Biomarker Development from Functional Precision Medicine Datasets via Explainable Machine Learning

Noah E. Berlow[1], Arlet M. Acanda De La Rocha[2], Maggie Fader[3], Ebony R. Coats[2], Cima Saghira[4], Paula S. Espinal[3], Jeanette Galano[3], Ziad Khatib[3], Haneen Abdella[3], Ossama M. Maher[3], Cristina M. Andrade-Feraud[2], Baylee Holcomb[2], Yasmin Ghurani[2], Lilliam Rimblas[3], Tomás R. Guilarte[2], Nan Hu[2], Daria Salyakina[3], Diana J. Azzam[1,2]

[1] First Ascent Biomedical, Beaverton, OR; [2] Florida International University, Miami, FL; [3] Nicklaus Children's Hospital, Miami, FL; [4] Nicklaus Children's Hospital, Miami, FL

BACKGROUND

Current large-scale studies and routine clinical use of genomics precision medicine demonstrates ~10% benefit rates through genomics-informed therapy. The gap between actionability and benefit remains a major clinical challenge attributed to multiple factors, including the complex relationship between molecular status and therapy response.

Functional precision medicine (FPM) – the integration of patient-specific drug sensitivity testing with molecular tumor profiling– represents the next generation of multi-modal tumor profiling approaches for clinical decision support, and provides new opportunities identify relationships between tumor molecular profiles and functional response, requiring novel computational approaches for multi-omics biomarker development.

To address this challenge, we developed and validated an explainable multi-modal, multi-omics biomarker development platform designed for practical oncology functional genomics datasets. We used this approach to analyze data from our FPM trial at Nicklaus Children's Hospital in Miami, FL (NCT05857969, NCT03860376). We also applied this same approach to functional genomics datasets and drug development datasets, demonstrating development of computationally- and biologically-optimized response biomarkers.

METHODS

Explainable Machine Learning for Multi-Modal Multi-Omics Biomarkers

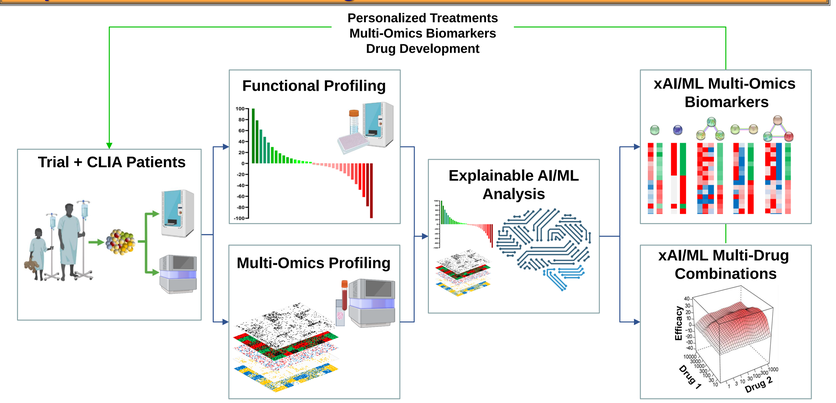

Figure 2: Workflow of xML multi-modal multi-omics biomarkers. Workflow diagram for xML development of multi-omics biomarkers and mechanism-based multi-drug combinations development from oncology functional genomics datasets. Functional (drug sensitivity testing) and molecular (whole exome, whole transcriptome, etc.) profiling data from cohorts of patients enrolled through clinical trials or CLIA services are ingested into First Ascent's PTM-Signature xML engine. Multi-modal, multi-omics signatures predictive of pharmacological response are generated by the machine learning system and are further refined by existing biological and disease knowledge-bases. These biomarkers are used to develop a biologically-driven network of the mechanisms of pharmacological response and identify individualized or cohort combination therapy approaches. Finally, individual signatures and the signature network identify expanded label or repurposing opportunities, and rank additional oncology models for follow-on studies.

