

Supplementary Materials:

Cancer type abbreviations: ACC, Adrenocortical Cancer; BLCA, Bladder Cancer; BRCA, Breast Cancer; CESC, Cervical Cancer; CHOL, Cholangiocarcinoma (bile duct cancer); COAD, Colon Cancer; DLBC, Large B-cell Lymphoma; ESCA, Esophageal Cancer; GBM, Glioblastoma; HNSC, Head and Neck Cancer; KICH, Kidney Chromophobe; KIRC, Kidney Clear Cell Carcinoma; KIRP, Kidney Papillary Cell Carcinoma; LGG, Lower Grade Glioma; LIHC, Liver Cancer; LUAD, Lung Adenocarcinoma; LUSC, Lung Squamous Cell Carcinoma; MESO, Mesothelioma; OV, Ovarian Cancer; PAAD, Pancreatic Cancer; PCPG, Pheochromocytoma & Paraganglioma; PRAD, Prostate Cancer; READ, Rectal Cancer; SARC, Sarcoma; SKCM, Melanoma; STAD, Stomach Cancer; TGCT, Testicular Cancer; THCA, Thyroid Cancer; THYM, Thymoma; UCEC, Endometrioid Cancer; UCS, Uterine Carcinosarcoma; UVM, Ocular melanoma.

Supplementary Tables:

Sup Table 1: Binding energies of top-ranking docked compounds from Drug Bank against C361 amino acid of BRF2.

Drugs	Binding-energy (kCal/mol)
Cetrorelix	-10.3
Goserelin	-9.8
Nafarelin	-9.7
Degarelix	-9.6
Ganirelix	-9.6
Fertagyl	-9.4
Nilotinib	-9.3
Lumacaftor	-8.9
Desmopressin	-8.8

Sup Table 2: Binding energies of top-ranking docked compounds from NCI against C361 amino acid of BRF2.

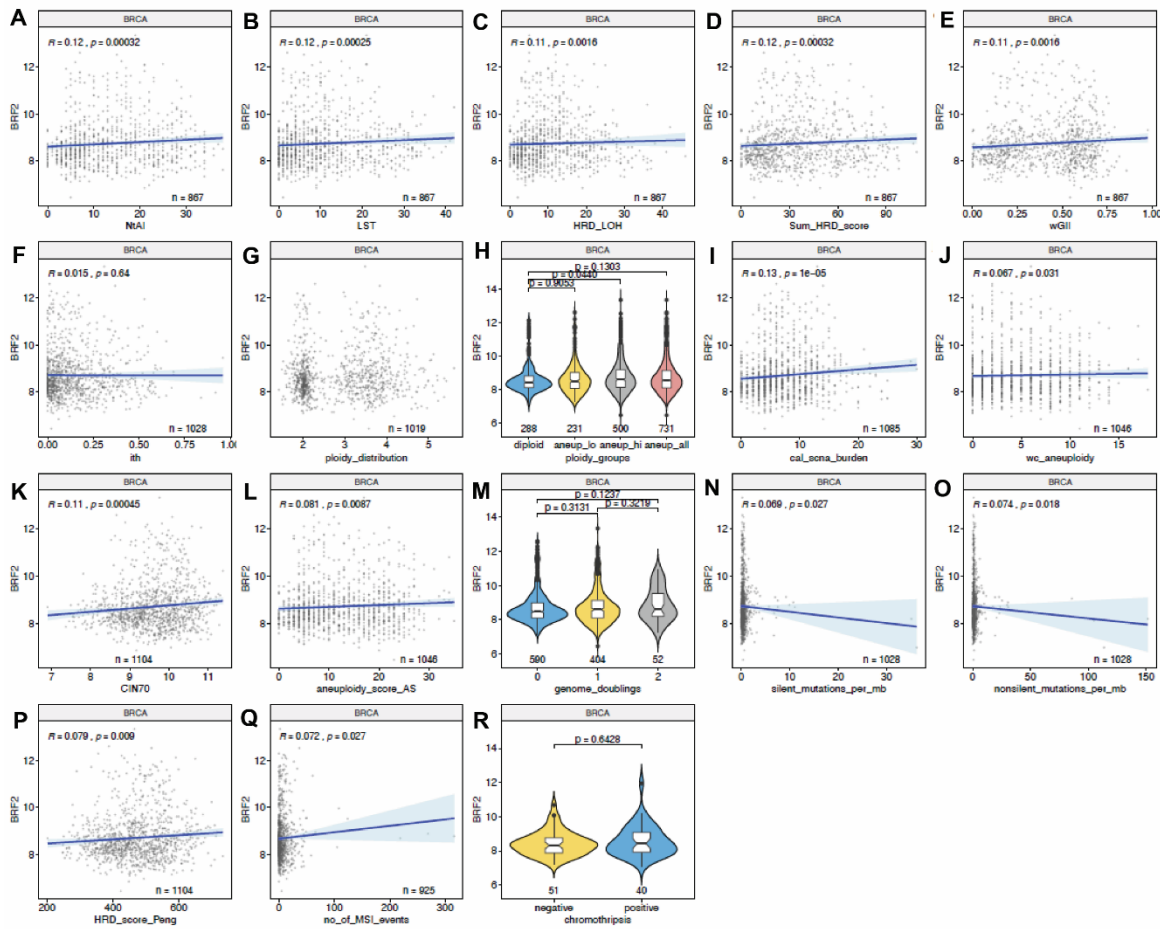
Drugs	Binding-energy (kCal/mol)
Nilotinib	-9.3
Imatinib	-8.2
Sonidegib	-8
Dabrafenib	-7.9
Osimertinib	-7.8
Vismodegib	-7.7
Niraparib	-7.6
Rucaparib	-7.6
Crizotinib	-7.3

