MIRACS: An Unsupervised Multi-Omics Strategy for Identifying Patterns of Therapeutic Resistance in Cancer.

Sara González-Carro and Jose Linares-Blanco

Machine Learning Lab in Life Sciences, CITIC, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: j.linares@udc.es

DOI: *https://doi.org/10.17979/spu.23.c52*

Abstract: Drug resistance is a major challenge in cancer therapy, often driven by poorly characterized molecular states. We developed MIRACS (Multiomic Integration for Resistance Association via Correlation Strategy), an unsupervised multi-omics framework to identify resistance programs across diverse cancer types and drug classes. Transcriptomic and proteomic profiles from 370 cell lines (22 tumor types, DepMap) were decomposed into 30 latent factors using Multi-Omics Factor Analysis (MOFA). Correlating these factors with drug response data from the PRISM dataset revealed over 700 significant factor–mechanism of action (MoA) associations across 10 cancer types.
Among these, a molecular program emerged as a cross-lineage resistance program linked to five drug classes—including topoisomerase, Bcr-Abl kinase, Aurora kinase, DNA synthesis, and HSP inhibitors—across eight cancer types. Functional analyses of gene and protein loadings highlighted stress response and inflammatory pathways (p53, NF-κB, TNF-α, JAK–STAT) and transcriptional regulators ATF6 and IRF1, consistent with a pre-existing stress-adapted state.
These results demonstrate that integrating multi-omics with drug perturbation data can uncover convergent, lineage-independent resistance programs, providing a systematic approach to identify targetable vulnerabilities across cancers.

1 Introduction

Drug resistance is a major obstacle in cancer therapy, often leading to treatment failure and disease progression. Resistance arises from heterogeneous molecular states, many of which remain poorly characterized (Vasan et al., 2019). Most studies focus on single tumor types or individual drugs, but this narrow scope risks overlooking lineage-independent resistance programs that converge across cancers. Previous studies suggested that resistant states can be driven by common molecular mechanisms that transcend tumor lineage and therapeutic class (Gu et al., 2025; Lei et al., 2023).

To investigate this, we developed MIRACS (Multiomic Integration for Resistance Association via Correlation Strategy) a four-step computational strategy that integrates multi-omic profiles with drug response data.

2 Materials and Methods

2.1 Datasets

Datasets were obtained from DepMap (DepMap, Broad, 2024). Transcriptomic data consisted of RNA-seq ($\log_2$(TPM+1)) for 19,138 genes in 1,673 cell lines, while proteomic data (Nusinow et al., 2020) included mass spectrometry-based quantification of 12,196 genes in 375 cell lines. After harmonization and duplicate filtering, 370 cell lines with both data types were retained.

Drug response data were obtained from PRISM Repurposing (19Q4), covering 1,502 compounds across classified in 444 MoAs (Corsello et al., 2019). Replicates were filtered, and z-score normalization was applied per drug.

2.2 MIRACS: *Multiomic Integration for Resistance Association via Correlation Strategy*

The workflow consists of four phases and starts from four matrices: Y_m, multi-omic data matrices (each represented as a view); R, drug response matrix (AUC per cell line); a matrix of cell line group annotations; and drug annotations (MoAs), grouping drugs by mechanism of action.

First, different omics (Y_m) are integrated across cell lines using Multi-Omics Factor Analysis (MOFA). MOFA yields latent factors that capture axes of variability, providing sample-level factors in a matrix Z for each group, and feature weights in matrices W (one per view) for each omic layer. Second, Spearman correlations are computed between latent factors (Z_g) and drug responses (R), yielding a correlation matrix C_g. Third, an enrichment analysis based on drug mechanisms of action (MoAs) evaluates whether a factor is consistently associated with drugs sharing the same MoA. This GSEA-like approach groups drugs by MoA and tests whether their correlations with a given factor show a consistent pattern, enabling the detection of programs shared across multiple drugs with the same MoA, even if not universal. The result is a matrix of enrichment scores (E) and adjusted p-values (p_{adj}), from which significant MoA–factor associations can be selected. These associations can then be biologically interpreted by exploring the weight matrices W to identify genes, proteins, or other features that contribute most to the variability captured by each factor.

2.3 MOFA modeling

Multi-Omics Factor Analysis (MOFA) is a statistical framework designed for unsupervised integration of heterogeneous omics datasets (Argelaguet et al., 2018). It decomposes high-dimensional data into a reduced number of latent factors that capture major sources of variability across different molecular layers. Each factor reflects shared or view-specific patterns of variability, while the weight matrices W_m indicate the contribution of individual features (genes, proteins, etc.) to each factor.