Results

Analysis of Functional Response Profiles – Clustering and Ethnicity

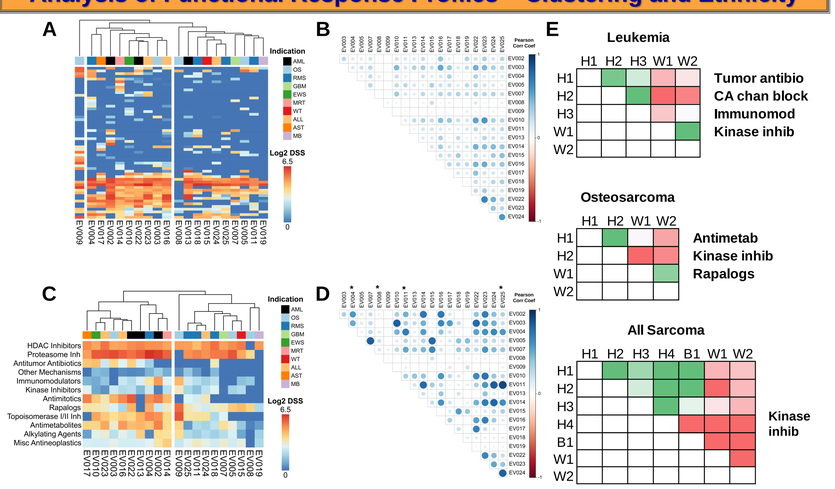

Figure 2: Clustering and functional response analysis of patient FPM data. A) Agglomerative hierarchical clustering of ex vivo drug sensitivity profiles of 20 DST-assayed patients across 56 common drugs. B) Inter-patient Spearman correlation coefficients of DSS response profiles. Correlation coefficients are visualized as squared values for visual clarity. C) Agglomerative hierarchical clustering of ex vivo drug sensitivity profiles of 20 DST-assayed patients across 56 common drugs, grouped by drug class. D) Inter-patient Spearman correlation coefficients of DSS response profiles grouped by drug class. Correlation coefficients are visualized as squared values for visual clarity. * represents patient DSS profiles most correlated with a patient sample of the same indication. E) Differences in functional response to oncology drug classes between patients of different racial/ethnic backgrounds, stratified by disease type.

RESULTS (cont)

Pan-Pediatric Cancer xML Biomarkers from FPM Data - Romidepsin

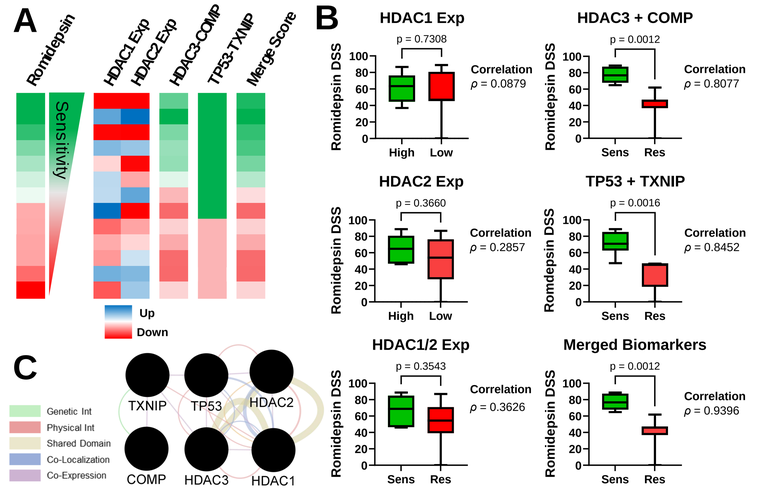

Figure 3: Romidepsin biomarkers in pan-pediatric cancer identified by xML analysis. A) Gene expression level of canonical romidepsin targets (HDAC1 and HDAC2) and scores from novel xAI-developed multi-omics biomarkers, and merged biomarkers. B) DSS distribution of individual and merged multi-omics romidepsin response biomarkers identified by xML analysis. Pearson correlation coefficient is provided alongside distribution plots. C) Interaction and association network for canonical romidepsin targets and genes identified in romidepsin response biomarkers.

