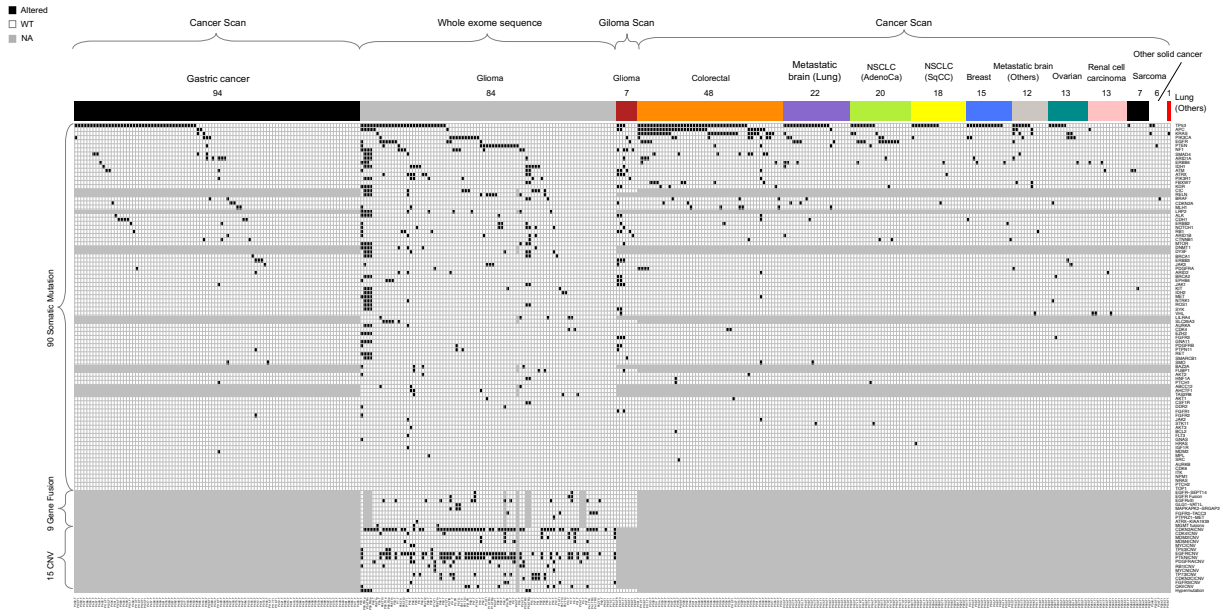
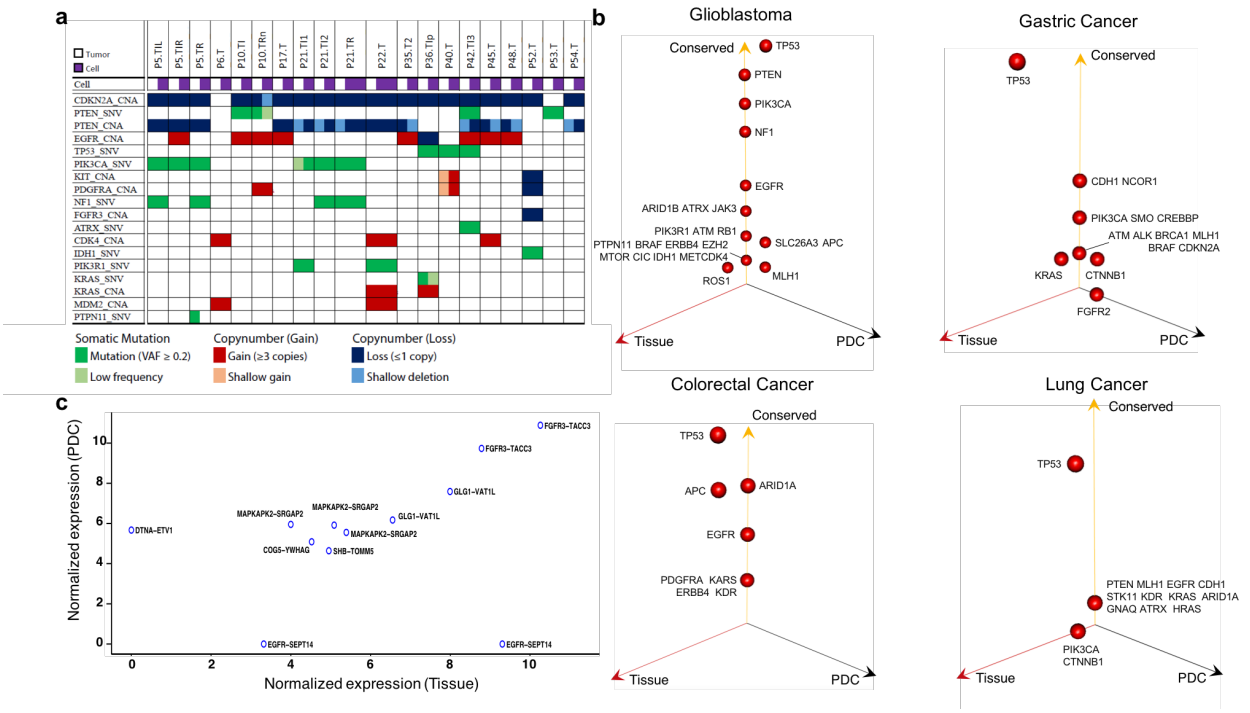
Pharmacogenomic landscape of patient-derived tumor cells informs precision oncology therapy

Jin-Ku Lee, Zhaoqi Liu, Jason K. Sa, Sang Shin, Jiguang Wang et al.

