bioenergetics and Enhance Antioxidant Defense and Escape from Chemotherapy. *Cell Metab.* 32, 829–843.e9 (2020).
37. Azadniv, M. *et al.* Bone marrow mesenchymal stromal cells from acute myelogenous leukemia patients demonstrate adipogenic differentiation propensity with implications for leukemia cell support. *Leukemia* 34, 391–403 (2020).
38. Shafat, M. S. *et al.* Leukemic blasts program bone marrow adipocytes to generate a protumoral microenvironment. *Blood* 129, 1320–1332 (2017).
39. Tabe, Y. *et al.* Bone marrow adipocytes facilitate fatty acid oxidation activating AMPK and a transcriptional network supporting survival of acute monocytic leukemia cells. *Cancer Res.* 77, 1453–1464 (2017).
40. Tabe, Y. *et al.* Inhibition of FAO in AML co-cultured with BM adipocytes: mechanisms of survival and chemosensitization to cytarabine. *Sci. Rep.* 8, (2018).
41. Valk, P. J. *et al.* Prognostically Useful Gene-Expression Profiles in Acute Myeloid Leukemia. *N Engl J Med* vol. 350 www.hovon.nl (2004).
42. Ma, S. *et al.* Moving towards genome-wide data integration for patient stratification with Integrate Any Omics. *Nat. Mach. Intell.* 7, 29–42 (2025).
43. Farge, T. *et al.* Chemotherapy-resistant human acute myeloid leukemia cells are not enriched for leukemic stem cells but require oxidative metabolism. *Cancer Discov.* 7, 716–735 (2017).
44. Stuani, L. *et al.* Mitochondrial metabolism supports resistance to IDH mutant inhibitors in acute myeloid leukemia. *Journal of Experimental Medicine* 218, (2021).

45. Gunn, K. et al. (R)-2-Hydroxyglutarate Inhibits KDM5 Histone Lysine Demethylases to Drive Transformation in IDH-Mutant Cancers. *Cancer Discov.* 13, 1478–1497 (2023).
46. Figueroa, M. E. et al. Leukemic IDH1 and IDH2 Mutations Result in a Hypermethylation Phenotype, Disrupt TET2 Function, and Impair Hematopoietic Differentiation. *Cancer Cell* 18, 553–567 (2010).
47. Zhang, Q. et al. Activation of RAS/MAPK pathway confers MCL-1 mediated acquired resistance to BCL-2 inhibitor venetoclax in acute myeloid leukemia. *Signal Transduct. Target. Ther.* 7, (2022).
48. Nwosu, G. O., Ross, D. M., Powell, J. A. & Pitson, S. M. Venetoclax therapy and emerging resistance mechanisms in acute myeloid leukaemia. *Cell Death and Disease* vol. 15 Preprint at <https://doi.org/10.1038/s41419-024-06810-7> (2024).
49. Pei, S. et al. Monocytic subclones confer resistance to venetoclax-based therapy in patients with acute myeloid leukemia. *Cancer Discov.* 10, 536–551 (2020).
50. Wang, T. et al. Sequential CRISPR gene editing in human iPSCs charts the clonal evolution of myeloid leukemia and identifies early disease targets. *Cell Stem Cell* 28, 1074–1089.e7 (2021).
51. Dembitz, V. et al. Stearoyl-CoA desaturase inhibition is toxic to acute myeloid leukemia displaying high levels of the de novo fatty acid biosynthesis and desaturation. *Leukemia* 38, 2395–2409 (2024).
52. Barbie, D. A. et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* 462, 108–112 (2009).
53. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. Circlize implements and enhances circular visualization in R. *Bioinformatics* 30, 2811–2812 (2014).
54. Hornik, K. A CLUE for CLUster Ensembles. *JSS Journal of Statistical Software* vol. 14 <http://www.jstatsoft.org/> (2005).
55. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. ClusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287 (2012).
56. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32, 2847–2849 (2016).
57. Wickham, H. *Ggplot2*. (Springer International Publishing, Cham, 2016). doi:10.1007/978-3-319-24277-4.
58. Ritchie, M. E. et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43, e47 (2015).
59. Wickham, H. et al. Welcome to the Tidyverse. *J. Open Source Softw.* 4, 1686 (2019).
60. Subramanian, A. et al. *Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles*. www.pnas.org/cgi/doi/10.1073/pnas.0506580102 (2005).
61. Döhner, H. et al. Genetic risk classification for adults with AML receiving less-intensive therapies: the 2024 ELN recommendations. *Blood* 144, 2169–2173 (2024).
62. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32, 2847–2849 (2016).