Ribociclib	-7.2
Celecoxib	-7.1
Methotrexate	-6.6

Supplementary figure legends

Sup 1 (A): Chart shows percent of alteration frequency (copy number alteration) for amplification (red) or copy number gain (pink) of *BRF2* across various cancers using the cBioPortal for Cancer Genomics. **(B-G):** Oncoprint of cBioPortal analysis shows amplification (red) or gain of function (pink) in *BRF2* for indicated cancer, dataset, and number of patients. Oncoprints from the TCGA and METABRIC breast cancer datasets showing mutually exclusive alterations between *BRF2* and *BRCA1/2*, with upregulation of *BRF2* rarely observed in tumors with *BRCA1/2* loss or downregulation. **(H):** Kaplan-Meier survival analysis for TCGA breast cancer patients for Disease-specific survival (upper) and Overall survival (lower). Patients were split into two groups, with high and low *BRF2* expression levels, split using the median *BRF2* expression level as the cut-off. Numbers of patients at risk are shown at the bottom. **(I):** The amplification or gain of function in *BRF2* comparing with *BRCA1* and/or *BRCA2* loss in breast cancer cell lines from Cancer Cell Line Encyclopedia (CCLE) (n=51, Mann-Whitney U test).

Sup 2

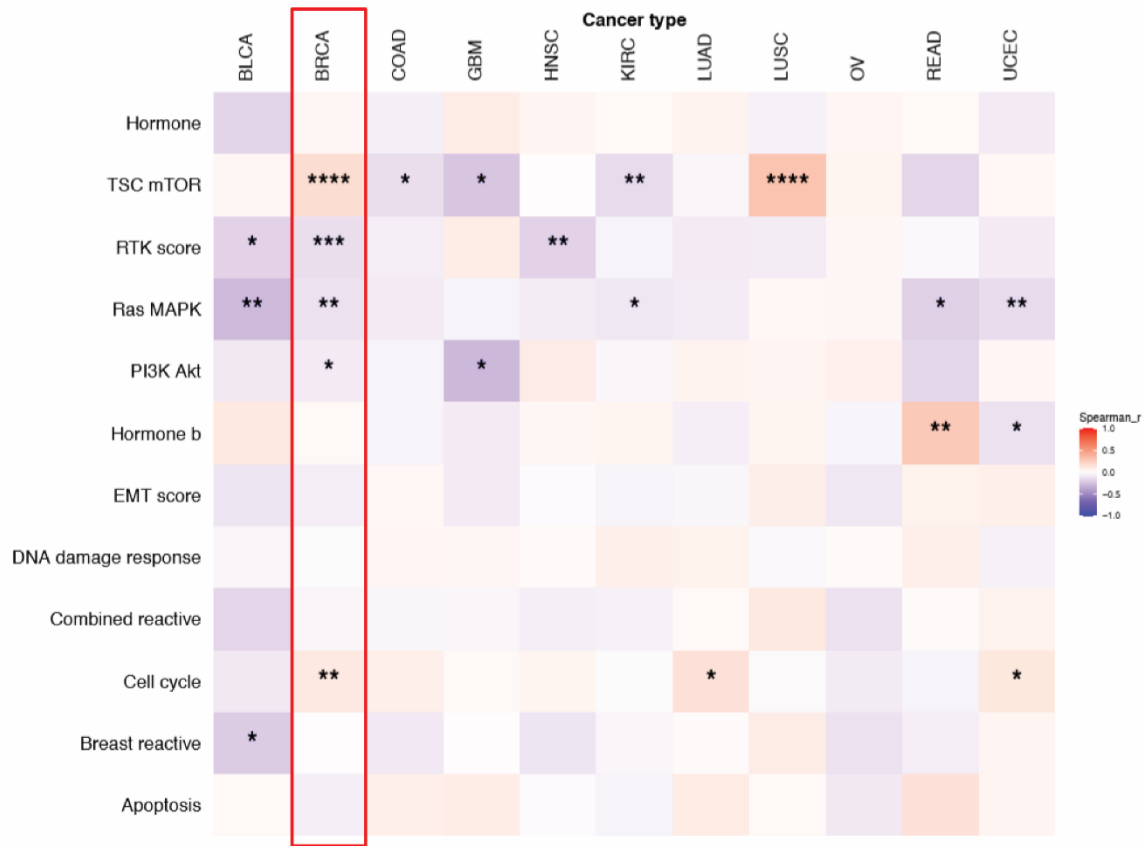


Sup 2A-R: The analyses show correlations between gene expression (from The Cancer genome Atlas (TCGA) RNAseq datasets) and 18 measures of genomic instability (GI). The 18 measures are labelled "a" through "r". For all panels, the y-axis shows the relative expression level of *BRF2*. Details on each of the GI variables on the x-axis of each panel is available in methods. In all scatter plots, linear regression analysis was performed and R and p values are given according to Spearman's rank correlation. Samples sizes, n , are also included. Panel (g) intentionally does not

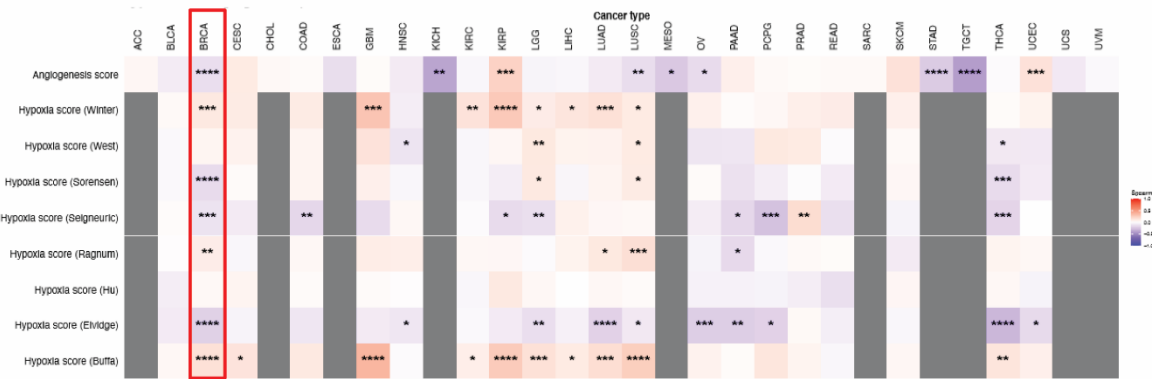
include linear regression analysis, because there is no scientific value for that for this data type. Instead, these data were binned and included in panel (h). Violin plots in panels (h), (m) and (r) show p values calculated according to Mann-Whitney U tests.

