

Molecular profiling establishes genetic features predictive of the efficacy of the p110 β inhibitor KIN-193

Isha Sethi^{1,2}, Zhenying Cai³, Thomas M. Roberts³, Guo-Cheng Yuan^{1,*}

¹: Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute and Harvard Chan School of Public Health, Boston, MA 02215

²: Department of Cardiology, Boston Children's Hospital and Harvard Medical School, Boston, MA 02115

³: Department of Cancer Biology, Dana-Farber Cancer Institute and Harvard Medical School, Boston, MA 02215

* Corresponding Author: Email: gcyuan@jimmy.harvard.edu; Phone number: 1-671-582-8532; Mailing Address: 450 Brookline Avenue, Mail Stop LC1052, Boston, MA 02215, USA

Conflict of Interest: The authors declare no potential conflicts of interest.

Abstract:

Aberrant activation of the PI3K pathway is a common alteration in human cancers. Therapeutic intervention targeting the PI3K pathway has achieved limited success due to the intricate balance of its different components and isoforms. Here, we systematically investigated the genomic and transcriptomic signatures associated with response to KIN-193, a compound specifically targeting the p110 β isoform. By integrating genomic, transcriptomic, and drug response profiles from the GDSC database, we identified mutational and transcriptomic signatures associated with KIN193 and further created statistical models to predict the treatment effect of KIN-193 in cell lines which may eventually be clinically valuable. These predictions were validated by analysis of the external CCLE data set. These results may assist precise therapeutic intervention targeting the PI3K pathway.

Statement of Significance: Findings provide new insights into molecular signatures associated with sensitivity of the p110 β inhibitor KIN-193, which may provide a useful guide for developing precise treatment methods for cancer.

Introduction

Phosphatidylinositol 3-kinase (PI3K) is an important signaling pathway mediating diverse cellular functions such as metabolism, cell growth and cell death [1, 2]. It contains a family of genes that are divided into three classes. Class I, which is further subdivided into class IA and class IB, is the most widely studied and has been implicated in promoting proliferation in many cancers [3, 4]. The class IA PI3K proteins are heterodimers composed of a catalytic and regulatory subunit. The catalytic subunit has three isoforms: p110 α , p110 β and p110 δ . p110 α and p110 β , encoded by PIK3CA and PIK3CB genes respectively, are the most ubiquitously expressed PI3K proteins across human cell types. Aberrant activation of the PI3K pathway is frequently observed in cancer. This is often driven by gain of function mutations in PIK3CA gene [4-6], while oncogenic PIK3CB mutations occur infrequently [7, 8]. However, some cancers are exclusively dependent on p110 β protein instead of p110 α [8-11]. These tumors mostly feature a wild-type (WT) PIK3CA gene but often have mutations and or deletions in the PTEN tumor suppressor, which is the main negative regulator of PI3K activity. Therefore, studying the action and effects of p110 β specific drugs in PTEN-null tumors can be a very rewarding avenue of research.

The small molecule KIN-193 was previously shown to specifically target the p110 β isoform and, hence, has been proposed as a potential drug to turn off PI3K signaling in PTEN null tumors [12]. In a previous study [12], it was shown that PTEN-null cell-lines indeed are highly enriched for KIN-193-sensitive status, although there remains an important subset that is KIN-193 resistant. Therefore, it is of clinical interest to identify the genetic signature(s) that may distinguish KIN-193-sensitive from resistant cell lines. In addition, many PTEN wild-type (WT) cell lines are also found to be KIN-193 sensitive, suggesting there are additional genetic and transcriptomic signatures that are associated with KIN-193 sensitivity beyond PTEN status.

Recent efforts by various consortia have led to massive online repositories containing drug sensitivity data coupled with genetic and transcriptomic information for thousands of cancer cell lines [13, 14], providing a great opportunity to conduct a systematic analysis by using computational methods. In this study, we primarily use information from the GDSC database [13] and build statistical models to predict KIN-193 sensitivity. We show that a 100-gene mutation signature is highly predictive for KIN-193 resistant, PTEN-mutated cancer cell-lines. We further explored transcriptional signatures associated with KIN-193 sensitivity in PTEN-WT cell-lines and identified a 203-gene transcriptomic signature that has significant prediction power. These predictions were validated by using the external CCLE database [14] and were an extension of a previous study [12]. Our analysis has provided novel insights into the mechanism for p110 β dependency and may be useful for predicting treatment outcome in the clinical setting.

Methodology

Data Sources:

We utilized data from 2 major databases: GDSC (Genomics of Drug Sensitivity in Cancer: <http://www.cancerrxgene.org/>) and CCLE (Cancer Cell Line encyclopedia: <http://www.broadinstitute.org/ccle>). The GDSC Project has extensively profiled 1001 cell-lines for their mutation background (using Whole-Exome sequencing), copy number variations (CNV), gene expression and DNA methylation status. They have also generated drug-response profiles for 265 drugs across a majority of these cell-lines. Similarly, the CCLE consortium has generated mutation, CNV and expression data for 1043 cancer cell-lines. We made two types of models: 1) based on mutation status and 2) derived from gene expression signature, using GDSC datasets, in cell-lines profiled for response to KIN-193 drug (alternative name: AZD6482). We then validated our models using

the external mutation/expression CCLE datasets and drug response data profiled by various other groups [12, 15, 16].

Modeling using mutation status of genes as predictive variables:

Preprocessing of dataset:

We z-score normalized the drug sensitivity profile of KIN-193 across the entire cell-line panel for each study (e.g.: 1001 GDSC cell-lines). We then categorized all cell-lines with z-score < -0.5 (an arbitrary cutoff) as KIN-193 sensitive and cell-lines with z-score > 0.5 as KIN-193 resistant. All other cell lines were characterized as Ambiguous.

PTEN-null cell-lines were defined as those mutated in the PTEN gene. From the GDSC database we selected for PTEN-null cell-lines that could be characterized as KIN-193 sensitive (57 cell-lines with z-score < -0.5) or resistant (10 cell-lines with z-score > 0.5). We next filtered to remove all mutations that were present at low frequency (< 3 cell-lines). This resulted in a feature matrix of 67 cell-lines X 5143 genetic mutations, which was used to make all further PTEN null models.

Mutual Information based Linear (MIL) model

The feature matrix was used as input for PARIS (Probability Analysis by Ranked Information Score), which has been implemented as part of Project Achilles (<https://portals.broadinstitute.org/achilles/resources/paris>). PARIS uses a mutual information-based metric (RNMI score: range -1 to 1) to rank features (in this case mutations) which is then used to determine most significant associations with the target profile (in this case: sensitivity/resistance to KIN-193 drug). To our knowledge, PARIS has not been previously used to predict drug sensitivity profiles.