Figures

Figure 1

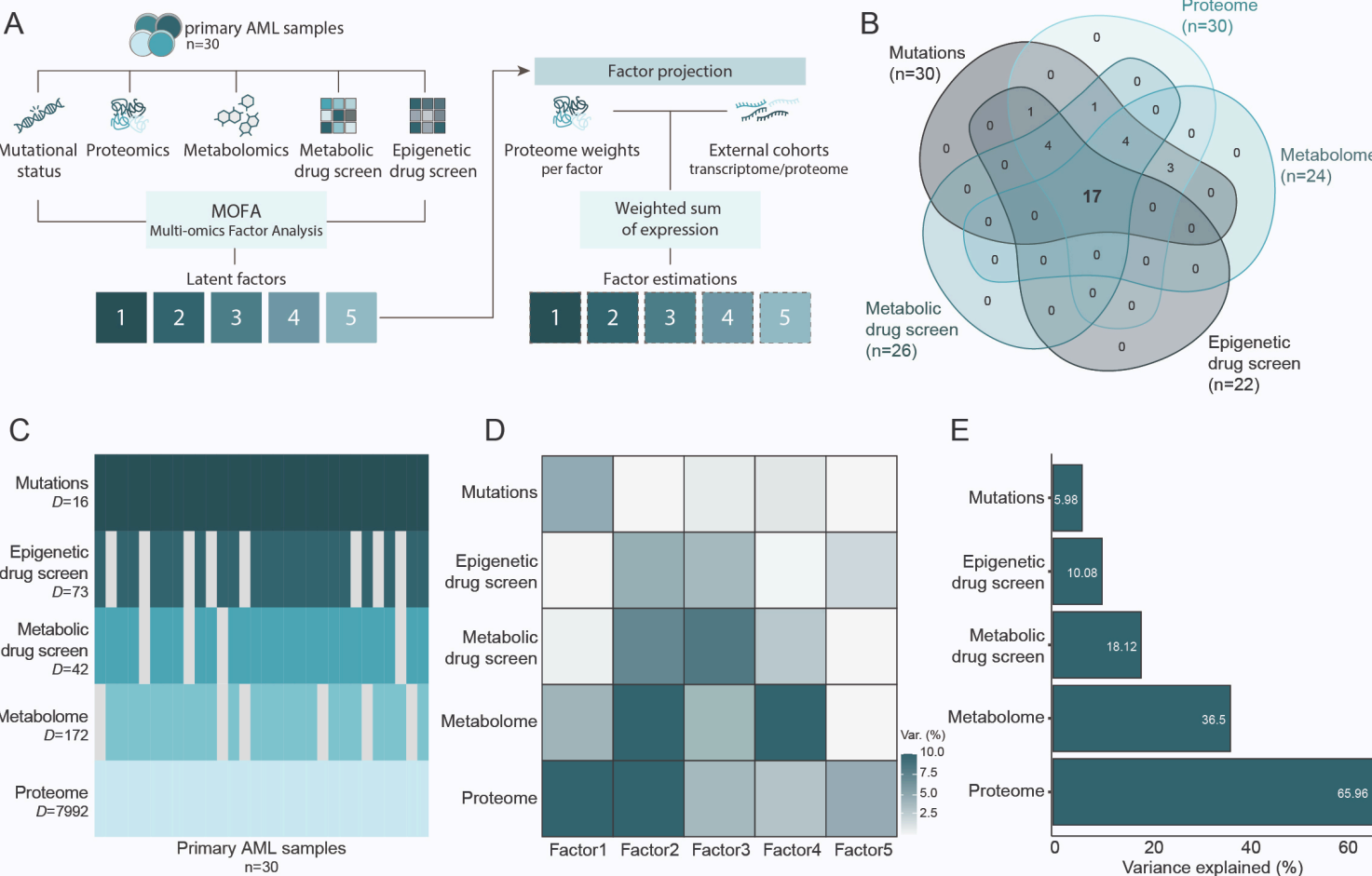


Figure 1

Overview of the multi-omics integration strategy and dataset structure.

(A) Schematic overview of the analytical strategy for multi-layer profiling of 30 primary AML samples using multi-omics factor analysis (MOFA). Proteome-based factor weights were subsequently extrapolated to external cohorts for biological and clinical validation. (B) Venn diagram representing the overlap between the five data layers, illustrating the number of samples were profiled in each combination of modalities. (C) Overview of the data types (in rows, D = number of features) and samples (in columns). Missing samples are represented in grey. (D) Heatmap depicting the proportion of variance explained by the MOFA model across factors and views. (E) Bar plot summarizing the total variance explained by each latent factor across all modalities. Numbers displayed in bars indicate total R^2 , rounded to two decimals.