We applied MOFA2 (`mofapy2`, Python implementation) to jointly model transcriptomic and proteomic data. Samples were grouped by tumor lineage to minimize factors dominated by inter-tissue differences. The model was trained with 30 latent factors.

2.4 MoA enrichment analysis

To link latent factors with drug response, we applied an enrichment strategy inspired by Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005). Instead of genes, the ranked elements were drug–factor correlation scores, grouped by mechanism of action (MoA). This approach tests whether drugs in the same MoA show consistent correlations with a factor, identifying latent drivers of sensitivity or resistance..

Correlations between each factor (Z_g) and drug response (AUC values) were computed per tissue using Spearman correlation, generating correlation matrices C_g. These matrices were analyzed with the `decoupler` package (v2.0.4), running 10,000 permutations to estimate normalized enrichment scores (NES) and adjusted p-values. Positive enrichments (high factor values associated with high AUC, i.e. resistance) and negative enrichments (high factor values with low AUC, i.e. sensitivity) were used to prioritize factor–MoA associations consistent across tissues.

3 Results

MIRACS decomposed transcriptomic and proteomic profiles into 30 latent programs, explaining 34% and 27% of their variability, respectively (Fig.1a). Spearman correlations between these programs and drug responses showed a mean absolute correlation of 0.2, with maximum absolute correlations reaching up to 0.96 (Fig.1b).

Drug enrichment analysis identified over 700 significant associations between latent factors and MoAs across 10 cancer types. Molecular programs related to resistance are those where high factor values associate with high AUC (resistant cells), resulting in positive correlation and positive enrichment. We prioritized latent factors consistently linked to resistance for the same MoA across multiple tumor types (Fig.1c). These cross-lineage associations highlighted recurrent resistance programs and motivated the selection of Factor 12 for detailed characterization.

Factor 12 was associated with resistance to multiple drug classes—including topoisomerase inhibitors, Bcr-Abl kinase inhibitors, Aurora kinase inhibitors, DNA synthesis inhibitors, and HSP inhibitors—across eight cancer types (Table1). Functional analyses of its transcriptomic and proteomic loadings revealed enrichment for stress response and inflammatory pathways (Ak and Levine, 2010). Gene-level enrichment highlighted p53 signaling, NF-κB, TNF-α, and JAK–STAT pathways in the transcriptomic component.

Additionally, transcription factor activity inference indicated overactivation of ATF6 and IRF1 across both omics layers (Fig.2 b), consistent with a pre-existing stress-adapted and inflammatory transcriptional program Stengel et al. (2020); Takaoka et al. (2008).

Among the MoAs asociated with factor 12 we can choose one, i.e. Aurore kinase inhibitors. To evaluate which drug responses within this selected MoA are influenced by factor 12, a forest plots summarizing individual drug–factor correlations was generated (Fig. 2c). Most drugs showed positive associations, although variability across tumor types was evident. Subsequently, Factor 12 values were visualized separately for each tissue, comparing sensitive and resistant cells (binarized between the 33rd and 67th percentiles). Fig. 2d–h illustrates this for the Aurora kinase inhibitor tozasertib, highlighting consistent tendency across multiple tumor types.