We ran this analysis using stratified 10-fold cross-validation (to maintain class balance) and found the top 100 features associated with sensitivity and

resistance each. We next assigned equal weight to each significant mutation and scored all sensitive mutations as +1 and resistant mutations as -1. This enabled us to calculate a cumulative mutation score for all cell-lines, which in turn could be used to predict their sensitivity/resistance to KIN-193. We also generated a cumulative RNMI score for each cell-line, which took into account the varied correlations of features with target profile, to predict sensitivity/resistance to KIN-193. We repeated this 10-fold cross-validation analysis 25 times to evaluate model performance.

Other Machine learning models: GLMNET & GBM

We also made use of more sophisticated machine learning models like GLMNET (type of penalized linear regression) and GBM (Gradient Boosting: typically an ensemble of decision trees) [17] to make predictions about sensitivity/resistance of cell-lines to KIN-193. These models were comprehensive in the sense that they also selected for significant features associated with prediction.

As we had a limitation of data especially for PTEN null models, we used nested 10-fold stratified cross-validation to generate precision-recall curves to validate the models. The inner fold CV was used to tune the parameters and make the model whereas the outer fold CV was used to repeatedly generate training and testing datasets to validate the models. The models were made using Caret package in R [17].

Determining PTEN-WT cell-lines that can be predicted as KIN-193 sensitive

Defining a gene expression signature characteristic of majority PTEN null sensitive cells

Utilizing the GDSC database, we generated an expression matrix of 57 PTEN-null sensitive cell-lines by 17417 genes. We log-normalized the matrix and then filtered for genes that had highly variable expression ($\log_2(\max \text{ expr.}) - \log_2(\min \text{ expr.}) \geq 1$). This gave a matrix of 57 cell-lines by 5313 genes which was then

hierarchically clustered using the Spearman rank correlation distance metric and Complete linkage clustering method. This resulted in a core cluster of 17 cell-lines (Spearman rank corr. ≥ 0.7). We next employed the Wilcoxon signed-rank sum test to identify genes that are differentially expressed between this core cluster and 967 PTEN WT cell-lines. 203 genes (at a stringent cutoff p-value of $1E-7$) were found to be significantly differentially expressed. These genes were used for subsequent downstream gene signature analysis. By calculating a median expression pattern for the core cluster of 17 PTEN-null sensitive cell-lines across the 203 genes, we were therefore able to define a PTEN null sensitive signature.

Finding PTEN WT cell-lines that cluster with PTEN null-sensitive signature

Next, to determine PTEN WT cell-lines that have a similar expression pattern to PTEN null sensitive signature, we took the matrix of 203 genes by 967 PTEN WT cell-lines and clustered (hierarchical clustering with Spearman rank correlation distance metric) them with the PTEN null sensitive signature defined above. Any WT cell-line with a correlation coefficient greater than 0.75 (as an arbitrary cutoff) was predicted as KIN-193 sensitive.

Results

Distinct genetic features are associated with KIN-193 resistance

In order to systematically identify the genetic features that are associated with KIN-193 sensitivity, we carried out an integrative analysis of the GDSC database [13], which contains genetic, transcriptomic, and drug response information of over 1001 cancer cell lines. As a starting point, we focused on the subset of 119 cell lines with PTEN mutation, which was previously recognized as a major determinant for KIN-193 sensitivity [12]. Among these cell lines, 67 had an unambiguous drug sensitivity outcome, including 57 sensitive and 10 resistant

cell lines (see Methods for details) (Supplementary Table 1A). After removing features that were infrequently mutated ($n \leq 3$) in these cell lines, we obtained a list of 5143 features for further analysis.

To quantify the degree of association, we used a mutual information based score (RNMI) as implemented in the PARIS software [18]. Briefly, the score measures correlation between each mutation feature and the KIN193 response status of cell-line (see Methods for details). Of note, many genetic mutations were significantly associated ($\text{abs}(\text{RNMI score}) > 0.25$) with KIN-193 resistance, whereas few mutations were strongly associated with KIN-193 sensitivity (Figure 1A and Supplementary Table 1B).

We started by focusing on those genes whose mutations have previously been causally linked to cancer. For this, we utilized the cancer genes census in the COSMIC database, which has catalogued 616 such genes [19]. We identified 255 of these genes in our 5143-feature list and used the RNMI metric to rank them by their significance of association with KIN-193 resistance (Supplementary Table 1B). A list of the top ranked features and their association with KIN-193 response in cells is shown in Figure 1B. One of these features was mutation in the KRAS gene (RNMI = -0.246). A closer examination indicates most of the identified mutations are oncogenic activation (Supplementary Figure 1A, 1B). Consistent with our analysis, KRAS gain of function activations have previously been associated with resistance to pan PI3K inhibition [20] and switch from p110 β to p110 α dependence [21]. Another example is the presence of a PTCH1 mutation. Loss of function PTCH1 mutations has been shown to contribute to uncontrolled SMO activity, which in turn leads to constitutive Hedgehog signaling [22]. Interestingly, Buonamici et al. showed in Medulloblastoma that a PI3K inhibitor can delay the resistance to SMO-antagonist [23]. Our study indicates that the cross-talk between Hedgehog signaling and PI3K pathway goes both ways. While anecdotal, these analyses suggest our predictions are consistent with existing knowledge in the literature. Notably, we also identified a number of novel features associated with KIN-193

resistance, such as CAMT1 and HIF1A, suggesting these factors may also be involved in mediating the PI3K pathway activity. Of note, two subunits of PI3K, PIK3CB and PIK3R1, were also frequently mutated in the resistant cell-lines with frequency at 60% and 50% respectively (details in Supplementary Table 2), although the functional consequence is still unclear. In the case of PIK3CB the mutations were likely to be loss of function which might be expected to yield resistance.

In addition, we carried out Gene Set Enrichment Analysis (GSEA) to identify pathways that are significantly associated with KIN-193 resistance, even though individual member genes may not be (Table 1, Supplementary Figure 2). Of note, we found mutations in YAP target genes (Cordenosi YAP conserved signature) to be enriched in KIN-193 resistance associated features (p -value = 0.026), although the association becomes statistically insignificant after correction for multiple hypothesis testing. YAP along with TAZ plays a central role in the Hippo pathway [24] and this finding indicates a cross-talk between YAP/Hippo and PI3K pathway in PTEN mutated tumors.

A mutation signature can predict PTEN mutated cell-lines that are resistant to KIN-193

Motivated by our previous analysis, we set out to build a simple statistical model to predict KIN-193 sensitivity based on the genetic profiles. To this end, we defined a cumulative score for each cell line as follows. First, we created a ranked list of genetic features according to the RNMI values (Supplementary Table 1B). Second, we selected the top and bottom 50 features, as they are most associated with KIN-193 sensitivity or resistance. Finally, we summarized the overall effect by evaluating the difference between the total number of sensitivity-associated mutations and resistance-associated mutations. This cumulative score is used to predict the KIN-193 treatment outcome for each cell line.