Figure 2

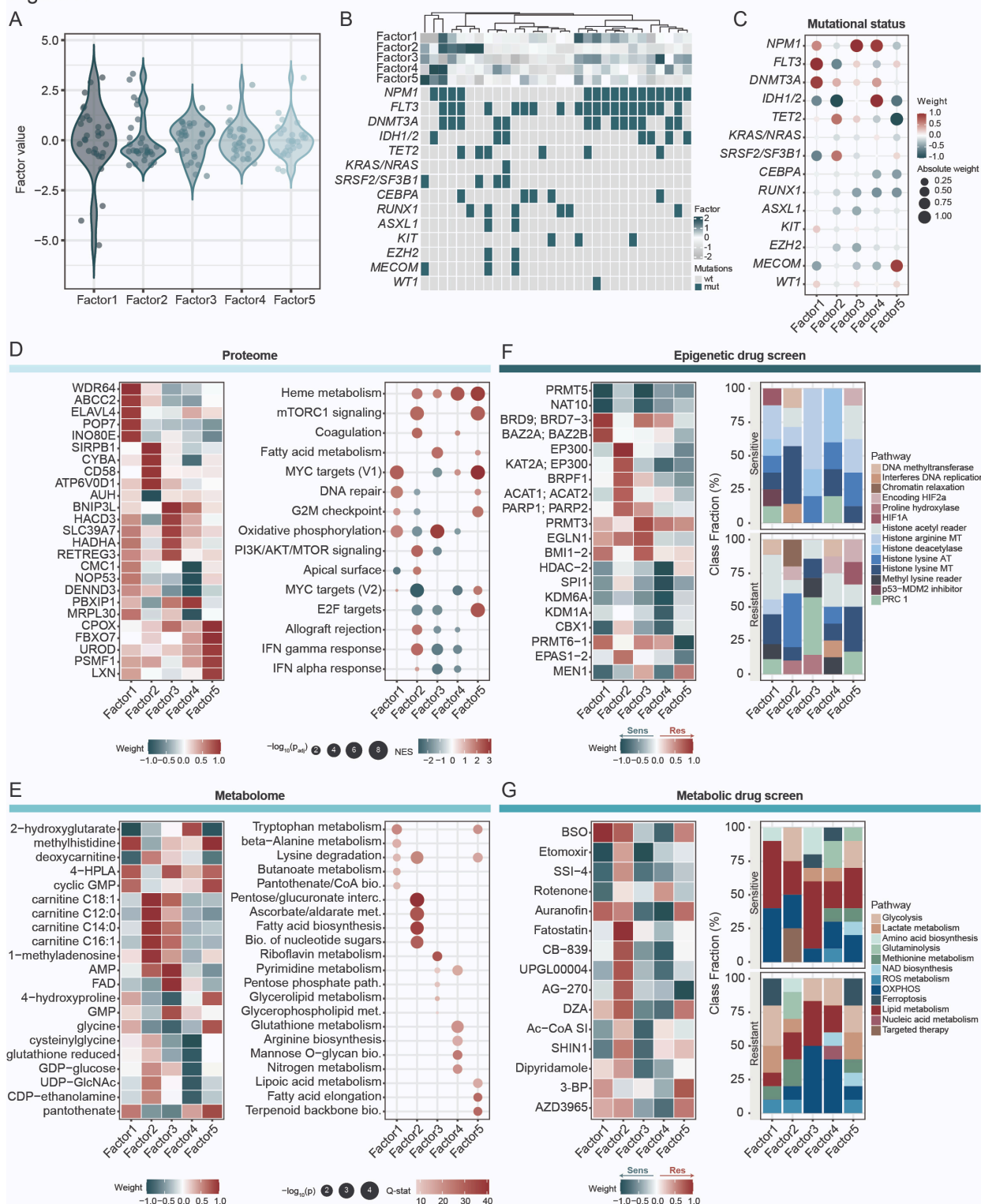


Figure 2

Characterization of MOFA latent factors linking molecular programs, mutations and drug responses. (A) Violin plots show the distribution of factor scores (z-scaled) per factor across samples. (B) Oncoprint of the mutational landscape from the samples used to train the MOFA model with z-scored factor scores (continuous color scale) as top annotation. (C) Bubble plot depicting mutation loadings by factor. The color encodes the scaled loading (value/max absolute value per factor); and size, the absolute scaled loading. Positive/negative values indicate opposing contributions of a mutation to the factor. (D-G) Top loadings heatmaps and enrichment summaries across views. For each view, left subpanels show heatmaps display

loadings normalized per factor by dividing by max absolute weight. Right subpanels summarize enrichment or class composition derived from the same factor-specific top features (D) *Proteomics*. Heatmap of top protein loadings per factor (scaled per factor) and gene set enrichment analysis (GSEA) dot plot showing the top pathways (top n) per factor; dot size = $-\log_{10}(p)$, color = normalized enrichment score (NES; blue/red as above).

(E) *Metabolomics*. Heatmap of top metabolite loadings per factor and MetaboAnalyst-derived pathway bubble plot displaying top metabolic pathways per factor (dot size: $-\log_{10}(p)$; color: Q-statistic). Abbreviations: 4-HPLA: 4-hydroxyphenyllactate; Interc.: interconversions; Bio.: biosynthesis; Met.: metabolism. (F) *Epigenetic drug screen*. Heatmap of selected inhibitors' loadings per factor (scaled) and stacked composition bars showing the fraction of drug classes represented among the top absolute weights per factor, separated based on sign of the scaled loading (positive = resistant; negative = sensitive). Bars are normalized per factor to allow comparison of class composition. Abbreviations: MT: methyltransferase; AT: acetyltransferase; PRC: Polycomb repressive complex. (G) *Metabolic drug screen*. Heatmap of top metabolic inhibitors per factor and class composition bars as in (F), summarizing the top absolute weights per factor by drug class in the sensitive vs resistant classifications.