Pan-Pediatric Cancer xML Biomarkers from FPM Data - Idarubicin

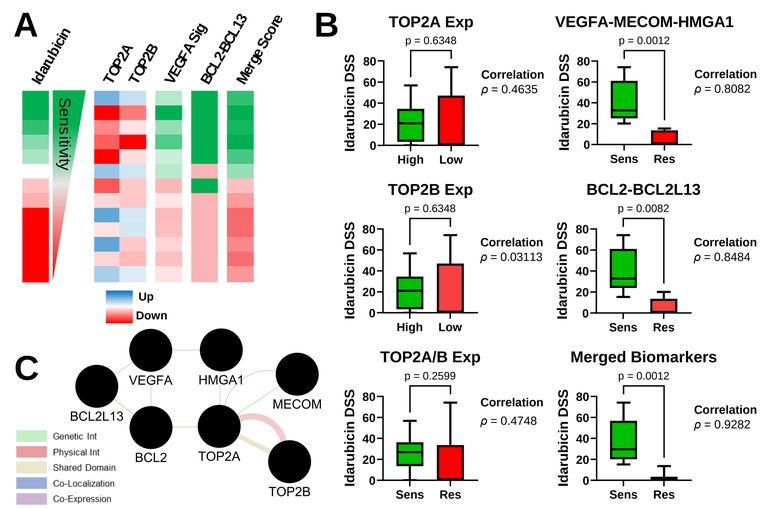

Figure 4: Idarubicin biomarkers in pan-pediatric cancer identified by xML analysis. A) Gene expression level of canonical idarubicin targets (TOP2A and TOP2B) and scores from novel xAI-developed multi-omics biomarkers, and merged biomarkers. B) DSS distribution of individual and merged multi-omics idarubicin response biomarkers identified by xML analysis. Pearson correlation coefficient is provided alongside distribution plots. C) Interaction and association network for canonical idarubicin targets and genes identified in idarubicin response biomarkers.

REFERENCES

1. A. A. de La Rocha, N. E. Berlow, … D. Azzam. "Feasibility of functional precision medicine for guiding treatment of relapsed or refractory pediatric cancers". Nature Medicine 2024 Apr 11. doi: 10.1186/s12885-019-5681-6
2. N. E. Berlow. "Probabilistic Boolean Modeling of Pre-clinical Tumor Models for Biomarker Identification in Cancer Drug Development". Current Protocols in Pharmacology 2021 October 18. https://doi.org/10.1002/cpz1.269

FUNDING SUPPORT

NIH National Institute on Minority Health and Health Disparities

NIH/NIMHD U54MD012333

NIH National Cancer Institute

techstars

Live like Bella Childhood Cancer Foundation Est. 2013

NSF

Treyfoote Foundation

RESULTS (cont)