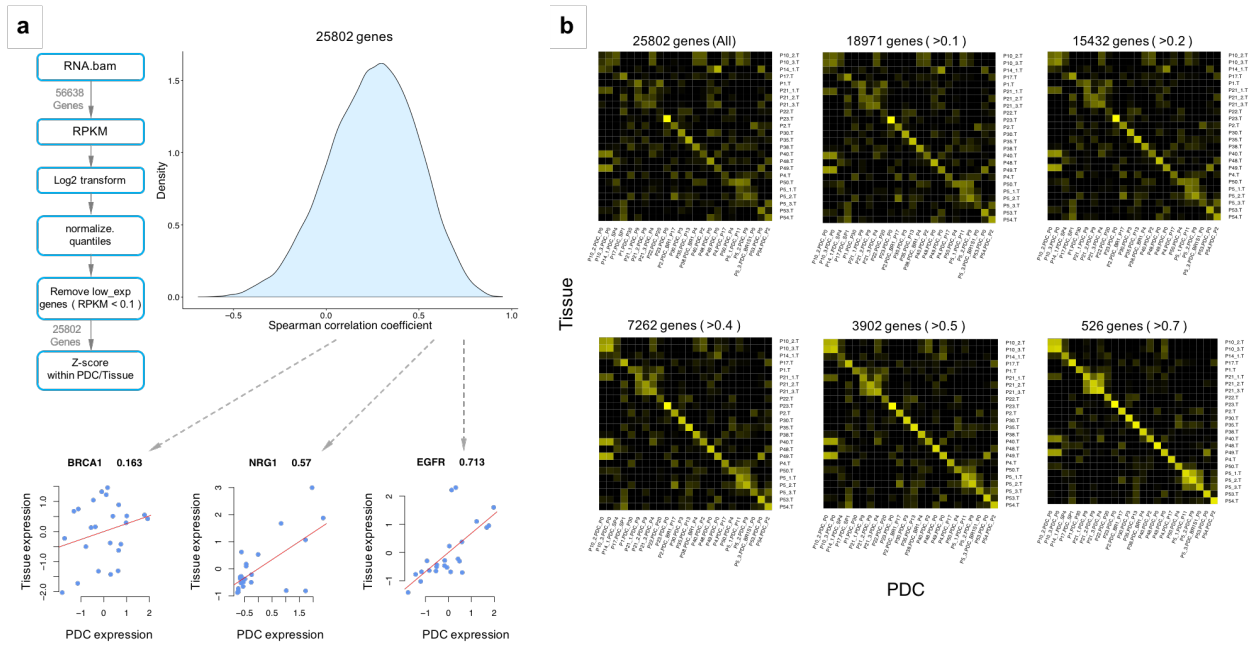
Supplementary note



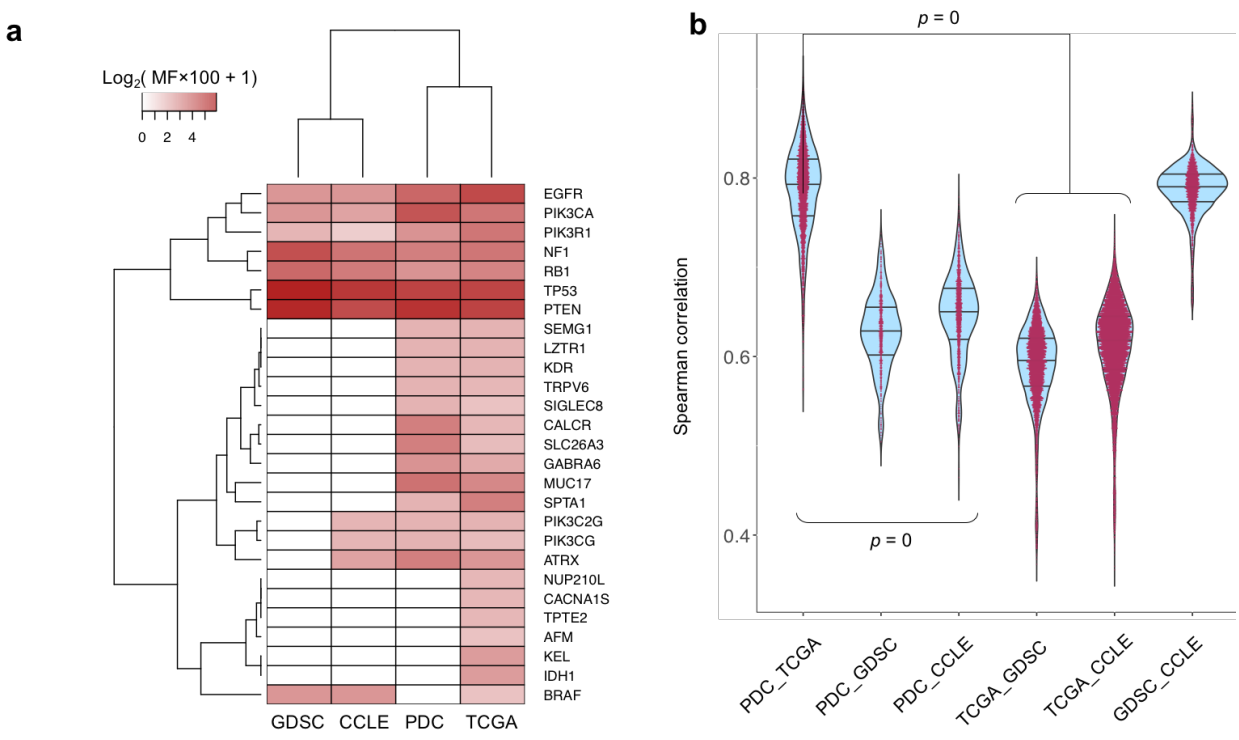
Supplementary Figure 1. Somatic genomic landscape of pan-cancer cohort. The somatic genomic landscape of 115 cancer-driven alterations of samples used for pharmacogenomics interactions. The number of samples was displayed according to the cancer types and sequencing platform, such as CancerSCAN^{TM}, WES and GliomaSCAN^{TM}. Black dots represent the presence of somatic variants of the indicated genes



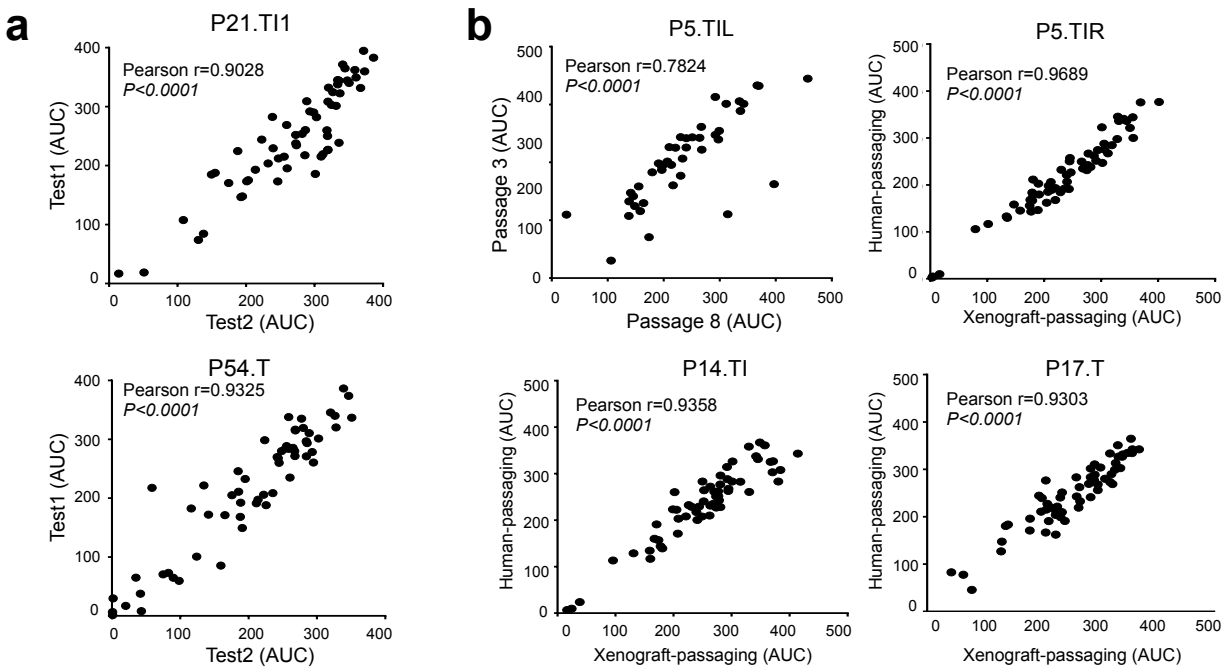
Supplementary Figure 2. Comparison of major cancer-driver genes between tumor tissues and PDCs. (a) Major somatic mutations and copy number alterations in tumors were propagated in tumor spheres. Targeted exome sequencing, comprising cancer and GBM-specific genomic variants was conducted to analyze point mutations and copy number changes in 20 different pairs of parental tumors and derived cultures. **(b)** Three-dimensional bubble plot showing the frequency of somatic nonsynonymous mutations exclusively in tissue (red; left axis), exclusively in PDC (black; right axis), and in common to the two (yellow; upper axis) based on tumor lineage type. **(c)** A correlation plot for fusion gene expressions between tissues and progeny PDCs.



