American Association For Cancer Research – April 14-18, 2018

Computationally Predicted Sensitivity of Clinical Cohorts Identifies Biomarkers Associated with Response to PCM-075, a PLK-1 Selective Inhibitor

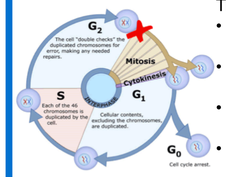
Abstract #2810

Penn Whitley[1], Peter J P Croucher[1], Barbara Valsasina[2], Dario Ballinari[2], Italo Beria[2], Jeffrey N. Miner[1] and Mark G. Erlander[1].
1. Trovagene, Inc., San Diego 92121, CA, United States 2. Nerviano Medical Sciences, S.R.L., 20014 Nerviano, MI, Italy

trovagene

Background

- Polo-like kinase 1 (PLK1), a serine-threonine kinase which regulates various cellular processes, including mitosis, DNA replication, and the stress response, is overexpressed in many malignancies
- PCM-075, a PLK1 selective inhibitor, is currently in a phase 1b/2 clinical trial in AML in combination therapy with low dose cytarabine (LDAC) (NCT03303339), and a phase 2 trial in metastatic Castration-Resistant Prostate Cancer (mCRPC) in combination with abiraterone (Zytiga®)(NCT03414034).
- A computational approach was developed to identify biomarkers associated with response to PCM-075



The approach consists of four steps:

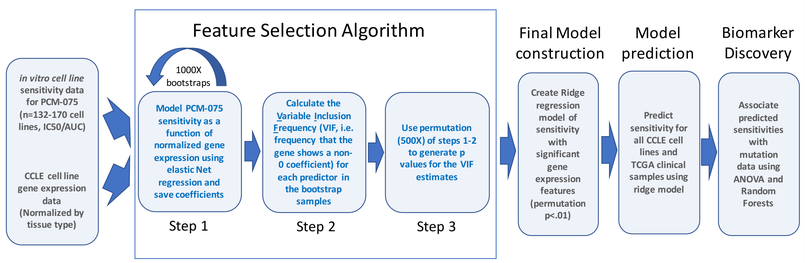
- Feature Selection: Determine critical RNA expression features associated with in-vitro sensitivity
- Final model construction: Generate regression model of sensitivity with selected RNA features
- Model Prediction: Apply the model to both clinical patient datasets (TCGA) and public cell line (CCLE) data
- Biomarker Discovery: Perform a statistical analysis to evaluate potential mutations and gene expression profiles associated with predicted sensitivity

Objective

- Discover potential RNA and DNA biomarkers that are associated with PCM-075 sensitivity using in-vitro cell line sensitivity data and predictive regression models

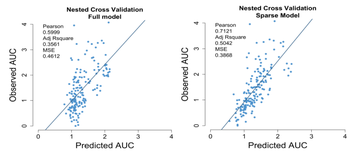
Methods

Biomarker Discovery pipeline



- PCM-075 in-vitro cell line sensitivity values (IC_{50}/AUC) are modeled using gene expression data from the CCLE
- A combination of elastic Net (eNet) regression, bootstrapping and permutation was used to perform feature selection prior to final regression modeling. 264 genes were found with p<0.01. 183 of these occurred in the TCGA RNA-seq dataset and were used for final modeling
- A final Ridge regression model was built with significant gene expression features from the selection procedure. The sparse model showed better performance than a full model starting with all gene expression features (see scatterplots below)
- The final model was used to predict sensitivity in all CCLE (n=819) and TCGA (n=9968) samples using a modified version of IDWAS (1)
- Significant gene expression features found during selection were evaluated using standard pathway analysis methods to better understand the model predictors
- Biomarker Discovery was performed with either ANOVA (CCLE analysis) or random forests and decision trees (TCGA analysis)

Regression model performance (Sparse versus full Ridge model)



Results

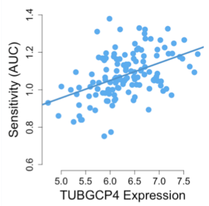
The gene expression profile associated with sensitivity to PCM-075 is positively enriched for pathways associated with highly proliferative/aggressive tumor growth