Figure 3

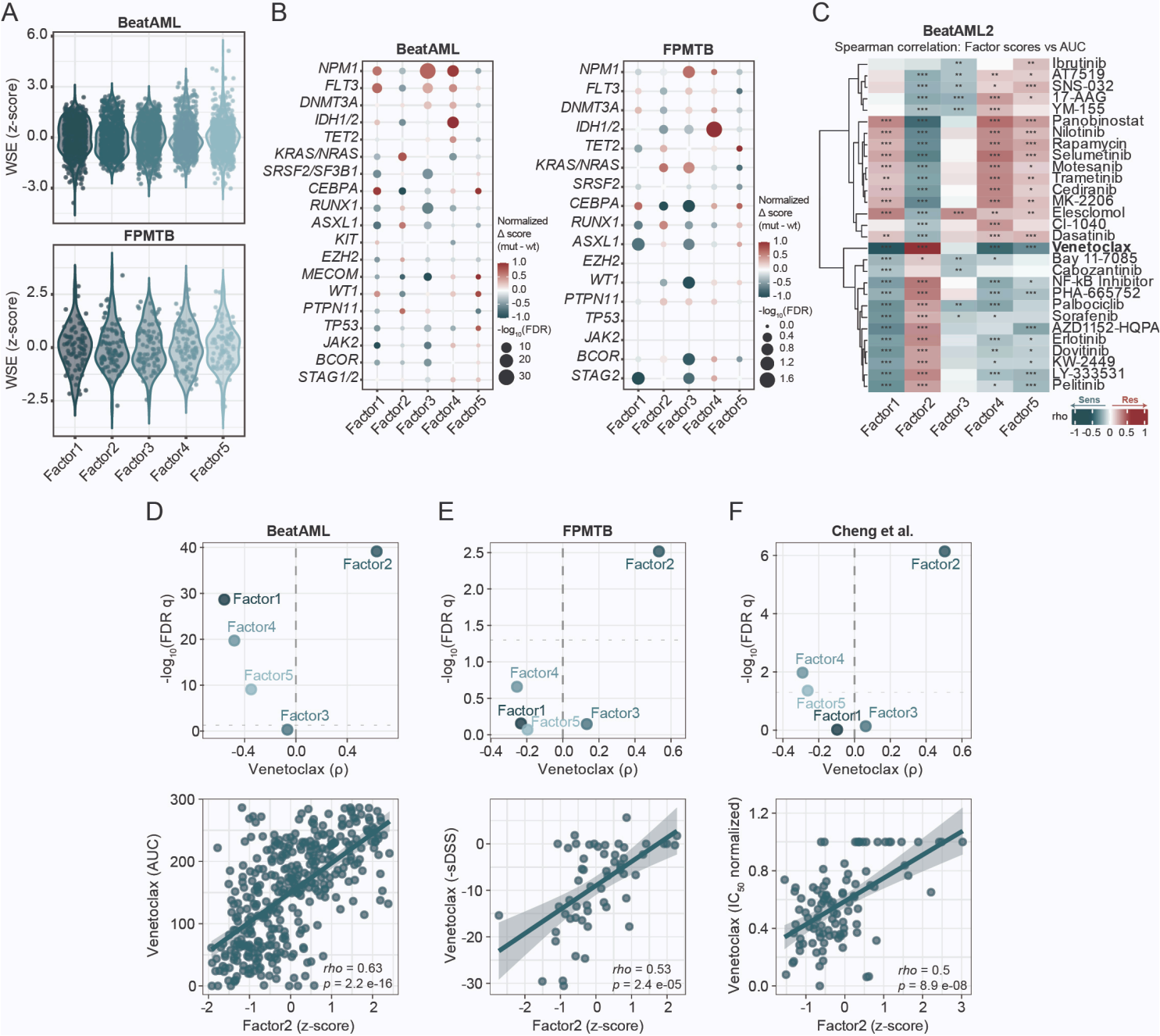


Figure 3

Proteome-derived MOFA factor projection to external cohorts links mutations and drug sensitivities. (A) Violin plots depicting weighted sum of expression (WSE) computed using the MOFA weights from the proteomic view extrapolated for the BeatAML and FPMTB dataset. Weights are scaled per factor, and the resulting scores are z-scored per factor across samples. (B) Bubble plot summarizing associations between factor signature scores and gene mutation status. For each factor x gene, a Wilcoxon rank-sum test compares scores in mutant vs wild-type samples (only genes with ≥ 3 mutants are tested). Dot size encodes $-\log_{10}(\text{FDR})$ (Benjamini-Hochberg per factor), and dot color encodes the direction and magnitude of the scaled mean difference. (C) Heatmap of Spearman correlations between factor signature z-scores and drug AUC (area under the curve) across inhibitors. For each factor, correlations are computed per inhibitor (minimum $n \geq 8$), with BH-FDR per factor; shown are the top 10 inhibitors per factor by FDR. Significance stars represent $q < 0.05$ (*), < 0.01 (**), < 0.001 (***) (D-F) For each cohort (D) BeatAML, (E) FPMTB, (F) Cheng et al.: on top, a volcano plot depicts the results of Spearman correlation (ρ) between factor scores and venetoclax AUC and respective $\log_{10}(\text{BH-FDR } q)$. Bottom plots with a

scatter plot of venetoclax AUC versus the per-sample Factor 2 signature z-scores, with a linear trend and Spearman ρ annotated.