We quantitatively evaluated the accuracy of this simple method to predict KIN-193 treatment outcome by using the 10-fold cross-validation approach. Specifically, we divided the 67 cell lines with known KIN-193 status into 10 groups of roughly equal size while maintaining the class balance (85% sensitive, 15% resistant, see Supplementary Figure 3). We next made 10 unique combinations of training and test datasets such that each cell-line was tested once (by combining 9 datasets as training and testing the remaining 10th). The prediction accuracy was evaluated by using Precision-Recall curves (Figure 2A, 2B). This procedure was repeated 25 times for robustness evaluation. Strikingly, this simple method is very accurate for predicting resistant cell lines: the median precision level remains 100% at the 50% recall cutoff. For comparison, we also considered a number of more complex models: 1) a weighted cumulative score approach, 2) a GLMNET (type of penalized linear regression) approach, and 3) a GBM (Gradient Boosting: typically an ensemble of decision trees) approach (see Methods for details). Interestingly, these more sophisticated models do not lead to better performance (Figure 2A, Supplementary Figure 4).

Defining a gene expression signature to predict KIN-193-sensitive PTEN WT cell-lines

While PTEN loss is strongly associated KIN-193 sensitivity, a large number of PTEN WT cell-lines are sensitive to KIN-193, suggesting there are additional genetic and non-genetic factors that are yet to be identified. To this end, we further integrated the gene expression data in the GDSC database in order to identify transcriptomic signatures that are associated with KIN193 sensitivity (Supplementary Figure 5). To search for a common signature, we hierarchically clustered the PTEN-mutated sensitive cell-lines based on their gene expression patterns. We found 17 cell-lines to form a core cluster (Spearman rank correlation > 0.7) using two different approaches (Figure 3, Supplementary Figure 6) of which 15 cells lines are derived from Gliomas.

We next determined genes that are differentially expressed between PTEN WT (967 cell-lines) and the 17 cell-lines comprising the core cluster of PTEN mutated-sensitive cell-lines, by employing Wilcoxon signed-rank sum test (Figure 3A). We ran GSEA to determine the main pathways enriched in genes over-expressed in PTEN WT sensitive and PTEN mutated sensitive cell lines (Figure 3B). We observed genes up-regulated due to YAP1 over-expression to be enriched in the PTEN mutated-sensitive cell-lines. This link between YAP pathway and PI3K signaling that we found both through mutation and expression analysis points towards a potential mechanism by which PTEN-null cell-lines can become resistant to KIN-193. We next filtered the significantly differentially expressed genes to 203 using a p-value cutoff of $1E-7$. Calculating median expression for the 203 genes across the 17 KIN193-sensitive cell-lines, we were able to define a KIN193 sensitive signature.

We next tried to determine whether there are PTEN WT cell-lines that have similar transcriptional landscapes to the PTEN-mutated sensitive cell lines. Among the 64 PTEN WT cell-lines that showed high similarity (Spearman rank correlation > 0.7) and were predicted as KIN-193 sensitive (Figure 3C), 57 were previously tested for KIN-193 sensitivity by GDSC, and 34 (~60%) were determined to be sensitive (compared to 36% expected by chance). Only 3 (~5%) were predicted as sensitive but experimentally determined to be resistant (compared to ~18.5% expected by chance). This analysis suggests that a gene expression signature similarity could explain the KIN-193 sensitivity of a significant fraction of PTEN WT cell lines (Figure 3D, 3E).

Analysis of an external dataset validates our prediction model

As an external validation, we applied the cumulative score described in the previous section to predict KIN-193 sensitivity for cell lines in the CCLE database [14]. Among the 100 features selected from the GDSC database, 97 were also assayed in the CCLE database. Therefore, we calculated the cumulative score based on these 97 features (see Methods for details). 6 PTEN-null CCLE cell-lines were predicted as KIN-193 resistant (by MIA-RNMI model).

Since drug-screening information is not available for the CCLE dataset, we evaluated the prediction accuracy by searching the literature [12, 15, 16]. For 3 out of the 6 cell-lines, we found experimental evidence supporting their resistance to p110 β targeting drugs (Figure 4, Supplementary Table 3).

We also tested the validity of our gene expression based prediction model for PTEN WT cell lines based on the CCLE database. Among the 1037 cell-lines that have gene expression profiles (through microarray sequencing), 295 were previously tested for KIN-193 drug sensitivity [12]. Due to the platform differences, only 136 of the 203 genes that were associated with KIN-193 sensitivity in GDSC were profiled in CCLE. Based on the expression level of these 136 genes, we identified 22 highly correlated cell-lines whose gene expression signature is similar to the 17 cell-line core cluster in GDSC (Spearman rank correlation > 0.55) (Figure 5A). In keeping with expectation, the predicted cell-lines were enriched for mutation in PTEN (7 cell-lines: ~32%, p-value = 0.003) (Figure 5B). Strikingly, 16 out of the 22 cell-lines are indeed KIN-193 sensitive, whereas the remaining 6 cell lines have ambiguous outcome (Figure 5C). These analyses provide additional support for the validity of our prediction model based on genetic and transcriptomic signatures.

Discussion

The PI3K pathway is often aberrantly activated in cancer. During the past decade, several small molecules targeting the Class I PI3K have been investigated for their applicability as clinical drugs [12, 25-28]. However, the simple pan-PI3K approach has failed to provide effective clinical outcome due to both the combined toxicity of inhibiting all isoforms and the aberrant activation of alternate signaling pathways [29, 30]. It is also now well established that the different isoforms of PI3K perform distinct functions and therefore drugs targeting specific PI3K isoforms should be both more effective and cause fewer side-effects. Indeed, the PI3K isoform specific drug: Idelalisib (targeting p110 δ) was

recently approved for hematologic malignancies (<https://www.cancer.gov/about-cancer/treatment/drugs/fda-idelalisib>). Another drug BYL719 targeting p110 α is currently being tested in the clinical setting and is expected to gain FDA approval in ER positive PI3CA mutant cancers [25]. Though there are advantages to such targeted therapeutic approaches, often a mechanistic knowledge of the action of drug is required for successful translational applications. Moreover, precise tumor signatures are essential for defining the target patient population and successful clinical testing. In the context of p110 β specific inhibitors KIN-193 has been shown to be especially potent in PTEN null tumors [12]. However, a number of PTEN null tumors are resistant to KIN-193, and many PTEN WT tumors are sensitive to KIN-193. Here we have presented a systematic approach to determine the KIN-193 sensitivity by integrating genetic, transcriptomic, and drug screening information that has recently become available.