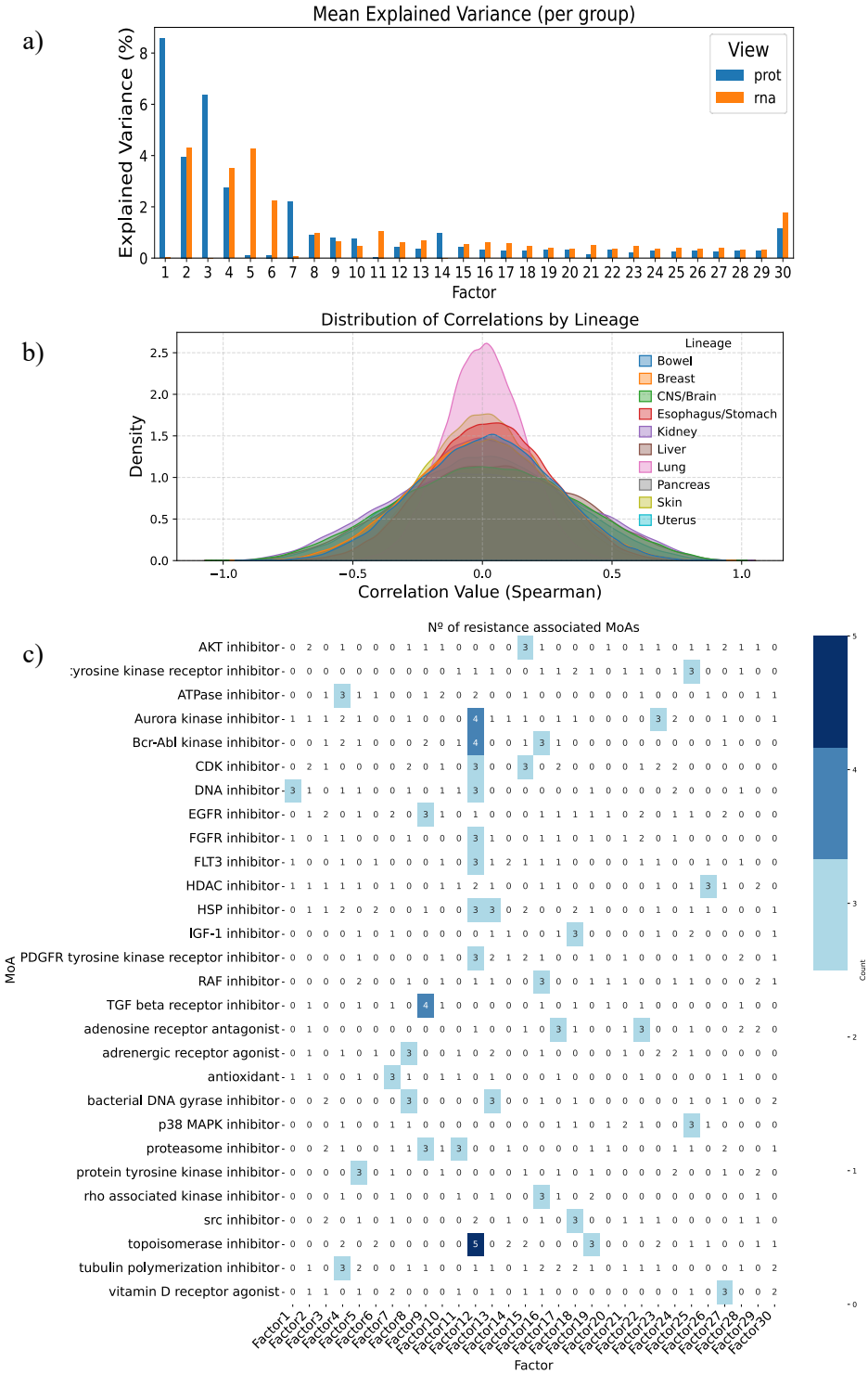

Figure 1: MIRACS applied to DepMap data (Phases 1–3). a) Mean R^2 by factor and view (RNA or protein) across groups (tumor type). b) Distribution of Spearman correlation of factors and drugs by group (tumor type). c) Number of positive enriched factors by MoA (highlighting those present in > 3 tissues).

Table 1: Selected Factor 12–MoA associations.

MoA	Number of tissues	Tissues
Topoisomerase inhibitor	5	Breast, Esophagus/Stomach, Kidney, Skin, Uterus
Bcr-Abl kinase inhibitor	4	Intestine, Esophagus/Stomach, Kidney, Lung
Aurora kinase inhibitor	4	Intestine, Lung, Pancreas, Skin
DNA inhibitor	3	Intestine, Lung, Uterus
HSP inhibitor	3	Kidney, Pancreas, Skin