Figure 4

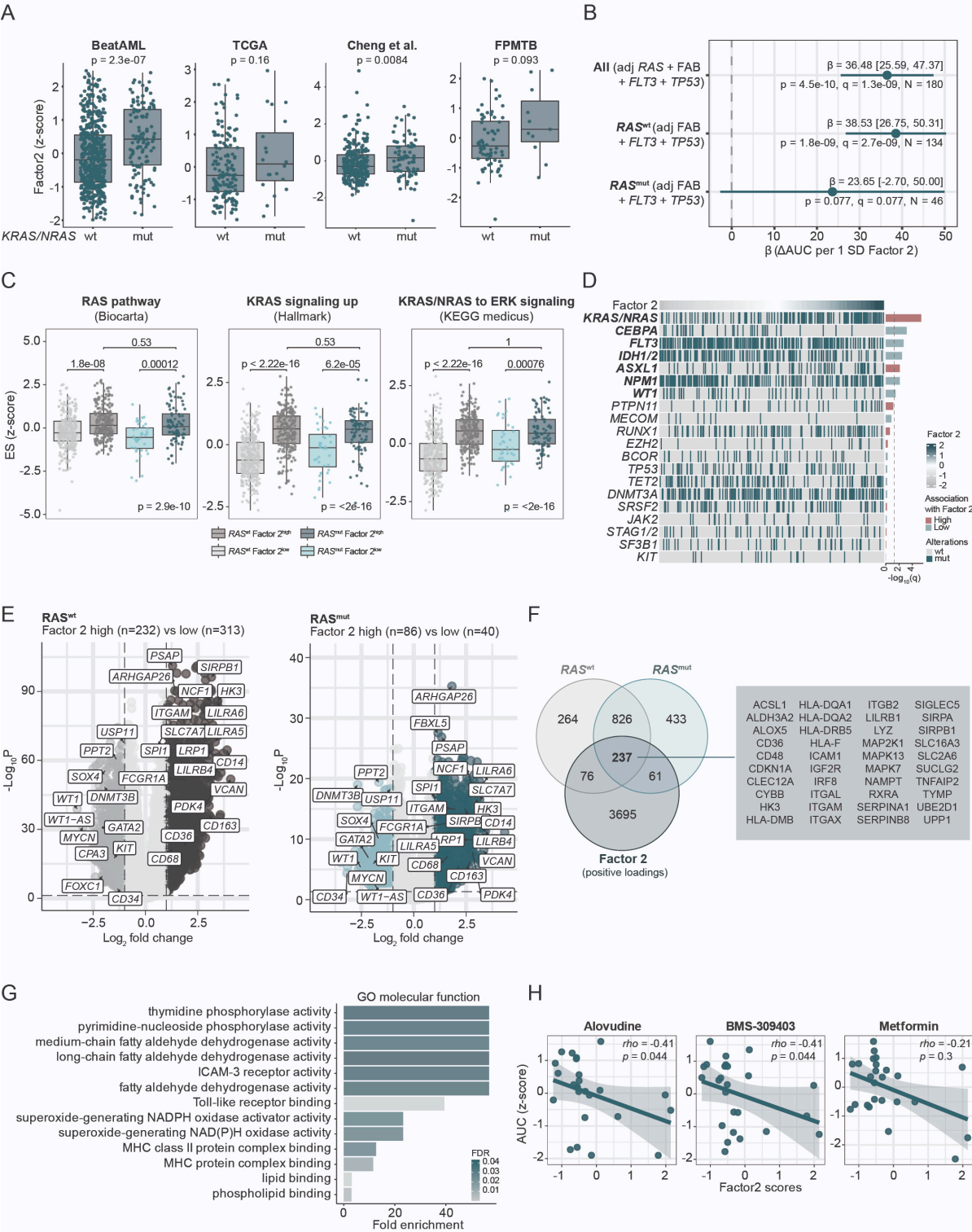


Figure 4

Factor 2 identifies a *RAS*-mutant-like subgroup with venetoclax resistance and metabolic vulnerabilities. (A) Boxplots showing Factor 2 scores in *RAS* wild-type (*RAS*^{wt}) versus mutant (*RAS*^{mut}) samples across four independent AML cohorts, clearly depicted in the graphs. Statistical analysis was done using Wilcoxon test and p-values are stated in the plot. (B) Forest plot shows the β coefficient ($\pm 95\%$ CI) for Factor 2 from linear models of venetoclax AUC, fitted in all samples

(adjusted for *RAS* status, FAB monocytic group, *FLT3*, *TP53*), and separately in *RAS*^{wt} and *RAS*^{mut} subsets (adjusted for FAB monocytic group, *FLT3*, *TP53*); β represents the adjusted change in venetoclax AUC per 1-standard-deviation increase in Factor 2; positive β values indicate higher AUC and reduced venetoclax sensitivity. (C) Single sample gene set enrichment analysis (ssGSEA) enrichment scores (z-scored) for *RAS*/MAPK gene signatures for samples stratified by *RAS* genotype and Factor 2 high (Factor 2^{high}) vs low (Factor 2^{low}) (Factor 2 > 0 vs $\leq$ 0) using the BeatAML dataset. Statistical analysis between groups were assessed using a Kruskal-Wallis test and post hoc pairwise Wilcoxon tests. Adjusted p-values (Benjamini-Hochberg) are depicted in the plots. (D) Oncoprint from the BeatAML cohort ranked on Factor 2 scores (z-score). Genes marked in bold