Gene Set Enrichment Analysis

geneset	description	Size	correlation	FDR
hsa03010	Ribosome - Homo sapiens (human)	101	positive	0.00E+00
hsa03008	Ribosome biogenesis in eukaryotes - Homo sapiens (human)	63	positive	0.00E+00
hsa03013	RNA transport - Homo sapiens (human)	135	positive	0.00E+00
hsa03040	Spliceosome - Homo sapiens (human)	118	positive	0.00E+00
hsa03030	DNA replication - Homo sapiens (human)	32	positive	4.74E-04
hsa03020	RNA polymerase - Homo sapiens (human)	27	positive	1.04E-03
hsa00240	Pyrimidine metabolism - Homo sapiens (human)	88	positive	5.42E-03
hsa04142	Lysosome - Homo sapiens (human)	116	negative	0.00E+00
hsa04141	Protein processing in endoplasmic reticulum - Homo sapiens (human)	157	negative	0.00E+00
hsa05110	Vibrio cholerae infection - Homo sapiens (human)	47	negative	4.17E-04
hsa04130	SNARE interactions in vesicular transport - Homo sapiens (human)	31	negative	1.38E-03
hsa04064	NF-kappa B signaling pathway - Homo sapiens (human)	86	negative	4.05E-03
hsa05162	Measles - Homo sapiens (human)	125	negative	4.73E-03
hsa00510	N-Glycan biosynthesis - Homo sapiens (human)	44	negative	4.84E-03
hsa04672	Intestinal immune network for IgA production - Homo sapiens (human)	32	negative	5.29E-03
hsa00600	Sphingolipid metabolism - Homo sapiens (human)	42	negative	7.69E-03

- Genes found during feature selection were enriched (FDR<0.01] for ribosome and rRNA processing gene ontology and KEGG pathways using hypergeometric test
- Univariate correlations of gene expression (n=17,419) with sensitivity were calculated and used in Gene Set Enrichment Analysis (GSEA). Ribosome biogenesis, and other fundamental replication and translation pathways were positively enriched (see table)
- Tumor cells with up-regulation of replication (cell cycle), transcription and translation pathways are more sensitive to PCM-075

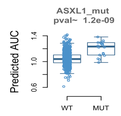
The top 2 genes with highest gene expression, TUBGCP4 and DVL1, are involved in mitotic activities associated with PLK1



- DVL1 and TUBGCP4 genes were the two highest ranked gene expression features found (ranked by VIF and p<0.01)
- Disheveled Segment Polarity Protein 1 (DVL1) is critical for cell division and microtubule stability. The DVL complex has been shown to be phosphorylated by PLK1
- Tubulin Gamma Complex Associated Protein 4 (TUBGCP4) is important for microtubule nucleation and is involved in the 'Regulation of PLK1 Activity at G2/M Transition' pathway
- High TUBGCP4 expression is associated with an aggressive subtype (STEM-A) of ovarian cancer
- PCM-075 has been shown to have significant anti-tumor activity in a xenograft A2780 STEM-A cell model (not shown)

Mutation biomarkers associated with predicted sensitivity in CCLE model cell lines

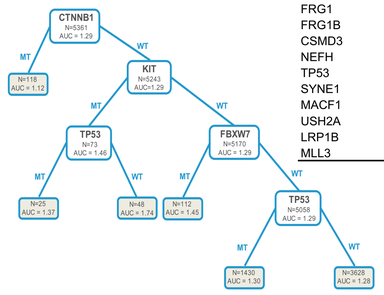
- Sensitivity (AUC) was predicted for all 819 cell lines using tissue normalized gene expression profiles from CCLE
- 1-way ANOVA was then used to search for potential driver mutations (curated by the GDSC) associated with predicted sensitivity values
- 19 gene mutations show a significant association (p<.05) with predicted sensitivity and are associated with aggressive tumor growth
- ASXL1 mutations was the highest ranked potential biomarker
- Tumors with ASXL1 mutations are highly aggressive and show poor prognosis in many indications



gene	mutants	coefficient	pval	fdr
ASXL1	18	0.153	1.21E-09	2.82E-07
BCR.ABL	6	0.160	2.55E-04	2.98E-02
ALK	10	0.102	2.70E-03	2.10E-01
TP53	546	0.022	4.68E-03	2.61E-01
FLT3	11	0.090	5.59E-03	2.61E-01
BRCA2	38	0.046	9.42E-03	3.66E-01
EWSR1.X	6	0.109	1.24E-02	4.12E-01
SMARCA4	47	0.038	1.69E-02	4.91E-01
MAP4K3	5	-0.110	2.14E-02	5.55E-01
TJP1	6	0.098	2.45E-02	5.70E-01
MKL1	8	0.082	2.98E-02	5.76E-01
EP300	62	0.030	3.47E-02	5.76E-01
EWSR1.FLI1	16	0.057	3.59E-02	5.76E-01
CHD4	14	0.060	3.82E-02	5.76E-01
BMPR2	20	0.050	3.97E-02	5.76E-01
USP6	8	0.077	4.33E-02	5.76E-01
WT1	5	0.097	4.39E-02	5.76E-01
SOS2	11	0.065	4.59E-02	5.76E-01
LARP4B	14	0.057	4.93E-02	5.76E-01