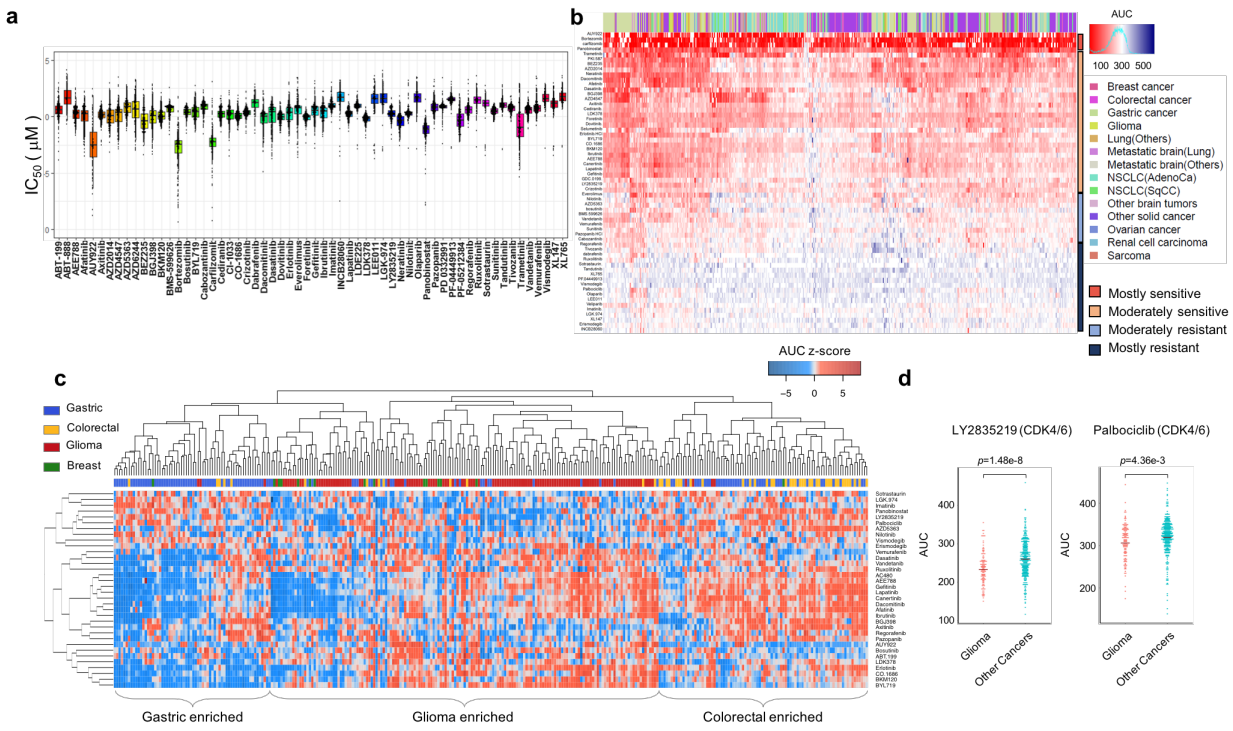
Supplementary Figure 3. Schematic overview for mRNA expression comparison between 24 primary tumors and their matched PDCs. (a) After counting reads from .bam file, expression was measured by RPKM, followed by log2 transformation, and quantile normalization. Genes with low expression level across the cohort are removed. Z score for each gene within the tissue/PDC cohort was calculated to eliminate batch effects. Spearman correlation was used to evaluate each gene's expression similarity between the two systems. **(b)** Based on different cutoffs on expression similarity, we calculated the pairwise spearman correlation coefficient between the filtered expression profiles for primary tissue and PDC on sample level. Spearman correlations between tissue and PDC are shown as a heat map. Paired samples are located along the diagonal.



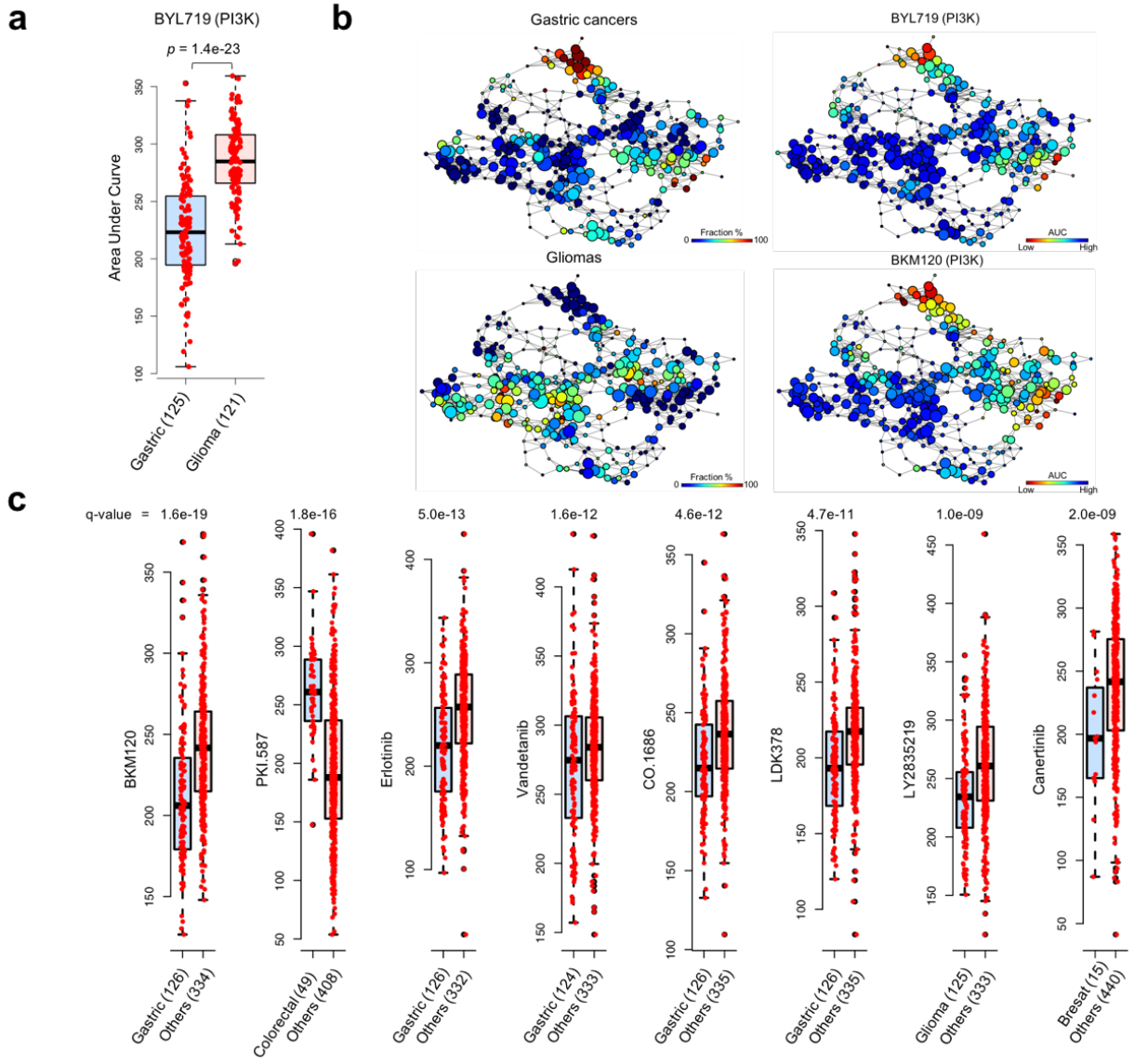
Supplementary Figure 4. Comparisons of molecular profiles among PDCs, tumor tissues (TCGA) as well as cancer cell lines (GDSC and CCLE) on glioma samples. (a) Hierarchical clustering of mutation frequency summarized on PDC, TCGA, GDSC and CCLE glioma samples (n=73, 287, 34, and 68 biologically independent samples, respectively). Recurrent mutations were selected based on their frequency from TCGA cohort. The value in the heatmap is the log2 transformation of mutation frequency (percentage of occurrences) times 100 plus 1. MF: mutation frequency. **(b)** Comparisons of transcriptome similarities among PDC, TCGA, GDSC and CCLE glioma samples (n=24, 172, 35, and 69 biologically independent samples, respectively). To compare between two cohorts, spearman correlations of mRNA expression were calculated for every pair of samples coming from the two different cohorts. The *P* values were derived form Kruskal-Wallis test. Horizontal lines within the violin plot represent 0.25, 0.50, and 0.75 quantiles.



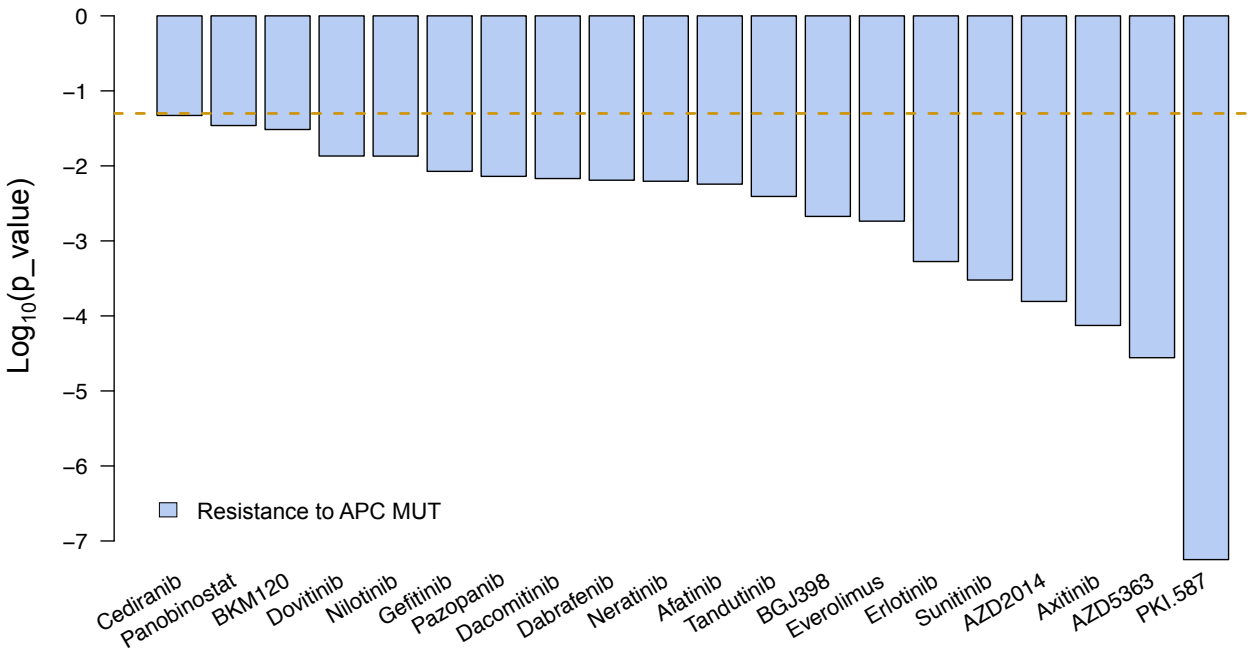
Supplementary Figure 5. Comparison of drug response between multiple tests. Scatterplots of the correlation in AUC values for 60-drug library screened in two different runs of each PDC **(a)**, and in PDCs of different in vitro or in vivo passages **(b).** (n=60 biologically independent samples). A pearson's correlation coefficiency test was performed to analyze the correlation.