represent statistically significant association between gene and Factor 2. Right-side bar plot depict the magnitude of the significance, expressed as -log₁₀ q-values from Wilcoxon rank-sum tests (adjusted using Benjamini-Hochberg), with bar color indicating whether mutations are associated with higher (red) or lower (teal) Factor 2 scores. (E) Volcano plots depicting differential expression for Factor 2^{high} vs Factor 2^{low} comparisons performed within *RAS*^{wt} and *RAS*^{mut} samples using the BeatAML cohort. Color coding is defined by genes significantly (adj. p-value < 0.05) upregulated (log₂FC > 1, darker tones) and downregulated (log₂FC < -1, lighter tones). (F) Venn diagram depicting the intersection between genes upregulated in Factor 2^{high} vs Factor 2^{low} from both *RAS*^{wt} and *RAS*^{mut} analysis and genes with positive Factor 2 loadings. Box show some of the genes found in the intersection between the 3 groups. (G) Gene Ontology (GO) enrichment analysis using the molecular function subset of the overlapping genes in the venn diagram. (H) Scatter plot of Spearman correlations between Factor 2 scores and drug response (z-scores) from the metabolic drug screen performed in liquid and MS5 co-culture (day 4). Rho and p-values are depicted in the plots for each condition.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SternadtetalMOFASupplementalTables.xlsx](#)
- [SternadtetalMOFASupplementalFigurelegends.pdf](#)