Predicting drivers of drug response from baseline omics data across breast cancer cell lines and models

Chiara Victor, Ben Gaudio, Mirra Chung, Marc Hafner, Artem Sokolov, Chen Chen, Sarah A. Boswell, Matthew Berberich, Robert A. Everley, Marian Kalocsay, Maulik Nariya, Gary Bradshaw, Caitlin E. Mills, Kartik Subramanian, Peter K. Sorger

Laboratory of Systems Pharmacology, Harvard Medical School, Boston, MA

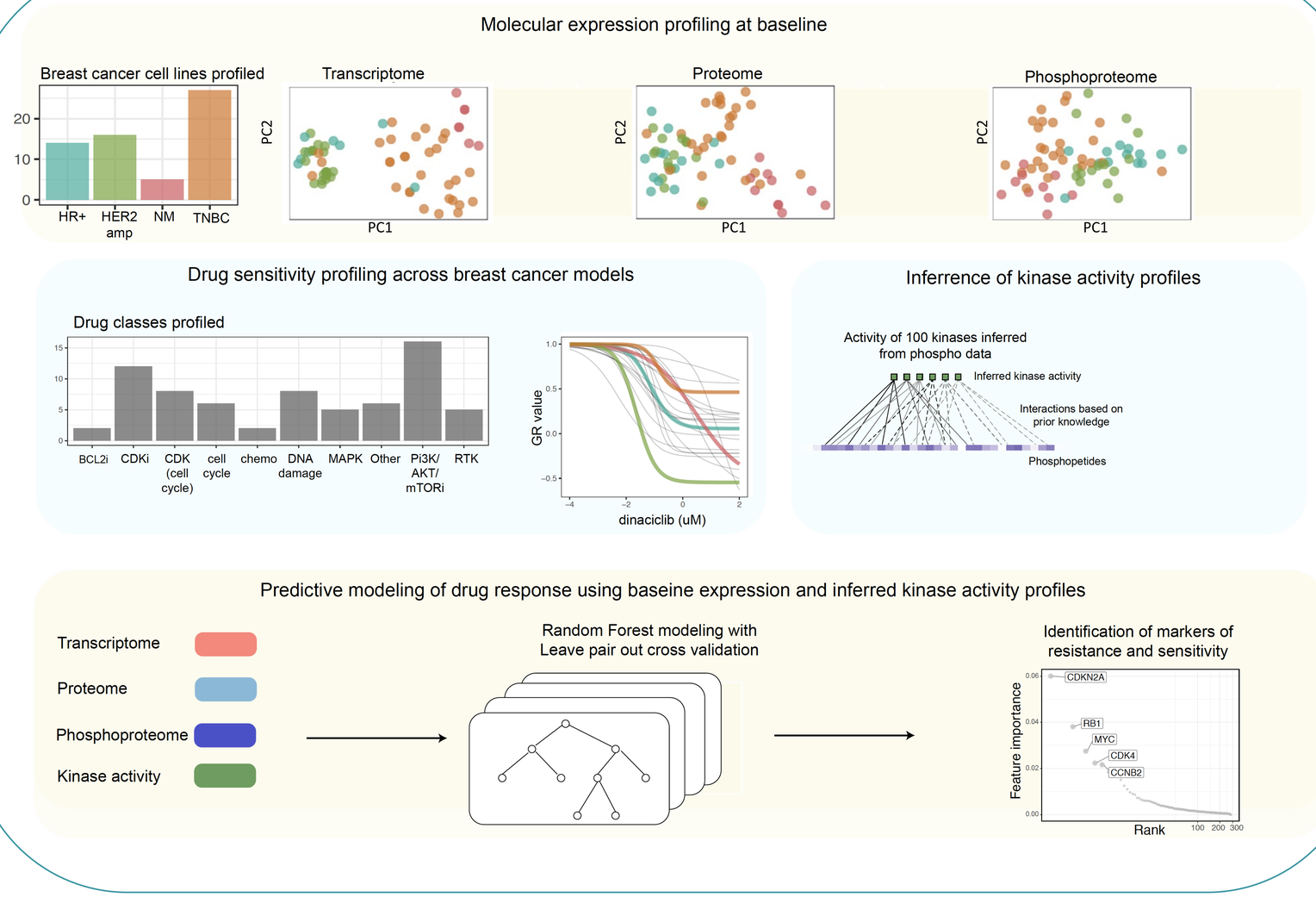
Harvard Program in Therapeutic Science

Laboratory of Systems Pharmacology

LINCS

HARVARD MEDICAL SCHOOL

Project overview and goals



Drug response predictions

Model performance & feature importance

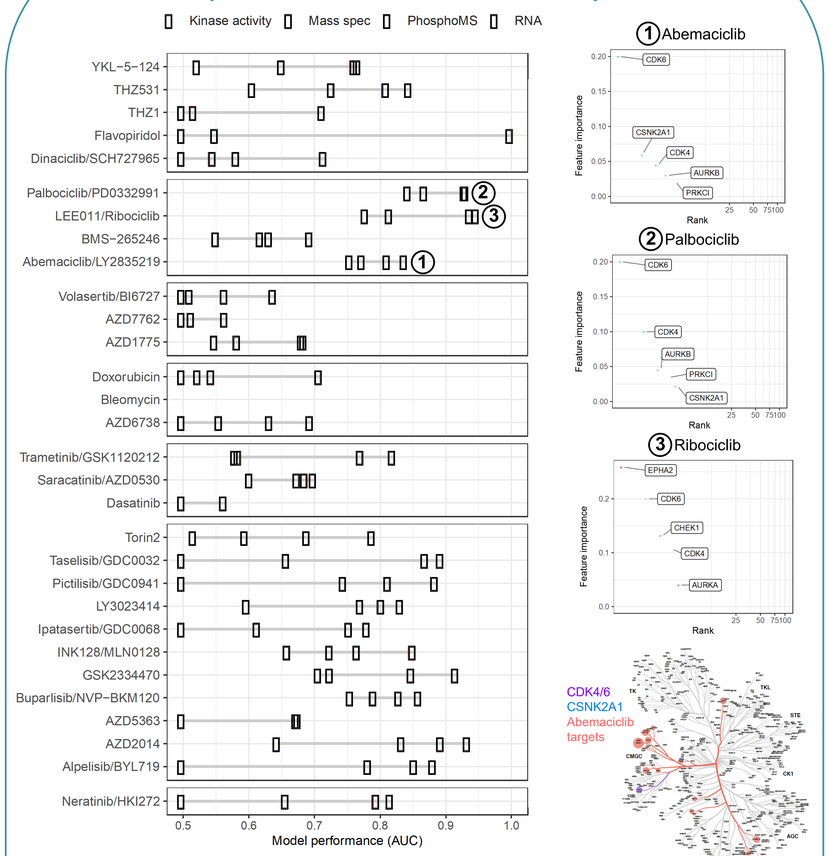


Figure 2: The performance of a random forest model for breast cancer drug response prediction using baseline RNA, protein, or phosphoprotein expression or inferred kinase activity. The most important Features for **1** abemaciclib, **2** palbociclib and **3** ribociclib show that kinase activity of the nominal drug targets (CDK4/6) influences cell line response, and reveal the importance of CSNK2A1 in abemaciclib sensitivity, a known off target as shown in the kinome tree (bottom right).

Rationale

Several publications have addressed concerns surrounding drug response screens by identifying sources of variability and by providing recommendations for improved experimental methods and more robust analytical approaches. In the HMS LINCS breast cancer profiling effort, we selected 70 breast cancer cell lines and evaluated their responses to a panel of 68 clinically relevant agents, using microscopy-based Dye Drop dose response assays to measure drug potency and efficacy in terms of growth rate inhibition (GR metrics), cell death and cell cycle fractions. This systematic dose response dataset is complemented by measurements of baseline transcript expression levels by mRNAseq, quantification of absolute abundance of ~12,000 proteins, and relative phosphoprotein levels by shotgun mass spectrometry across all cell lines. Additionally, the baseline activity of kinases were inferred from phosphoprotein data using a custom kinase enrichment analysis. The three baseline expression datasets and the inferred kinase activity dataset were used to build random forest predictive models of drug response. Overall these datasets will be a valuable resource for understanding drug response in breast cancer models and the molecular mechanisms that influence them.

Dose response results

Figure 1: The Area Over the GR dose response Curves (GR AOC) for select drugs in 70 breast cancer cell lines. A GR AOC value of 0 is indicative of no response. Green: HR+ lines, Blue: HER2amp, Orange: TNBC, Red: non-malignant.

Conclusions

Predictive models of breast cancer cell line response to targeted kinase inhibitors reveal known biology: molecular markers that confer resistance/sensitivity to inhibition of the target or are off-targets of the drugs; and novel biology: molecular markers that suggest unknown off-targets of the drugs, and biomarkers for drug sensitivity. We are exploring several findings for *in vivo* and clinical biomarker potential, and for novel combination treatments. A manuscript is in preparation.

Acknowledgements

We thank ICCB-L for assistance setting up the library plates, and pin transfers, and the Gray lab for contributing compounds. This work was funded by NIH LINCS grant U54HL127365 to PKS.