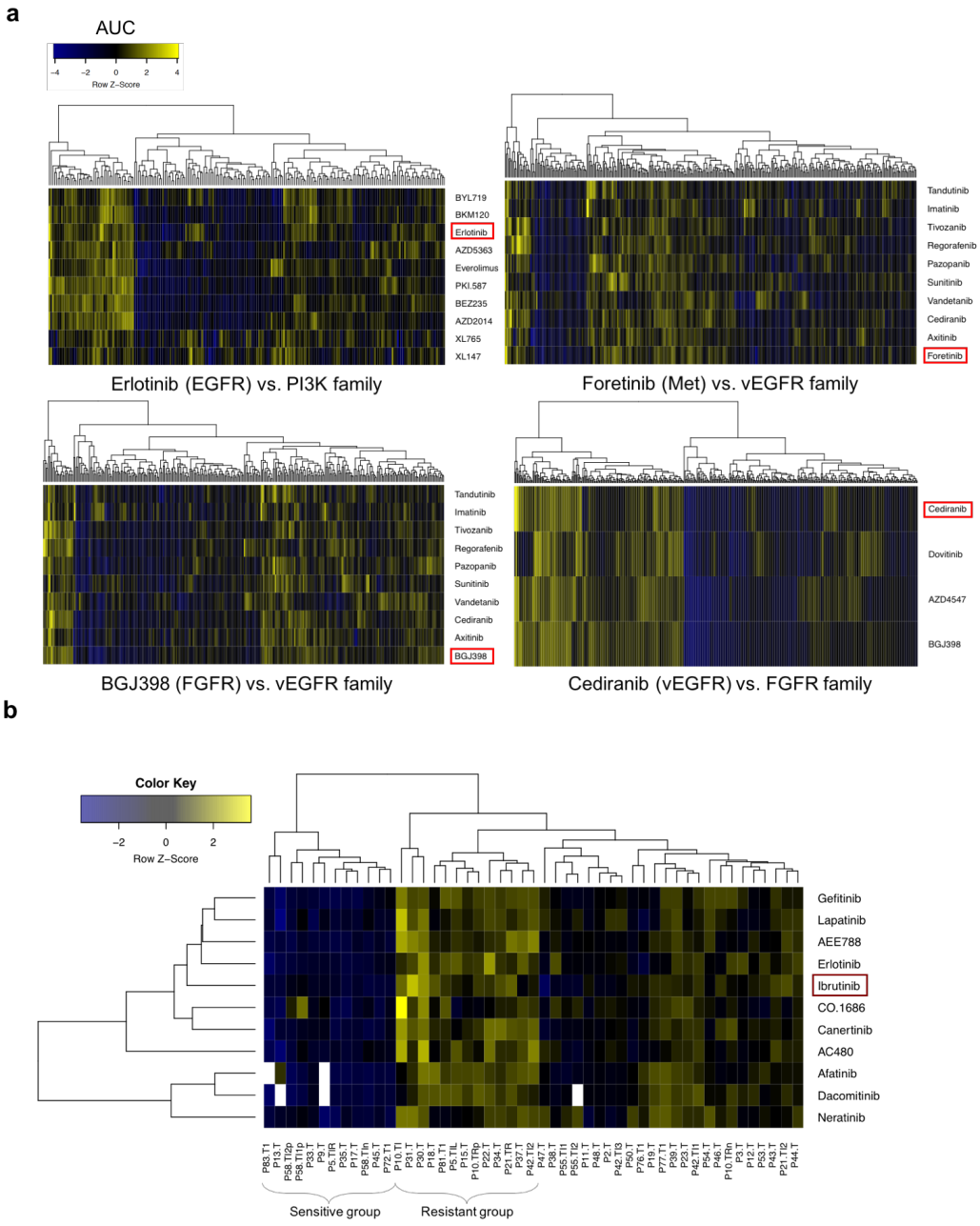
Supplementary Figure 6. Pharmacological drug response of pan-cancer cohort (n=462 biologically independent samples). (a) Box plots of the IC_{50} values (microM) of 60 molecular-targeted drugs. Box plot spans from the first to third quartiles. See also **Supplementary Table 8**. **(b)** Heat map for AUC-based drug sensitivities, annotated by their origin of tissues. The red color on the heat map represents sensitivity, while the blue color represents resistance. Drugs were classified according to their general responses: drugs sensitive to most PDCs, moderately sensitive or resistance, and resistant to most samples. **(c)** Hierarchical clustering of drug sensitivity data on four tumor types (glioma, gastric, colorectal, and breast) with pre-selected drug, which show significant different response on at least one of the four tumors. **(d)** Vertical scattered plot for AUCs comparison of LY2835219 and palbociclib (CDK4/6) between gliomas (n=125 biologically independent samples) and other cancers (n=337 biologically independent samples). The *P* values were obtained using two-tailed unpaired t-test. The horizontal line indicates median.