Results

Biomarkers associated with predicted sensitivity in solid tumor TCGA cohorts



Marker	P-value
CTNNB1	6.92E-17
KIT	8.91E-15
ZFPM1	1.36E-13
FBXW7	1.15E-12
FRG1	3.98E-09
FRG1B	8.47E-07
CSMD3	4.03E-06
NEFH	6.53E-06
TP53	3.05E-05
SYNE1	9.11E-05
MACF1	1.05E-03
USH2A	1.30E-03
LRP1B	4.98E-03
MLL3	8.33E-03

AML and Prostate cancer cohort specific TCGA associations

Indication	Mutation	P-value
LAML	NPM1	2.69E-02
LAML	KIT	4.25E-02
LAML	FAM5C	8.49E-02
LAML	NRAS	1.68E-01
LAML	U2AF1	2.20E-01
PRAD	SPOP	1.80E-06
PRAD	TNXB	8.00E-04
PRAD	5q15del (RGMB)	8.30E-03
PRAD	ZMYM3	3.50E-03
PRAD	BRCA2	4.10E-03

- Sensitivity (AUC) was predicted using gene expression data for subjects in the TCGA (n = 9968 samples)
- Selection of critical mutations was performed using random forests with mutation and CNV data
- Analysis was restricted to one of; solid tumors, prostate samples (PRAD) or Acute Myeloid Leukemia samples (LAML)
- The solid tumor analysis resulted in CTNNB1 and KIT mutations showing the strongest effect on sensitivity and were placed at the top of a final decision tree
- A random forest analysis of AML (n = 155) resulted in selection of NPM1 (12.5% of TCGA-AML cases) and KIT (6.9% of TCGA-AML cases). Other markers of high importance were found (e.g. FAM5C, NRAS, U2AF)
- An analysis of Prostate cancer samples (n = 320) resulted in selection of SPOP (11% of TCGA-prostate cases). Other markers of high importance were TNXB, ZMYM3, BRCA2 and a CNV at 5q15 del (RGMB)

Conclusions

- A feature selection and modelling method was developed which leverages penalized regression models, bootstrapping and permutation analysis to better identify biomarkers associated with drug sensitivity
- Overall, tumor cells with hyperactive pathways in DNA replication (cell cycle), gene transcription and translation are significantly associated with increased sensitivity to PCM-075
- The top gene expression signatures associated with PCM-075 sensitivity include DVL1 and TUBGCP4, both are involved in microtubule stability and are associated with PLK1 activity
- 19 putative driver mutations were associated with predicted sensitivity across 819 cancer cell lines, including ASXL1, BCR-ABL and TP53; these mutations are associated with aggressive tumor growth and poor prognosis
- The model was used to extrapolate to a large solid tumor patient cohort and analysis was performed to find potential pan-cancer biomarkers associated with PCM-075 sensitivity. A final decision tree was constructed with these mutations suggesting a possible role for CTNNB1, KIT, FBXW7 and TP53 mutations
- A prostate and AML cohort specific analysis was also performed due to the ongoing PCM-075 clinical trials in those indications. Drivers found in both indications are critical prognostic markers (NPM1 in AML and SPOP in prostate) that may be predictive of greater PCM-075 efficacy

References

1. Model developed was based on Geeleher P. et al. (2017), Discovering novel pharmacogenomic biomarkers by imputing drug response in cancer patients from large genomics studies. Genome Research 27*: 1743-1751*

Public data sources used in the current study:

1. CCLE – Cancer Cell Line Encyclopedia - *Nature* volume 483, pages 603–607 (29 March 2012)
2. GDSC – Genomics of Drug Sensitivity in Cancer - *Nucleic Acids Research*, Volume 41, Issue D1, Pages D955–D961 (1 January 2013)
3. CTRP – Cancer Therapeutics Response Portal V2 – *Cell* volume 154, Issue 5, pages 1151-1161 (29 August 2013)
4. TCGA – The Cancer Genome Atlas - http://cancergenome.nih.gov/

Mark Erlander, PhD
Trovagene, Inc.
11055 Flintkote Ave.
San Diego, CA 92121