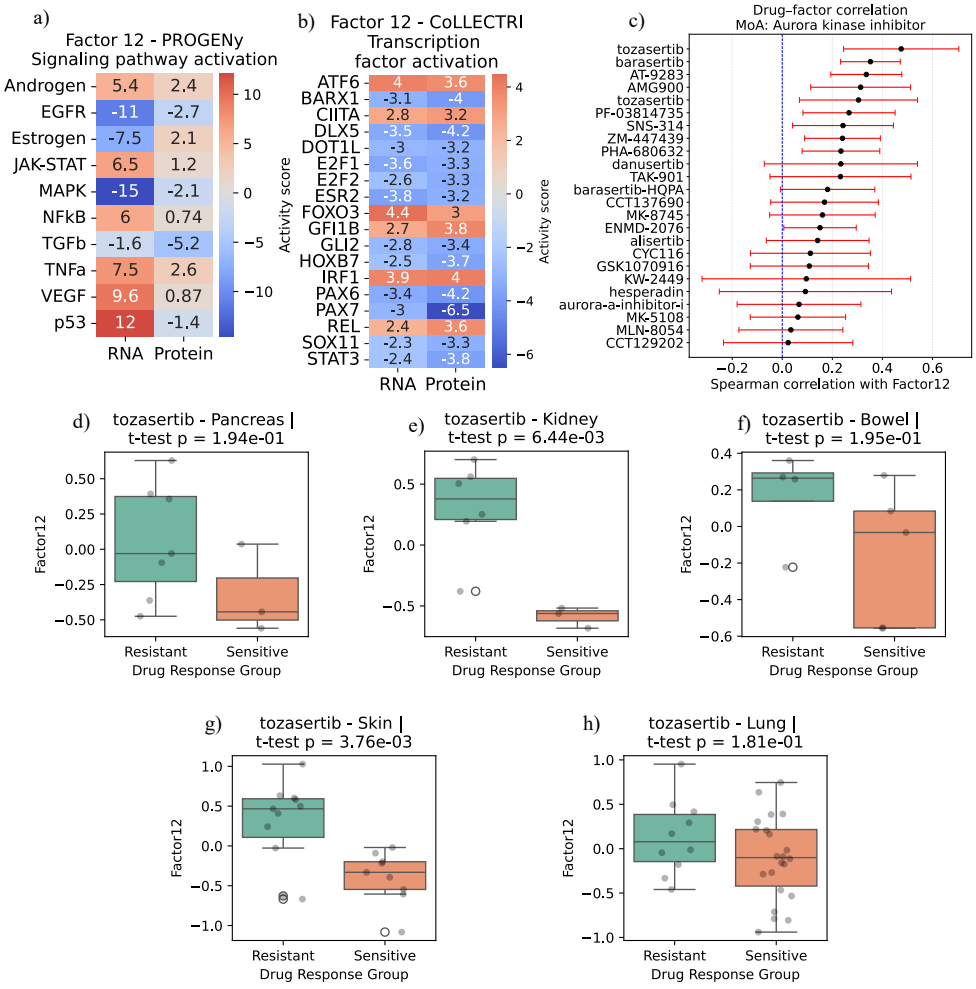

Figure 2: MIRACS applied to DepMap (Phase 4). Example analysis of factor 12. a) Functional analysis of factor 12 loadings with PROGENy pathways. b) Functional analysis of factor 12 loadings with CollecTRI transcription factors. c) Correlation of each Aurora kinase inhibitor drug with factor 12 in selected tissues. d–h) Comparison of factor 12 values for cell lines sensitive and resistant to tozasertib in cell lines from tumours of the pancreas (d), kidney (e), intestine (f), skin (g), and lung (h).

4 Discussion

MIRACS enables unsupervised, interpretable identification of shared and tissue-specific resistance programs across cancers. By capturing molecular variability independently from drug response, it avoids biases from prior annotations while retaining biological interpretability.

Applied to DepMap, MIRACS uncovered convergent resistance programs, exemplified by Factor 12. Unlike prior studies restricted to single tumor types (Linares-Blanco et al., 2025) or relying on black-box models (Wu et al., 2025), this framework systematically detects cross-tumor associations while remaining interpretable.

Limitations include reliance on cell lines, linear modeling assumptions, and dependence on MoA annotations. Future work should validate MIRACS-derived factors in clinical cohorts and expand integration to other omics layers such as mutations and methylation.

5 Conclusions

Our approach identified pre-existing molecular states potentially linked to drug resistance. Integrating pan-cancer transcriptomic and proteomic data revealed potential resistance mechanisms transcending tumor types or drug classes.

Bibliography

P. Ak and A. J. Levine. p53 and NF-κB: different strategies for responding to stress lead to a functional antagonism. *The FASEB Journal*, 24(10):3643–3652, 6 2010. URL *https://doi.org/10.1096/fj.10-160549*.

R. Argelaguet, B. Velten, D. Arnol, S. Dietrich, T. Zenz, J. C. Marioni, F. Buettner, W. Huber, and O. Stegle. Multi-Omics Factor Analysis—a framework for unsupervised integration of multi-omics data sets. *Molecular Systems Biology*, 14(6), 6 2018. URL *https://doi.org/10.15252/msb.20178124*.