Analyzing PTEN-null cell-lines, we identified several secondary genetic features associated with KIN-193 resistance. Some of the features are well-known, such as KRAS mutations (switches dependence of a PTEN null cell from p110 β to p110 α isoform) [21] and TNK2 mutations (aberrantly activates AKT pathway) [31], whereas others are previously unrecognized, such as Hedgehog signaling and the YAP/hippo pathway. Our results are therefore able to inform on the genetic heterogeneity of PTEN-null KIN-193 resistant tumors.

It is important to recognize that p110 β dependency is not limited to PTEN-null cell lines. Little is known about how PTEN WT cell lines confer KIN-193 sensitivity. Our integrative analysis provides a plausible explanation that transcriptomic profile similarity to PTEN-null cell-lines may lead to similar drug sensitivity phenotype, although this only accounts for a small fraction of the cell lines.

In summary, our analysis has provided new insights into molecular signatures associated with KIN-193 sensitivity, which in turn may provide a useful guide for developing precise treatment methods for cancer.

Acknowledgements

We thank Dr. Lan Jiang for his initial contribution to this project and members of the Yuan and Roberts Laboratories for helpful discussions. This research was supported by NIH grant R01CA187918 (to T.M.R. and G.-C.Y.).

References

1. Engelman JA, Luo J, Cantley LC: **The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism.** *Nat Rev Genet* 2006, **7**:606-619.
2. Vivanco I, Sawyers CL: **The phosphatidylinositol 3-Kinase AKT pathway in human cancer.** *Nat Rev Cancer* 2002, **2**:489-501.
3. Thorpe LM, Yuzugullu H, Zhao JJ: **PI3K in cancer: divergent roles of isoforms, modes of activation and therapeutic targeting.** *Nat Rev Cancer* 2015, **15**:7-24.
4. Zhao L, Vogt PK: **Class I PI3K in oncogenic cellular transformation.** *Oncogene* 2008, **27**:5486-5496.
5. Samuels Y, Wang Z, Bardelli A, Silliman N, Ptak J, Szabo S, Yan H, Gazdar A, Powell SM, Riggins GJ, et al: **High frequency of mutations of the PIK3CA gene in human cancers.** *Science* 2004, **304**:554.
6. Kan Z, Jaiswal BS, Stinson J, Janakiraman V, Bhatt D, Stern HM, Yue P, Haverty PM, Bourgon R, Zheng J, et al: **Diverse somatic mutation patterns and pathway alterations in human cancers.** *Nature* 2010, **466**:869-873.
7. Dbouk HA, Khalil BD, Wu H, Shymanets A, Nurnberg B, Backer JM: **Characterization of a tumor-associated activating mutation of the p110beta PI 3-kinase.** *PLoS One* 2013, **8**:e63833.
8. Li B, Sun A, Jiang W, Thrasher JB, Terranova P: **PI-3 kinase p110beta: a therapeutic target in advanced prostate cancers.** *Am J Clin Exp Urol* 2014, **2**:188-198.
9. Jia S, Liu Z, Zhang S, Liu P, Zhang L, Lee SH, Zhang J, Signoretti S, Loda M, Roberts TM, Zhao JJ: **Essential roles of PI(3)K-p110beta in cell growth, metabolism and tumorigenesis.** *Nature* 2008, **454**:776-779.
10. Wee S, Wiederschain D, Maira SM, Loo A, Miller C, deBeaumont R, Stegmeier F, Yao YM, Lengauer C: **PTEN-deficient cancers depend on PIK3CB.** *Proc Natl Acad Sci U S A* 2008, **105**:13057-13062.
11. Torbett NE, Luna-Moran A, Knight ZA, Houk A, Moasser M, Weiss W, Shokat KM, Stokoe D: **A chemical screen in diverse breast cancer cell lines reveals genetic enhancers and suppressors of sensitivity to PI3K isoform-selective inhibition.** *Biochem J* 2008, **415**:97-110.
12. Ni J, Liu Q, Xie S, Carlson C, Von T, Vogel K, Riddle S, Benes C, Eck M, Roberts T, et al: **Functional characterization of an isoform-selective inhibitor of PI3K-p110beta as a potential anticancer agent.** *Cancer Discov* 2012, **2**:425-433.

13. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, Bindal N, Beare D, Smith JA, Thompson IR, et al: **Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells.** *Nucleic Acids Res* 2013, **41**:D955-961.
14. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, Wilson CJ, Lehar J, Kryukov GV, Sonkin D, et al: **The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity.** *Nature* 2012, **483**:603-607.
15. Edgar KA, Wallin JJ, Berry M, Lee LB, Prior WW, Sampath D, Friedman LS, Belvin M: **Isoform-specific phosphoinositide 3-kinase inhibitors exert distinct effects in solid tumors.** *Cancer Res* 2010, **70**:1164-1172.
16. Weigelt B, Warne PH, Lambros MB, Reis-Filho JS, Downward J: **PI3K pathway dependencies in endometrioid endometrial cancer cell lines.** *Clin Cancer Res* 2013, **19**:3533-3544.
17. Kuhn M: **Building Predictive Models in R Using the caret Package.** *Journal of Statistical Software* 2008, **Vol 28**.
18. Cowley GS, Weir BA, Vazquez F, Tamayo P, Scott JA, Rusin S, East-Seletsky A, Ali LD, Gerath WF, Pantel SE, et al: **Parallel genome-scale loss of function screens in 216 cancer cell lines for the identification of context-specific genetic dependencies.** *Sci Data* 2014, **1**:140035.
19. Forbes SA, Beare D, Boutselakis H, Bamford S, Bindal N, Tate J, Cole CG, Ward S, Dawson E, Ponting L, et al: **COSMIC: somatic cancer genetics at high-resolution.** *Nucleic Acids Res* 2017, **45**:D777-D783.
20. Castellano E, Downward J: **RAS Interaction with PI3K: More Than Just Another Effector Pathway.** *Genes Cancer* 2011, **2**:261-274.
21. Schmit F, Utermark T, Zhang S, Wang Q, Von T, Roberts TM, Zhao JJ: **PI3K isoform dependence of PTEN-deficient tumors can be altered by the genetic context.** *Proc Natl Acad Sci U S A* 2014, **111**:6395-6400.
22. Pricl S, Cortelazzi B, Dal Col V, Marson D, Laurini E, Fermeglia M, Licitra L, Pilotti S, Bossi P, Perrone F: **Smoothened (SMO) receptor mutations dictate resistance to vismodegib in basal cell carcinoma.** *Mol Oncol* 2015, **9**:389-397.
23. Buonamici S, Williams J, Morrissey M, Wang A, Guo R, Vattay A, Hsiao K, Yuan J, Green J, Ospina B, et al: **Interfering with resistance to smoothened antagonists by inhibition of the PI3K pathway in medulloblastoma.** *Sci Transl Med* 2010, **2**:51ra70.
24. Hong W, Guan KL: **The YAP and TAZ transcription co-activators: key downstream effectors of the mammalian Hippo pathway.** *Semin Cell Dev Biol* 2012, **23**:785-793.
25. Dienstmann R, Rodon J, Serra V, Tabernero J: **Picking the point of inhibition: a comparative review of PI3K/AKT/mTOR pathway inhibitors.** *Mol Cancer Ther* 2014, **13**:1021-1031.
26. Massacesi C, Di Tomaso E, Urban P, Germa C, Quadt C, Trandafir L, Aimone P, Fretault N, Dharan B, Tavorath R, Hirawat S: **PI3K inhibitors as new cancer therapeutics: implications for clinical trial design.** *Onco Targets Ther* 2016, **9**:203-210.