A Signal Transduction Pathways (BRF2)

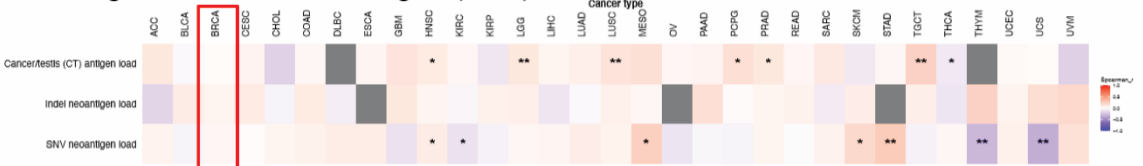
Sup 3



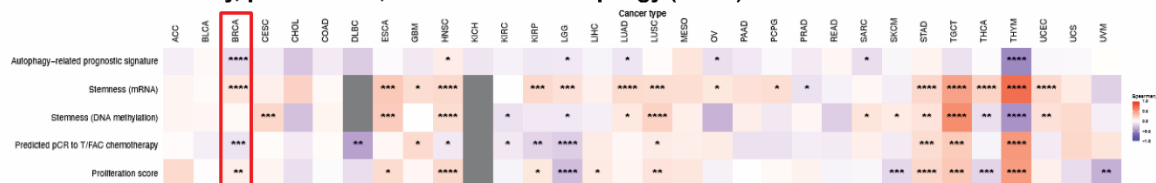
B Hypoxia and angiogenesis (BRF2)



C Neoantigens and cancer testis antigens (BRF2)

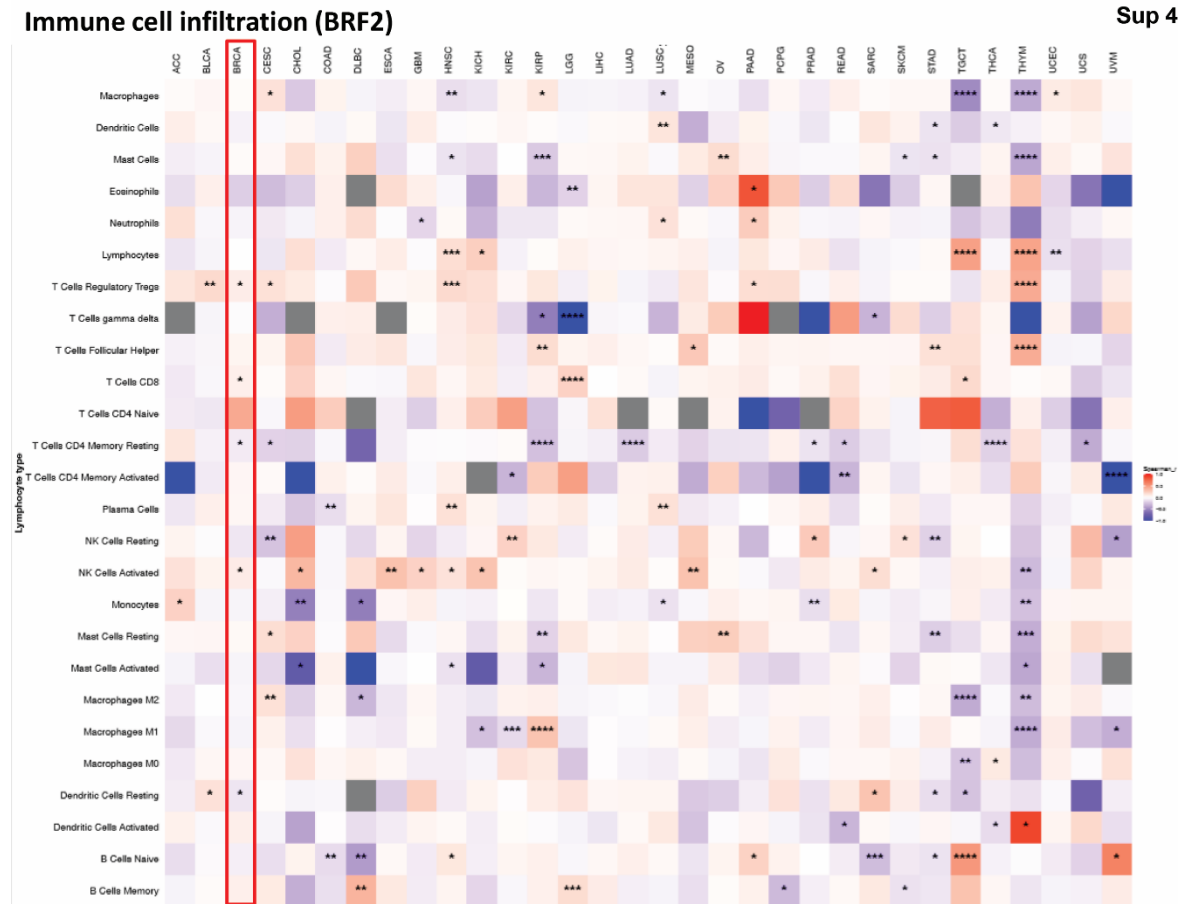


D Chemosensitivity, proliferation, stemness and autophagy (BRF2)