S. M. Corsello, R. T. Nagari, R. D. Spangler, J. Rossen, M. Kocak, J. G. Bryan, R. Humeidi, D. Peck, X. Wu, A. A. Tang, V. M. Wang, S. A. Bender, E. Lemire, R. Narayan, P. Montgomery, U. Ben-David, Y. Chen, M. G. Rees, N. J. Lyons, J. M. McFarland, B. T. Wong, L. Wang, N. Dumont, P. J. O'Hearn, E. Stefan, J. G. Doench, H. Greulich, M. Meyerson, F. Vazquez, A. Subramanian, J. A. Roth, J. A. Bittker, J. S. Boehm, C. C. Mader, A. Tsherniak, and T. R. Golub. Non-oncology drugs are a source of previously unappreciated anti-cancer activity. *bioRxiv* (*Cold Spring Harbor Laboratory*), 8 2019. URL *https://doi.org/10.1101/730119*.

DepMap, Broad. Depmap 24q4 public. *https://doi.org/10.25452/figshare.plus.27993248.v1*, 2024. Dataset.

Y. Gu, R. Yang, Y. Zhang, M. Guo, K. Takehiro, M. Zhan, L. Yang, and H. Wang. Molecular mechanisms and therapeutic strategies in overcoming chemotherapy resistance in cancer. *Molecular Biomedicine*, 6(1), 1 2025. URL *https://doi.org/10.1186/s43556-024-00239-2*.

Z. Lei, Q. Tian, Q. Teng, J. N. D. Wurpel, L. Zeng, Y. Pan, and Z. Chen. Understanding and targeting resistance mechanisms in cancer. *MedComm*, 4(3), 5 2023. URL *https://doi.org/10.1002/mco2.265*.

J. Linares-Blanco, I. Chicote, J. Ros, A. M. Alcantara, J. Martinez-Quintanilla, E. Elez, H. G. Palmer, C. Fernandez-Lozano, and J. A. Seoane. Transcriptomic signature reveals sensitivity to tubulin inhibitors in colon cancer. *bioRxiv* (*Cold Spring Harbor Laboratory*), 1 2025. URL *https://doi.org/10.1101/2025.01.14.633026*.

D. P. Nusinow, J. Szpyt, M. Ghandi, C. M. Rose, E. R. McDonald, M. Kalocsay, J. Jané-Valbuena, E. Gelfand, D. K. Schweppe, M. Jedrychowski, J. Golji, D. A. Porter, T. Rejtar, Y. K. Wang, G. V. Kryukov, F. Stegmeier, B. K. Erickson, L. A. Garraway, W. R. Sellers, and S. P. Gygi. Quantitative Proteomics of the Cancer Cell Line Encyclopedia. *Cell*, 180(2):387–402.e16, 1 2020. URL *https://doi.org/10.1016/j.cell.2019.12.023*.

S. T. Stengel, A. Fazio, S. Lipinski, M. T. Jahn, K. Aden, G. Ito, F. Wottawa, J. W. Kuiper, O. I. Coleman, F. Tran, D. Bordoni, J. P. Bernardes, M. Jentzsch, A. Luzius, S. Bierwirth, B. Messner, A. Henning, L. Welz, N. Kakavand, M. Falk-Paulsen, S. Imm, F. Hinrichsen, M. Zilbauer, S. Schreiber, A. Kaser, R. Blumberg, D. Haller, and P. Rosenstiel. Activating transcription factor 6 mediates inflammatory signals in intestinal epithelial cells upon endoplasmic reticulum stress. *Gastroenterology*, 159(4):1357–1374.e10, 7 2020. URL *https://pmc.ncbi.nlm.nih.gov/articles/PMC7923714/*.

A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, and J. P. Mesirov. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, 102(43):15545–15550, 9 2005. URL *https://doi.org/10.1073/pnas.0506580102*.

A. Takaoka, T. Tamura, and T. Taniguchi. Interferon regulatory factor family of transcription factors and regulation of oncogenesis. *Cancer Science*, 99(3):467–478, 1 2008. URL *https://pmc.ncbi.nlm.nih.gov/articles/PMC11159419/#:~:text=Stress%E2%80%90responsive%20IRF:%20IRF1%20and%20IRF5.&text=IRF1%20and%20IRF2%20were%20initially,some%20facets%20of%20TLR%20signaling.&text=However%2C%20many%20studies%20have%20revealed,genotoxic%20stresses%20(see%20below)*.

N. Vasan, J. Baselga, and D. M. Hyman. A view on drug resistance in cancer. *Nature*, 575(7782):299–309, 11 2019. URL *https://doi.org/10.1038/s41586-019-1730-1*.

Y. Wu, M. Chen, and Y. Qin. Anticancer drug response prediction integrating multi-omics pathway-based difference features and multiple deep learning techniques. *PLoS Computational Biology*, 21(3):e1012905, 3 2025. URL *https://doi.org/10.1371/journal.pcbi.1012905*.