27. Maira SM, Pecchi S, Huang A, Burger M, Knapp M, Sterker D, Schnell C, Guthy D, Nagel T, Wiesmann M, et al: **Identification and characterization of NVP-BKM120, an orally available pan-class I PI3-kinase inhibitor.** *Mol Cancer Ther* 2012, **11**:317-328.
28. Fritsch C, Huang A, Chatenay-Rivauday C, Schnell C, Reddy A, Liu M, Kauffmann A, Guthy D, Erdmann D, De Pover A, et al: **Characterization of the novel and specific PI3Kalpha inhibitor NVP-BYL719 and development of the patient stratification strategy for clinical trials.** *Mol Cancer Ther* 2014, **13**:1117-1129.
29. Carracedo A, Ma L, Teruya-Feldstein J, Rojo F, Salmena L, Alimonti A, Egia A, Sasaki AT, Thomas G, Kozma SC, et al: **Inhibition of mTORC1 leads to MAPK pathway activation through a PI3K-dependent feedback loop in human cancer.** *J Clin Invest* 2008, **118**:3065-3074.
30. Serra V, Scaltriti M, Prudkin L, Eichhorn PJ, Ibrahim YH, Chandarlapaty S, Markman B, Rodriguez O, Guzman M, Rodriguez S, et al: **PI3K inhibition results in enhanced HER signaling and acquired ERK dependency in HER2-overexpressing breast cancer.** *Oncogene* 2011, **30**:2547-2557.
31. Mahajan K, Mahajan NP: **PI3K-independent AKT activation in cancers: a treasure trove for novel therapeutics.** *J Cell Physiol* 2012, **227**:3178-3184.

PATHWAYS	SIZE	ES	NES	Pvalue	FDR
CSR_EARLY_UP.V1_UP	32	0.593	1.349	0.022	1
CORDENONSI_YAP	21	0.62	1.338	0.026026	0.83373
SRC_UP.V1_UP	43	0.571	1.308	0.01	0.90925
TBK1.DF_DN	95	0.505	1.203	0.022	1
PIGF_UP.V1_UP	71	0.504	1.188	0.044	1

Table 1. GSEA analysis predicts the Top 5 pathways associated with KIN-193 resistance.

Figure Legends

Figure 1: A. Binary matrix showing top 10 features associated with KIN193 sensitivity (green) and resistance (purple). Red represents mutation in features whereas white represents WT. The features are ordered by RNMI score and shown across 67 PTEN-null cell-lines from GDSC. **B.** Boxplots of 6 resistance associated features that have been causally linked to cancer (COSMIC database). Cell lines with mutation in these features are more likely to be KIN-193 resistant.

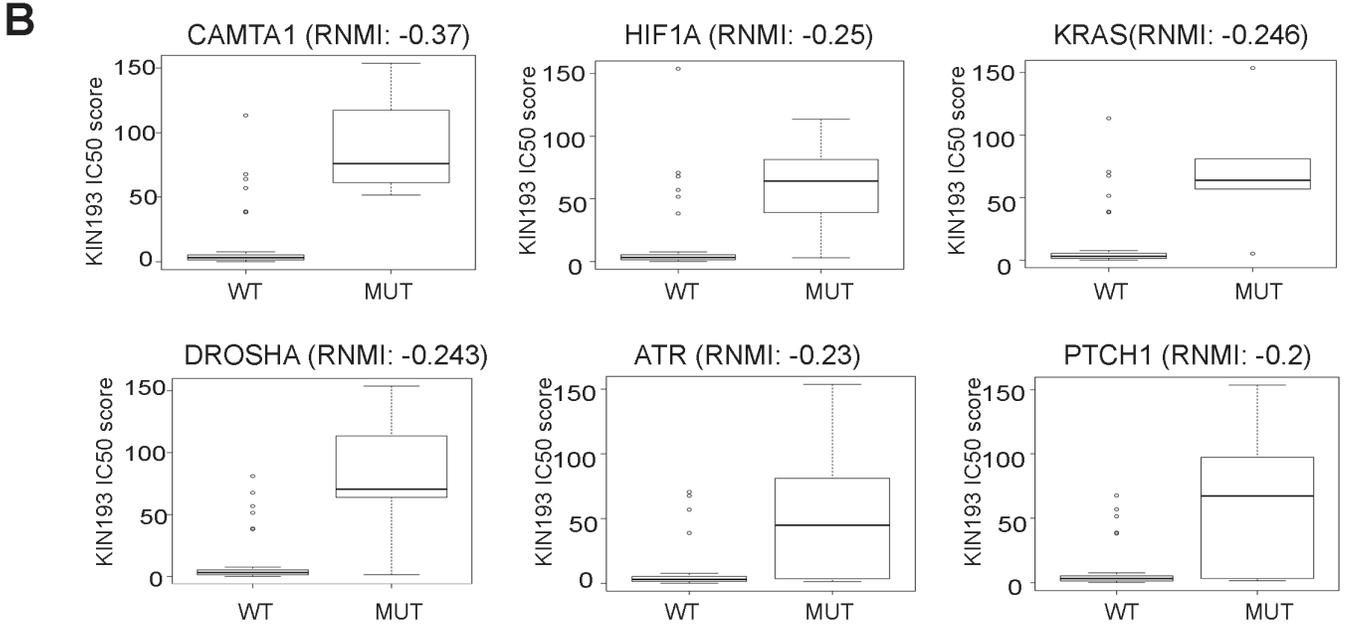
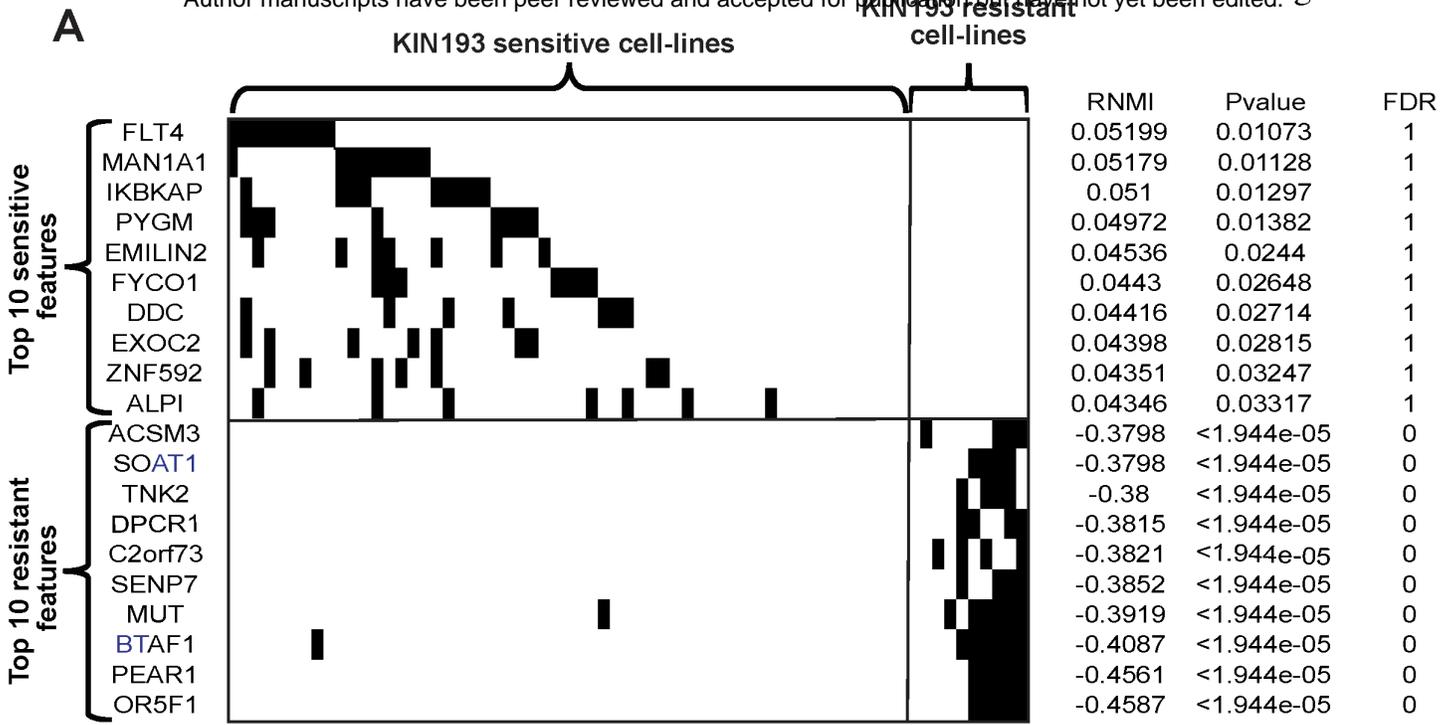
Figure 2: A. Precision Recall curve depicts the accuracy of Mutual Information Aggregate models, using either equal weights (Red) or RNMI score (green) for features. The boxplots represent the distribution of precision scores obtained by making the models 25 times. Random line represents the precision scores expected at those recall values by random chance. **B.** Barplots show that, as the resistance of the cell-lines to KIN-193 increases, so does our accuracy in predicting the said cell-lines as resistant.