xML Biomarkers from Functional Genomics Datasets - Erlotinib

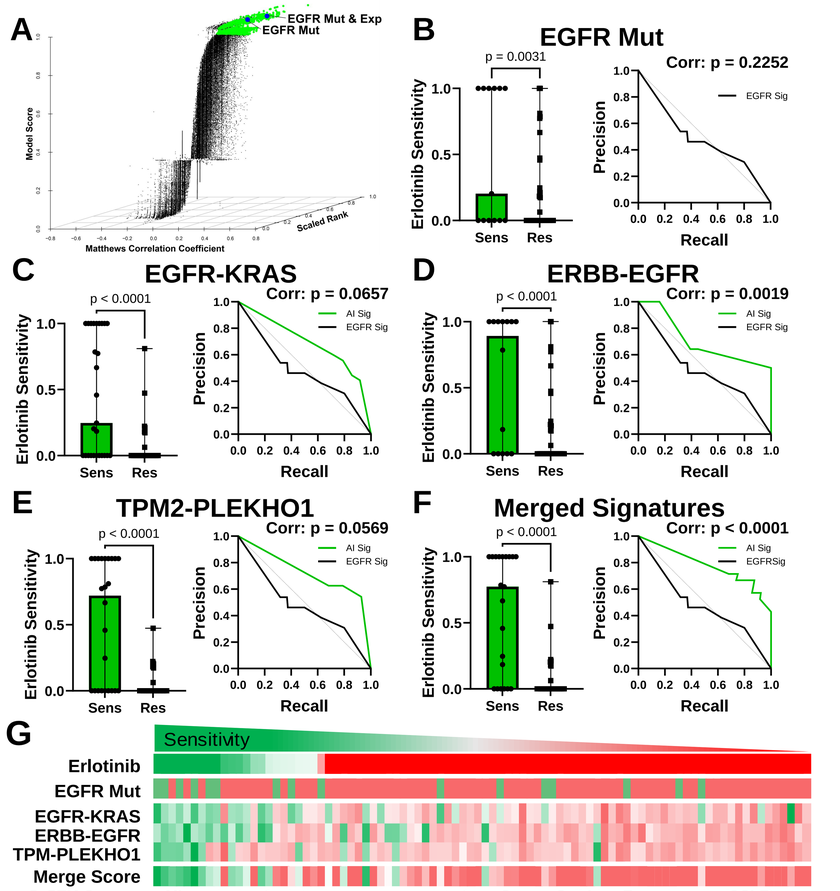

Figure 5: PTM-Signature xML analysis of erlotinib-lung cancer data from Cancer Cell Line Encyclopedia (CCLE). A) xML analysis of mono-features, identifying EGFR status among top 0.1% of features. In B-F, distributions are predicted response to erlotinib, p-values are from Kolmogorov-Smirnov tests, correlations are from Spearman tests, and precision-recall curves are presented. B) Analysis of standard EGFR mutation biomarker. C) Analysis of xML EGFR-KRAS biomarker. D) Analysis of ERBB-EGFR biomarker. E) Analysis of novel TPM2-PLEKH01 biomarker. F) Analysis of merged biomarker signature. G) Erlotinib sensitivity and biomarker predicted scores.

xML Biomarkers from Functional Genomics Datasets - PARPi

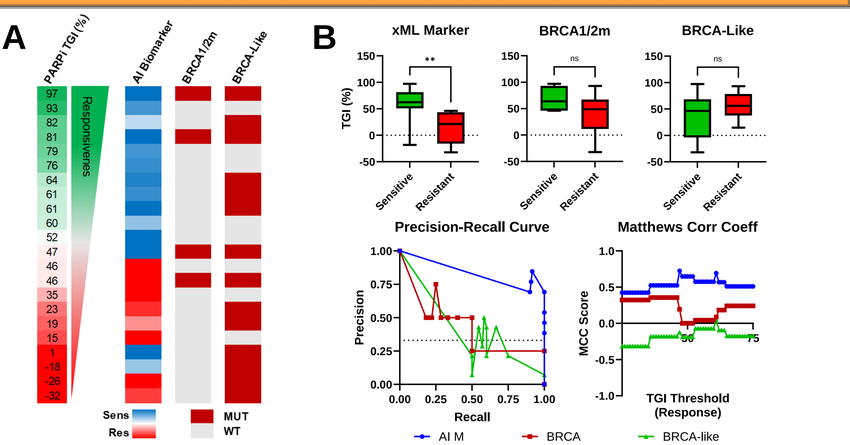

Figure 6: PTM-Signature xML analysis of PARPi treatment data from ovarian cancer PDX models. A) Tumor growth inhibition from ovarian cancer PDX models treated with the PARPi, with response predictions from PTM-Signature xML-defined biomarker, BRCA1/2m marker or BRCA-like mutation markers. B) TGI response distributions of PDX response predicted by PTM-Signature xML biomarker, BRCA1/2m biomarker, and BRCA-like biomarkers. Precision-recall curves and Matthew's Correlation Coefficient analysis of resulting response predictions.

CONCLUSIONS

Our studies demonstrate the potential for predictive and generative artificial intelligence and machine learning approaches to address critical challenges multi-omics, multi-modal biomarker design to support clinical decision-making and enhance the clinical utility of multi-omics tumor profiling. Improving therapy assignment through more robust pharmacogenomic biomarkers, which can improve outcomes and reduce ineffective therapy assignment. These novel biomarkers can be built both to expand labeling for existing drugs and support novel drug development.