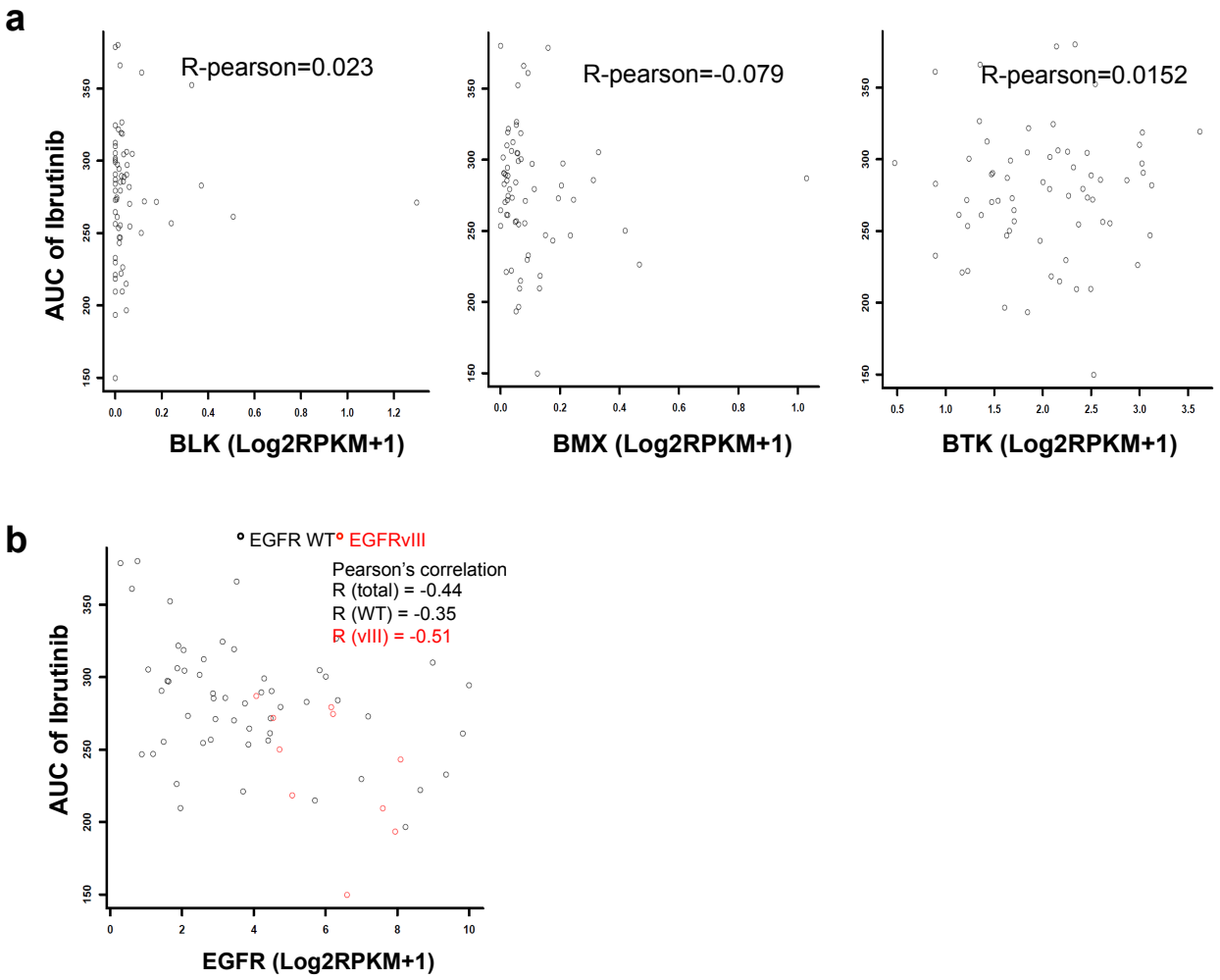
Supplementary Figure 7. Cancer lineage-specific drug sensitivities. (a) Comparison of AUC values of BYL719 between gliomas and gastric PDCs. The *P* value is calculated from two-sided wilcoxon rank-sum test. **(b)** Distribution of GC, glioma PDCs and AUC profile of BYL719 and BKM120 drugs over the topological representation of 462 PDCs. **(c)** Boxplots showing top Tumor type-specific drug interactions. Q-values are calculated from TDA, and multiple tests corrected. Box plots in **(a)** and **(c)** spas from the first to third quartiles and the whiskers represent the 1.5 interquartile range.



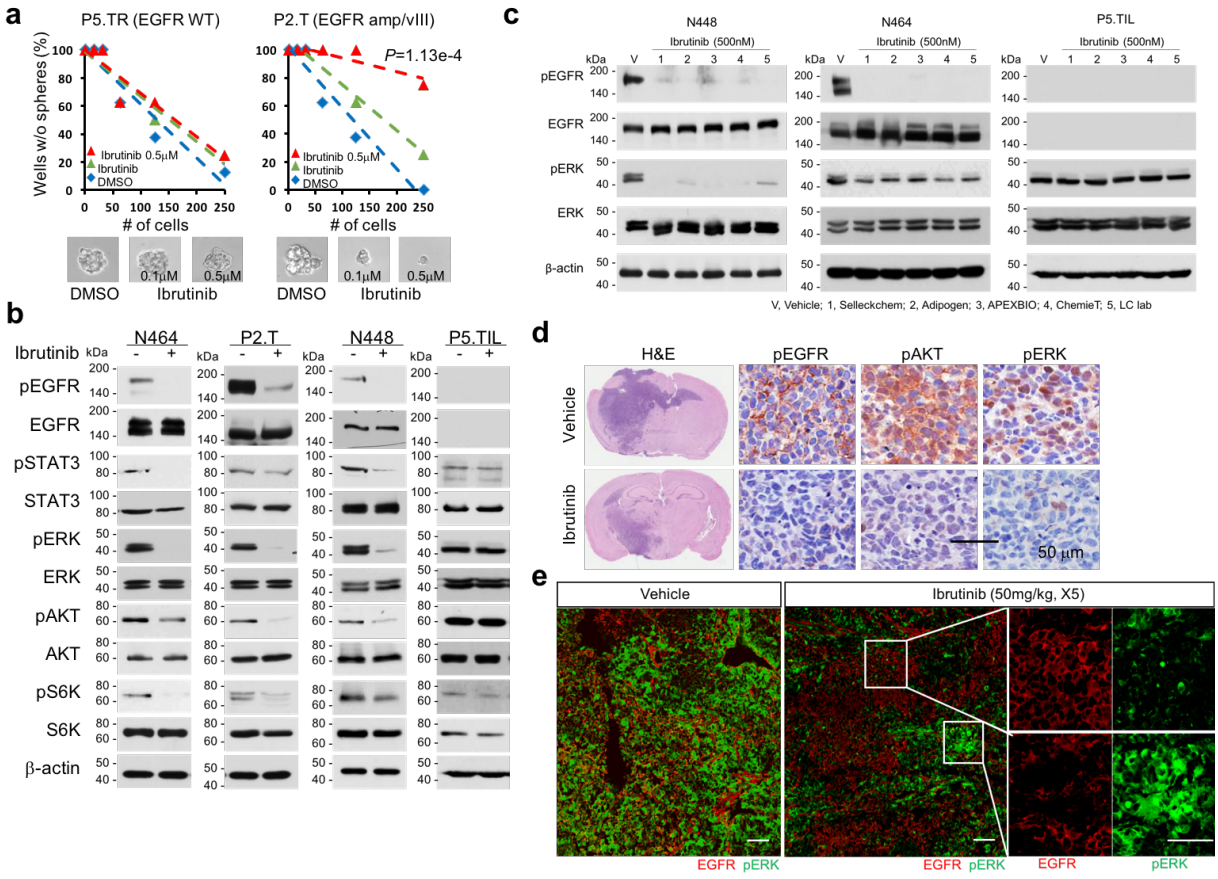
Supplementary Figure 8. *APC* mutation and its drug associations. Waterfall plot enumerates all significant interactions between APC mutation and drug sensitivity (n=360 biologically independent samples). The horizontal dashed lines indicate the threshold of statistical significance of 0.05. The *P* values were obtained using two-sided wilcox rank sum test.



Supplementary Figure 9. Unsupervised clustering of drug response profiles. (a) Unsupervised clustering of drug response profiles of Erlotinib with PI3K inhibitors (Top left panel), Foretinib with VEGFR inhibitors (Top right panel), BGJ398 with VEGFR inhibitors (Bottom left panel), and Cediranib with FGFR inhibitors (Bottom right panel) (n=462 biologically independent samples). **(b)** Unsupervised clustering of response profiles of 10 EGFR inhibitors and Ibrutinib. Clustering was performed on 49 GBM PDCs, detected with EGFR alterations.