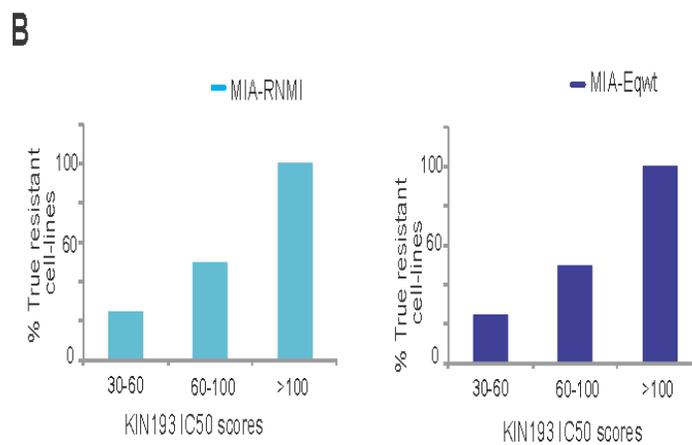
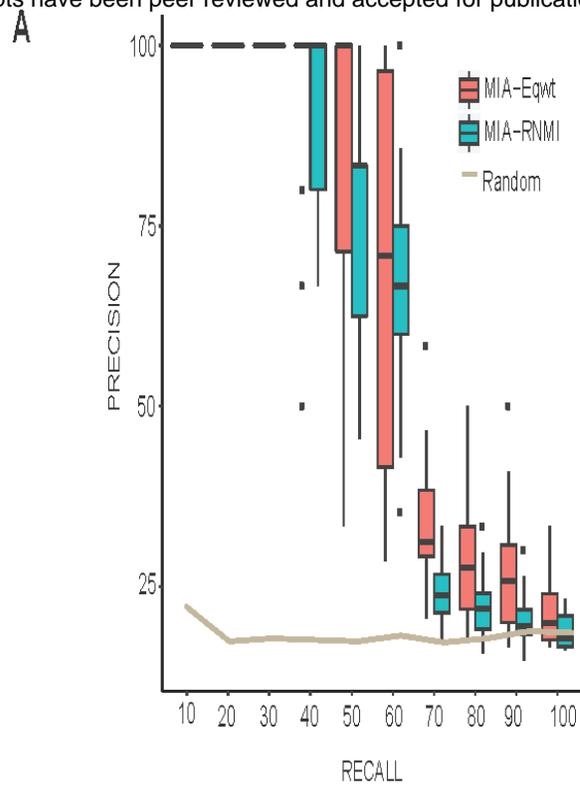
Figure 3: A. Bar plot shows the top 5 genes that are over-expressed in PTEN WT (dark blue) and PTEN null-sensitive (light blue) cell-lines. **B.** GSEA results represent the top pathways enriched in PTEN WT and PTEN null-sensitive cell-lines. **C.** PCA plot shows 903 PTEN WT (grey) and 17 PTEN-null sensitive (red) cell-lines. 64 PTEN WT cell-lines (blue) have a transcriptional signature similar to the PTEN-null sensitive cell-lines (Spearman correlation > 0.75). **D.** Barplot depicts experimental data (IC50 values from GDSC) for 57 PTEN WT cell-lines predicted to be sensitive (Blue). 34 are experimentally validated to be sensitive whereas only 3 are wrongly predicted, The remaining 20 cannot be classified. As control, the corresponding values for all cell-lines are shown. **E.** Logistic regression curve depicts the relationship between the correlation coefficient (PTEN WT with PTEN-null sensitive signature) and the likelihood of a PTEN-WT

cell-line being experimentally determined sensitive (left y-axis). Frequency of cell-lines is represented on the right y-axis.

Figure 4: Mutual information based aggregate models using either equal weights (MIA-Eqwt) or RNMI score (MIA-RNMI) made using GDSC data were applied to the CCLE databases. A binary matrix shows mutation in top 10 features associated with sensitivity (green) and resistance (blue) to KIN193. Data is shown for 6 PTEN-null cell-lines predicted to be resistant by MIA RNMI model, 3 of which are validated by published datasets (dark green), 1 is Ambiguous (light green) and the remaining has unknown status.

Figure 5: The KIN193 gene expression signature (made from GDSC datasets) is used to predict sensitivity in CCLE cell-lines. **A.** Logistic regression curve depicts the relationship between the correlation coefficient (PTEN WT with PTEN-null sensitive signature) and the likelihood of a PTEN-WT cell-line being experimentally determined sensitive (left y-axis). Frequency of cell-lines is represented on right y-axis. **B.** PTEN mutation frequency depicted in % barplot for the 22 WT cell-lines predicted sensitive vs. control of 295 cell-lines. **C.** Matrix showing the prediction score, PTEN mutation status (purple: mutated, white: WT) and experimental data on the 22 CCLE WT predicted sensitive cell-lines.





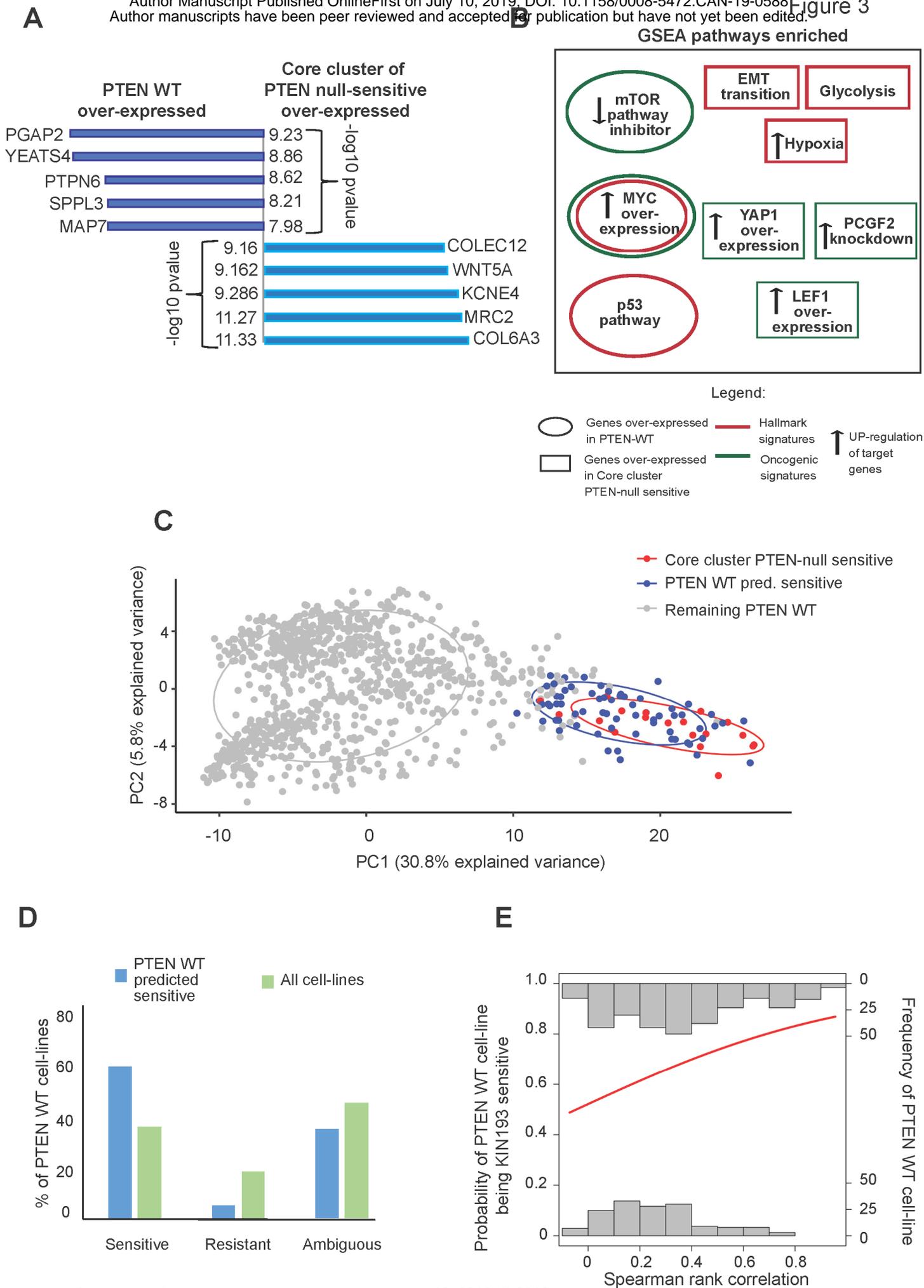