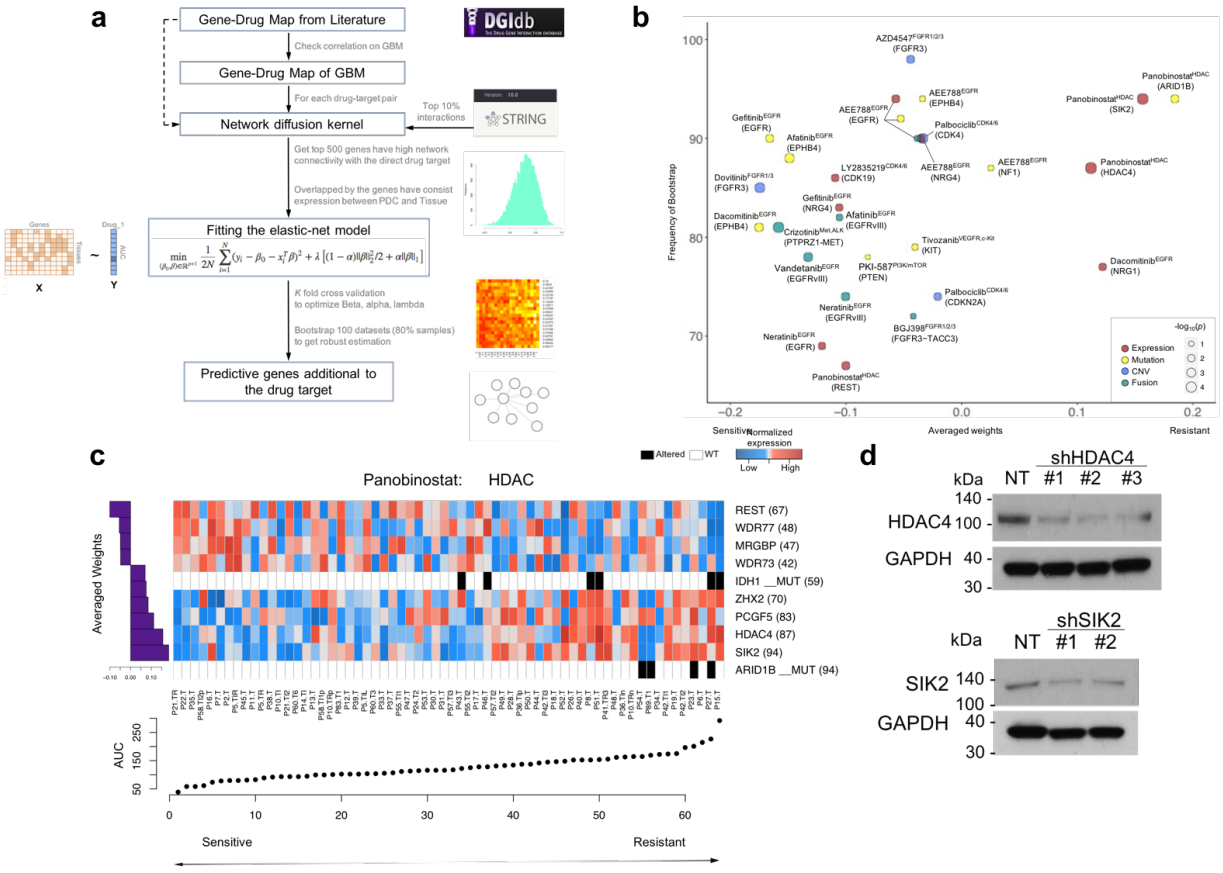


Supplementary Figure 10. Transcriptomic correlates of ibrutinib drug response. (a) Scatter plots of AUC values for ibrutinib and mRNA levels of the indicated genes in GBM PDCs (n=65 biologically independent samples). The Pearson's correlation coefficient analyses of AUC-based sensitivity to ibrutinib and the expression of its original targets such as *BLK*, *BMX* and *BTK* are shown. **(b)** Scatter plot of AUC values for ibrutinib and *EGFR* mRNA levels (n=65 biologically independent samples). *EGFR*vIII-harboring samples (n=10 biologically independent samples) are marked as red circles. A Pearson's correlation coefficiency test was performed to analyze the correlation between ibrutinib sensitivity and *EGFR* expression in the WT *EGFR* (n=55 biologically independent samples) and/or vIII group.



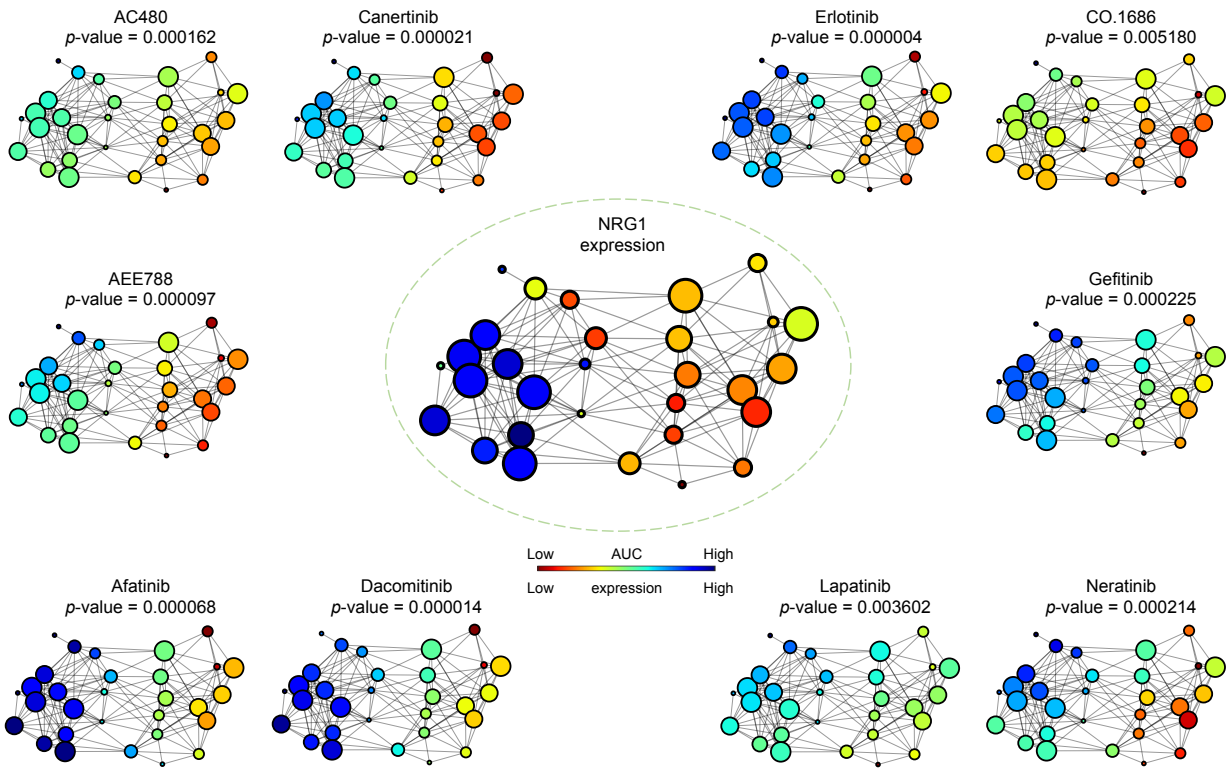
Supplementary Figure 11. Functional role of ibrutinib-*EGFR* association and its biological effects *in vitro* and *in vivo* (a) Effects of ibrutinib on GBM PDC clonogenic growth. *In vitro* tumor sphere forming limiting dilution assays were performed using P5.TR (*EGFR* WT), and P2.T (*EGFR* amp/vIII) cells treated with DMSO control or ibrutinib (100 or 500 nM). Cells were plated at 1-250 cells per well; 24 wells per condition, and cultured for 2 weeks. The p value was obtained using Extreme Limiting Dilution Analysis (ELDA). **(b)** Western blot analysis of pEGFR (Y1068), EGFR, pSTAT3, STAT3, pERK, ERK, pAKT, AKT, pS6K and S6K using lydates isolated from *EGFR* WT (PDC#5-1) and *EGFR* amplification and/or vIII (N464, PDC#2, and N448) harboring PDCs. **(c)** 500 nM of ibrutinib obtained from the indicated vendors (Selleckchem, Adipogen, AEXBIO, ChemiT, and LC lab). **(d)** After completion of 5 cycles of ibrutinib, mice for histological analyses were sacrificed (n=5 per group). Hematoxylin and eosin (H&E), and immunohistochemical assays were conducted in paraffin embedded blocks of the indicated groups. Bar represents 50 microns. **(e)** Subcutaneous tumors of cells with *EGFR* amplification were harvested after the treatment of vehicle or ibrutinib (50mg/kg/day, 5

consecutive days orally; n=5 per group). Representative immunofluorescence images for EGFR and pERK were demonstrated. Scale bar indicates 50 micron. **(b,c)** Data were representatives of three independent experiments with similar results.