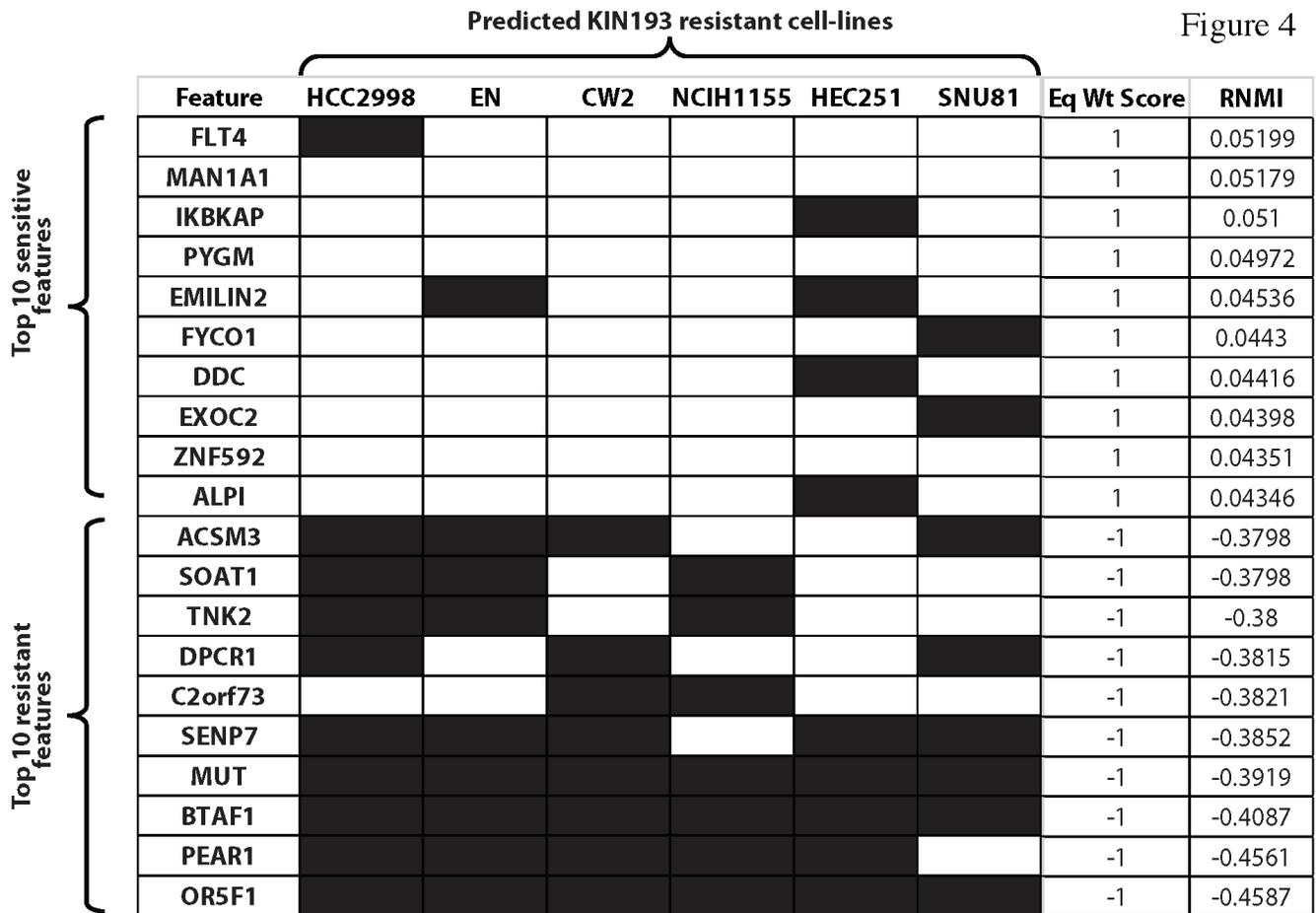
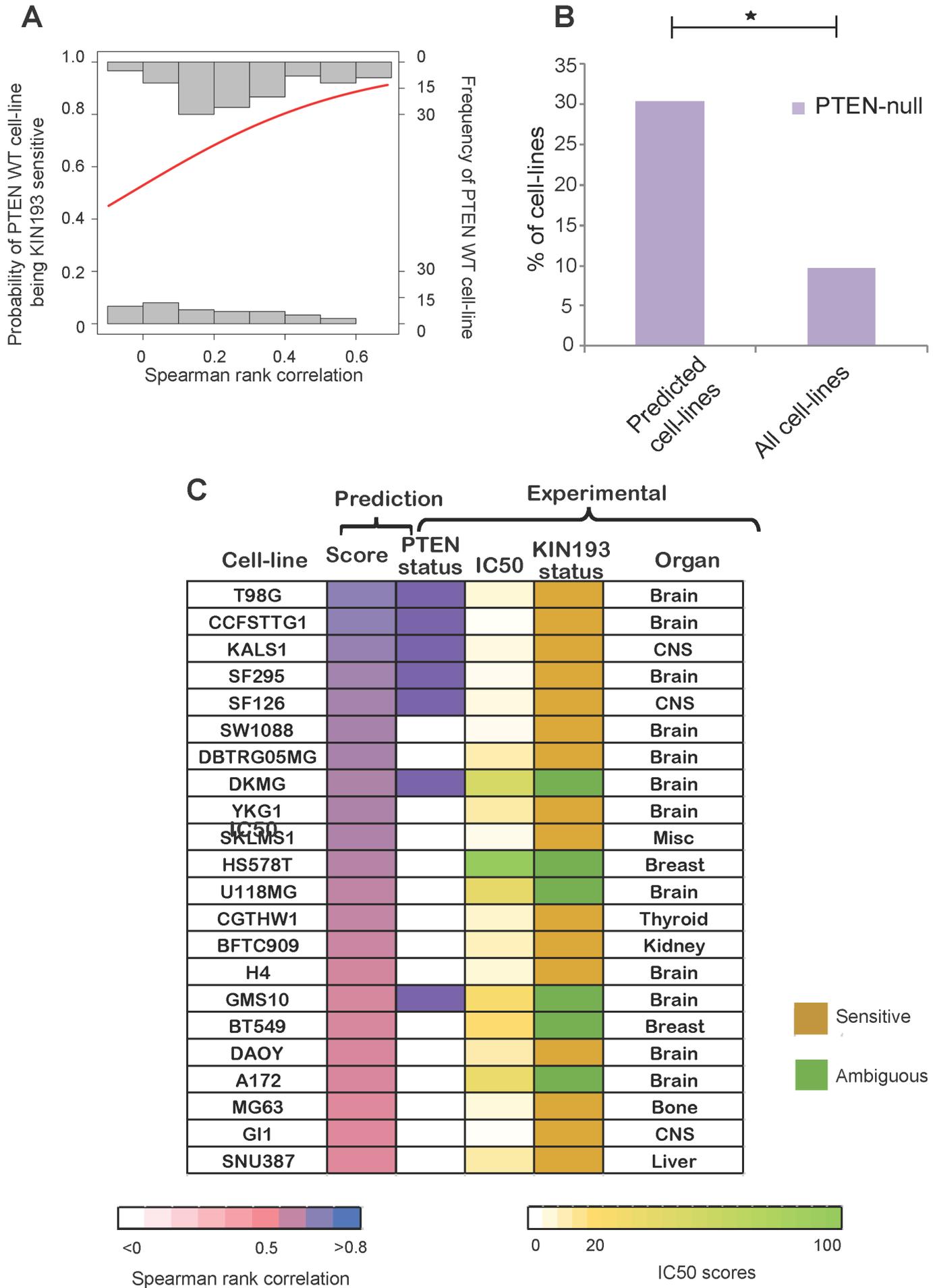


Figure 4



MIA RNMI	-14.1	-13.9	-13.1	-10.1	-8.7	-8.2
MIA EqWt	-40	-39	-34	-26	-9	-11
Experimental	✓	✓		—	✓	

 Strong evidence of resistance
  Ambiguous evidence
  No evidence



Cancer Research

The Journal of Cancer Research (1916–1930) | The American Journal of Cancer (1931–1940)

Molecular profiling establishes genetic features predictive of the efficacy of the p110 β inhibitor KIN-193

Isha Sethi, Zhenying Cai, Thomas M. Roberts, et al.

Cancer Res Published OnlineFirst July 10, 2019.

Updated version	Access the most recent version of this article at: doi: 10.1158/0008-5472.CAN-19-0588
Supplementary Material	Access the most recent supplemental material at: http://cancerres.aacrjournals.org/content/suppl/2019/07/10/0008-5472.CAN-19-0588.DC1
Author Manuscript	Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts [Sign up to receive free email-alerts](#) related to this article or journal.

Reprints and Subscriptions To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, use this link <http://cancerres.aacrjournals.org/content/early/2019/07/10/0008-5472.CAN-19-0588>. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.