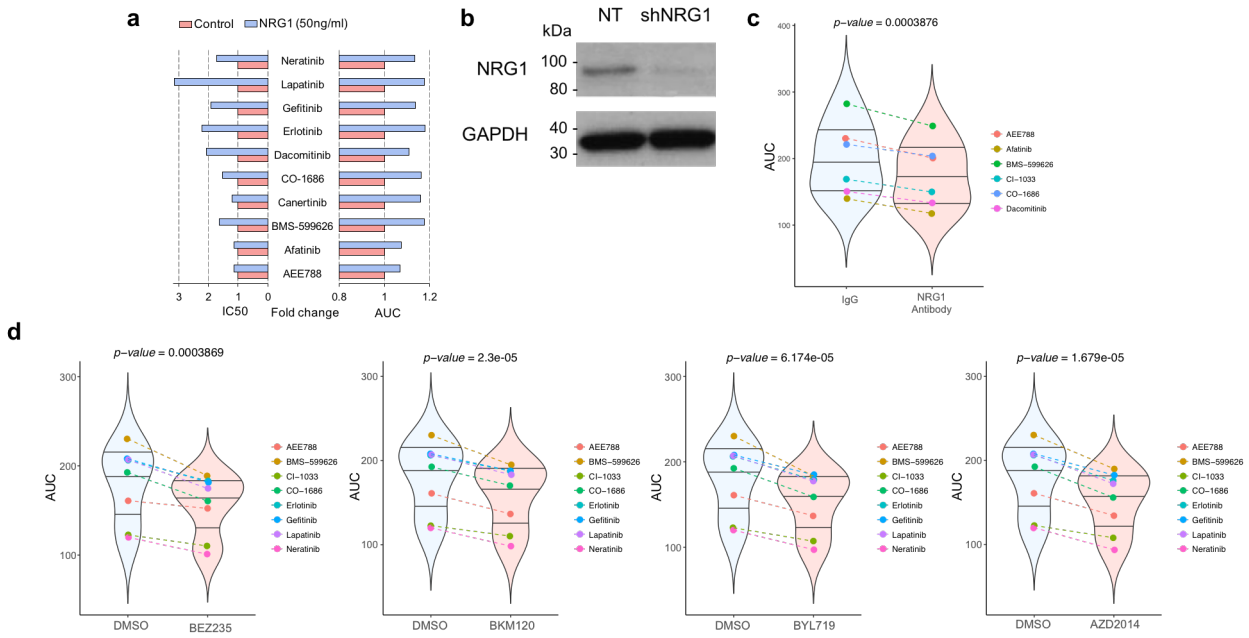


Supplementary Figure 12. Genomic and transcriptomic correlates of panobinostat sensitivity and its experimental validation. (a) A workflow chart of dNetFS pipeline. Generally, for a given drug, the known targeted gene was found from current database. Then, a diffusion kernel was applied to select genes, which has a close interaction with drug targets defined by biological network topology. Next, an elastic net regression algorithm was applied, using genomic alterations, tumor subtyping, and mRNA expression of previously select genes as the input features. Ten-fold cross validations were used to optimize the parameters. Finally, we adopt bootstrapping strategy for 100 times to obtain a robust evaluation of the predictive power of features. **(b)** Selected drug-feature associations identified by the dNetFS are plotted for their frequency and effect size (n=71 biologically independent samples). Node size is proportional to the single drug–feature linear correlation. Nodes are labeled with drug name, targets (upper, right), and genomic features (below, in parentheses). **(c)** dNetFS results of genomic features that predict response to panobinostat. The bottom scatters indicates drug response (AUC) for GBM PDCs. The central heat map shows the top 10 features in the model (continuous z-score for expression, black for binary genetic alteration calls), across all PDCs. Bar plot (left):

averaged weight of the predictive features for sensitivity (bottom) or insensitivity (top). Parentheses indicate the time of appearances (non-zero fitting coefficient) out of the 100 bootstrapping. **(d)** Immunoblot assay demonstrating shRNA-mediated knockdown of *HDAC4* using 3 different shRNA vectors (upper panel) and *SIK2* using 2 different shRNA vectors (bottom panel). Data were representatives of two independent experiments with similar results.

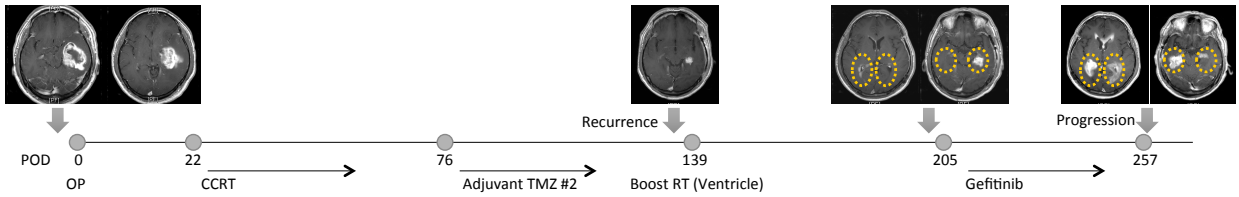


Supplementary Figure 13. Comparison of *NRG1* expression with EGFR inhibitors. Gene expression profiles of *NRG1* and AUC drug response profiles of 10 EGFR inhibitors over the topological representation of EGFR-altered PDCs (n=44 biologically independent samples). The *P* values are computed based on the pearson correlation of the distributions over the network.

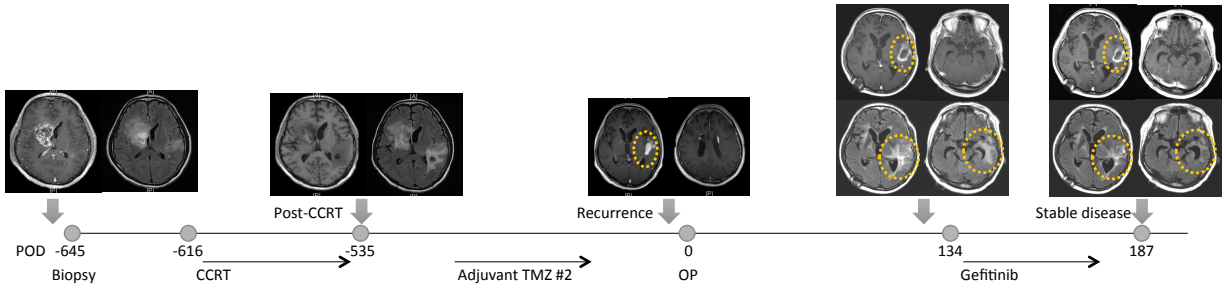


Supplementary Figure 14. Experimental validation of predictive biomarkers for response to EGFR inhibitors. (a) Bar plots represent fold changes of IC_{50} (left) or Area Under the Curve (AUC) of 10 different EGFR inhibitors after 7 days of exposure in a fourfold, seven-point dilution series from 4.88 nM to 20 uM, in a GBM PDC (P5.TIR) harboring *EGFR* amplification/vIII, incubated with NRG1 (50ng/ml), normalized to values of vehicle group. Data were representative of three independent experiments. **(b)** Immunoblot assay demonstrating shRNA-mediated knockdown of NRG1 using shRNA vector. Data were representatives of two independent experiments with similar results. **(c)** Drug response assessment of EGFR inhibitors with NRG1 neutralizing antibody or IgG. Cell viability for each dose was normalized to NRG1 antibody or IgG treated cells only (n=6 independent experiments with 3 technical replicates). The *P* value was calculated by two-sided wilcox rank sum test. Horizontal lines within the violin plot represent 0.25, 0.50, and 0.75 quantiles. NRG1 neutralizing antibody (AF-296) was purchased from R&D Systems. **(d)** Drug response assessment of EGFR inhibitors with individual PI3K-AKT-mTOR (PAM) inhibitors or DMSO. Cell viability for each dose was normalized to each individual PAM inhibitor of DMSO treated cells only. (n=8 independent experiments with 3 technical replicates). The *P* values were calculated by two-sided wilcox rank sum test.

a B1031 (53/M): Glioblastoma (IDH1 WT, EGFR amp with mutations, PTEN deletion)



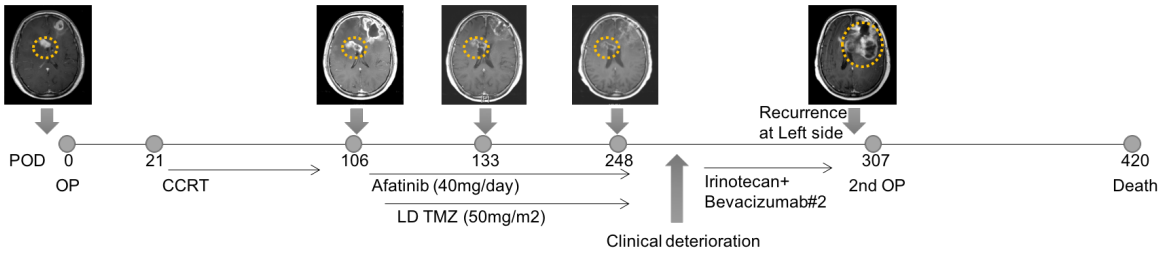
b B1089 (51/M): Glioblastoma (IDH1 WT, EGFR amp with mutations, PTEN deletion)



c

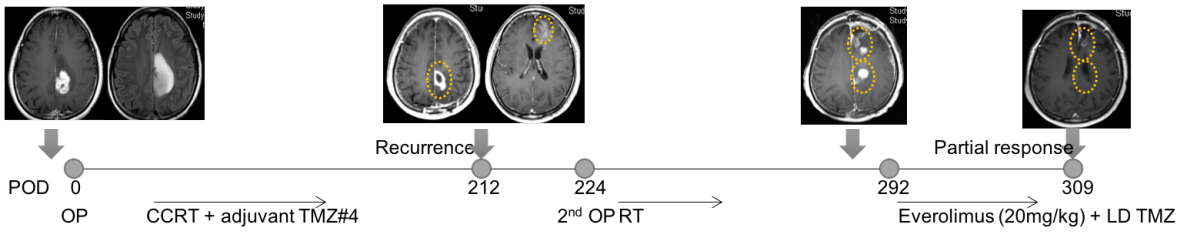
P5 (65/F): Glioblastoma

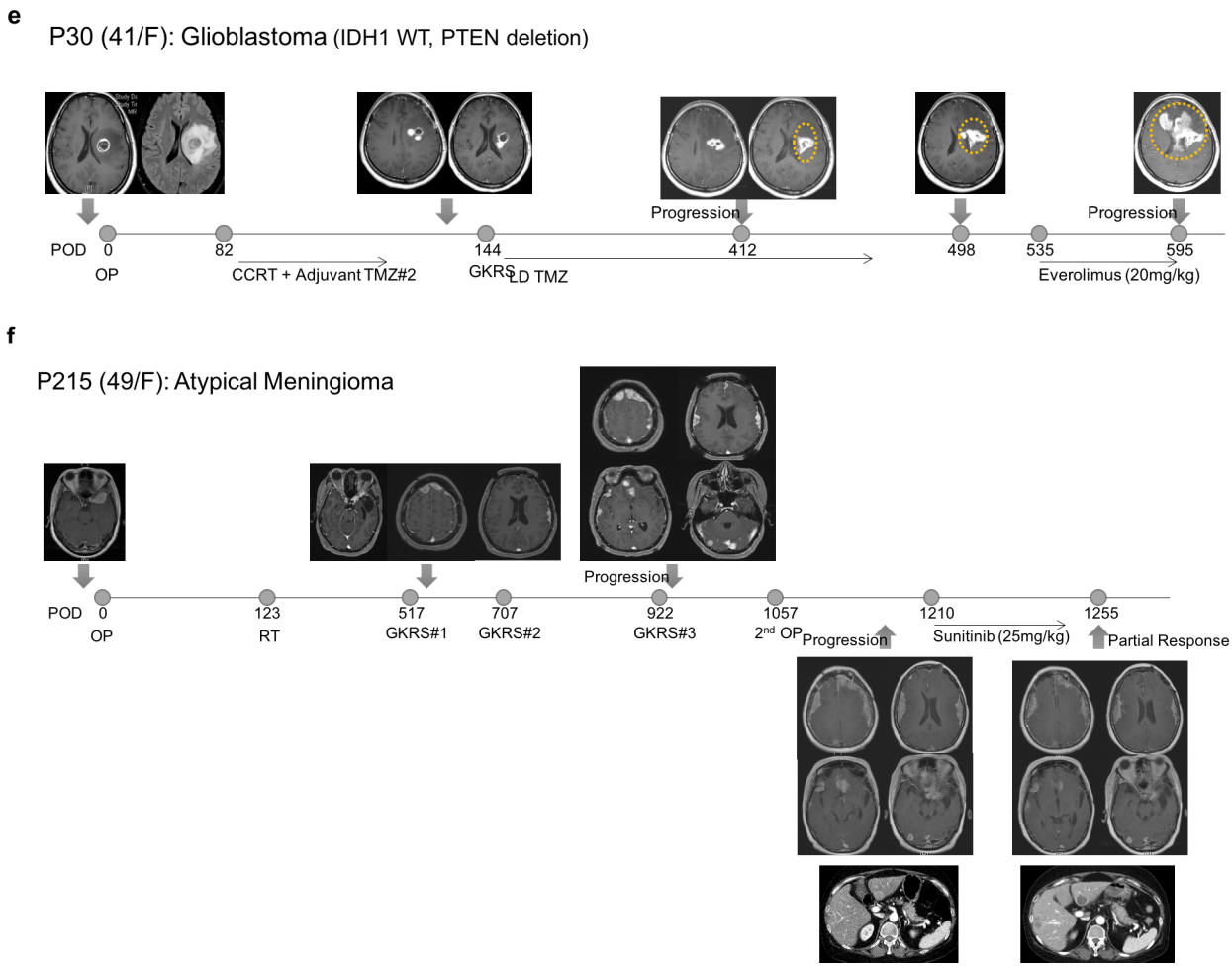
	Genomic profiles
Left	PIK3CA (p.F1016C); NF1 (p.S1112*,p.L386*)
Right	EGFR(p.A289V,p.G403V,p.G598V); PIK3CA(p.F1016C), ARID2 (p.E1314*) EGFR-SEPT14 fusion; EGFR vIII (5.6%); EGFR amp (log2R: 4.15) PTEN del (log2R: -2.9); CDKN2A del(log2R: -2.7)



d

P10 (52/F): Glioblastoma (IDH1 WT, PTEN mutation)





Supplementary Figure 15. Clinical course of targeted-therapy. (a-b) Clinical course of gefitinib treatment on two GBM patients. **(c)** Clinical course and genomic representation of a multicentric GBM patient, treated with afatinib. **(d-e)** Clinical courses of everolimus treatment on two GBM patients harboring *PTEN* mutation **(d)** or copy number deletion **(e)**. **(f)** Clinical course of sunitinib treatment on a patient suffering from atypical meningioma.

Supplementary Table 1. Descriptive characteristics of 462 PDCs used in the drug screening and the status of genomic evaluations

Supplementary Table 2. List of genes and sequencing protocol for CancerSCAN^{TM}, detecting cancer-driven somatic variant

Supplementary Table 3. List of genes and sequencing regions for GliomSCAN^{TM} to analyze glioma-associated as well as cancer-driven mutations.

Supplementary Table 4. The genomic profile of 115 cancer-driven alterations for PDC samples. The list is made up of 91 mutations, 9 gene fusions, and 15 copy number variances. This profile is a binary calls matrix, that '1' indicates the present of alteration, while 'NA' means status unknown. 'NA' could be results from the unavailability of relevant sequencing data, or genomic region not covered by targeted sequencing.

Supplementary Table 5. A list of the 60 drugs used in drug sensitivity screens in PDCs. We catalogued the chemical and/or generic names of the drugs, their respective targets, clinical phases and the blood–brain barrier penetration prediction values based on Lipinski's rule of 5.

Supplementary Table 6. A document for mass spectrometry-based quality assurance data of 60-drug library, provided by Selleckchem.

Supplementary Tables 7-8. AUC (Supplementary Table 7) or IC_{50} (Supplementary Table 8) results of 60 drugs in 462 PDCs, with the exclusion of non-fitted DRC that resulted in a non-convergent or ambiguous curve.

Supplementary Table 9. Tumor type-specific drug association identified using 60 drugs. Wilcox rank sum test was applied to determine the relative differences of drug sensitivity between certain tumor type and all the other samples. Sheet1 gives the significant level as q-value. Sheet2 indicates the Resistance/ Sensitive direction for associations with $q < 0.05$.

Supplementary Table 10. Pearson correlation of tumor type-specific distribution and mean AUC drug response values associated with nodes of topological network of 462 PDCs.

Supplementary Table 11. Pharmacogenomic associations identified from pan-cancer analyses. For each drug, sensitivity data were compared between pan-cancer subgroups, which harbor or not the genomic alteration using Wilcoxon rank-sum test.

Supplementary Table 12. dNetFS results for panobinostat. Features were ranked by their frequency of appearance in the 100 bootstrapping. P.value was calculated for each single feature drug interaction (Pearson correlation test).

Supplementary Table 13. dNetFS results for 10 EGFR inhibitors. Analysis was conducted on the 49 GBM PDCs with *EGFR* alteration. For each inhibitor, only top 20 predictive features were reported based on their frequency of appearance in the 100 bootstrapping. P value was calculated for each single feature drug interaction (Pearson correlation test).

Supplementary Table 14. Clinical response and duration for retrospective studies.