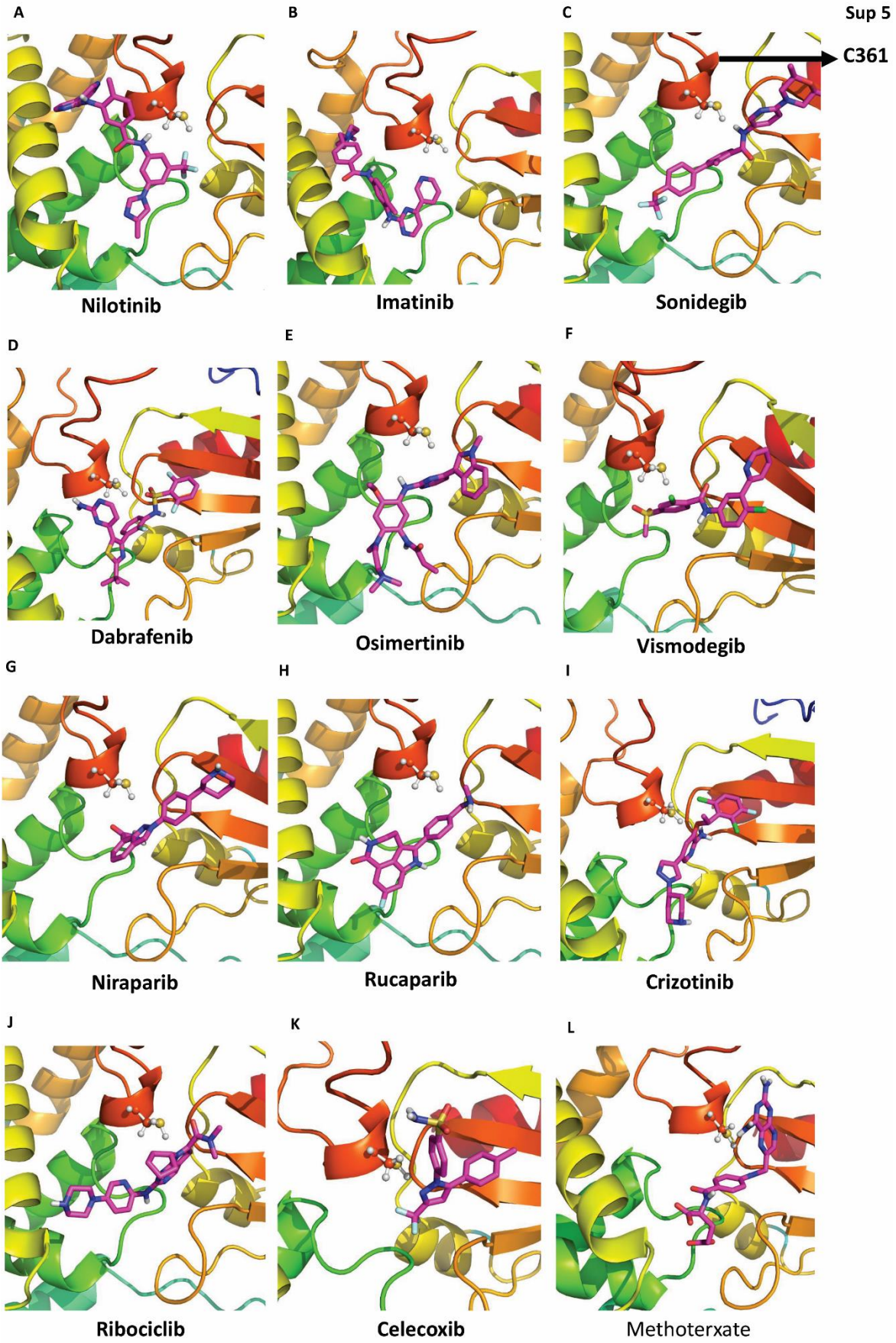


Sup

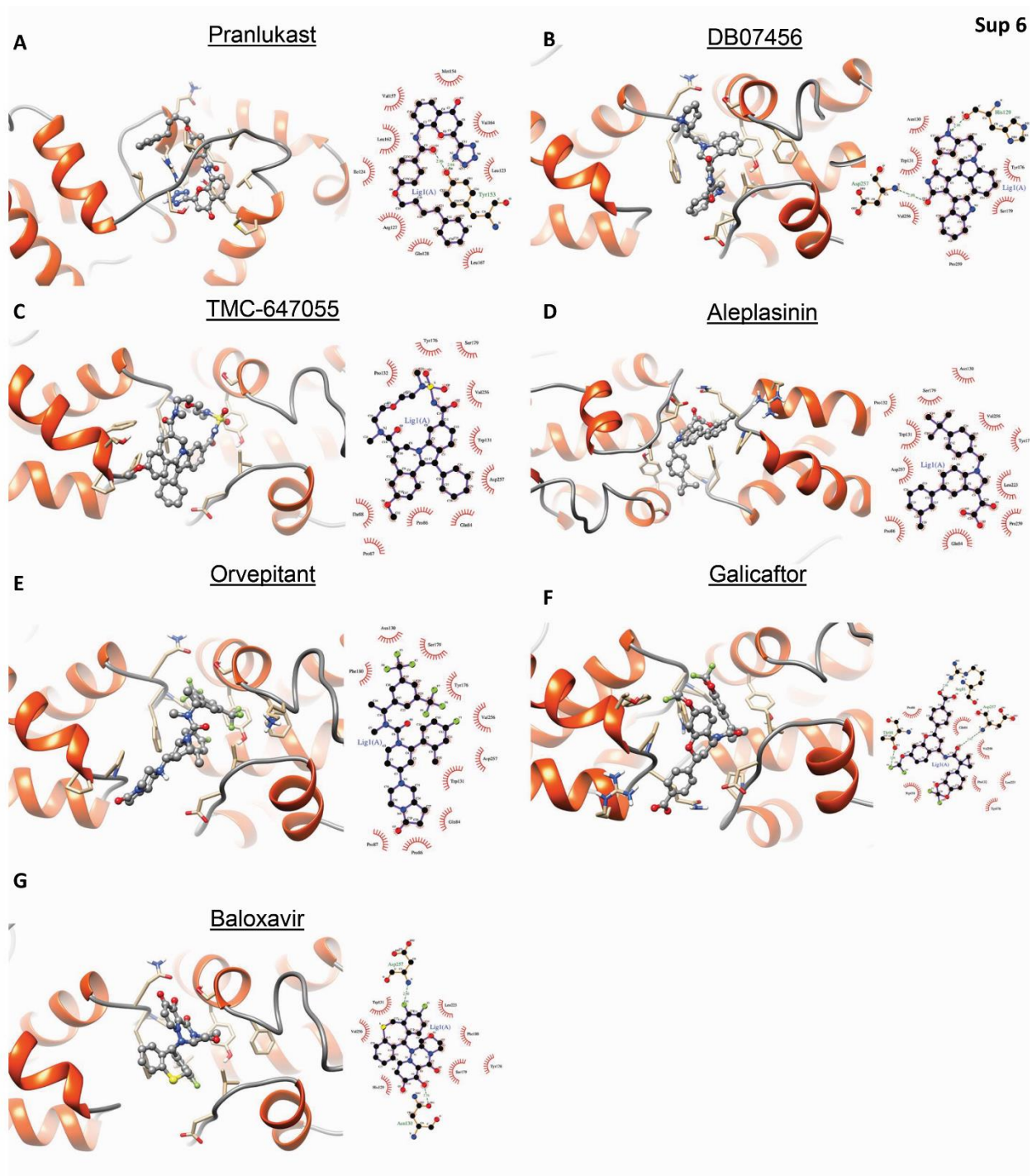
3A-D: The heatmaps show the correlations between gene expression levels and the variables listed on the y-axis in up to 32 cancer types, whose abbreviations are listed on the top x-axis. For more details see the methods. Each tile in the heatmaps shows the Spearman correlation between the gene expression levels (mRNA levels from The Cancer Genome Atlas (TCGA) RNAseq datasets)) and the variables listed below. The colour of each tile reflects the Spearman correlation coefficient (r), as indicated by the colour key on the right. Each tile shows the Spearman p value abbreviations in the following format: blank tiles: $p > 0.05$; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$.



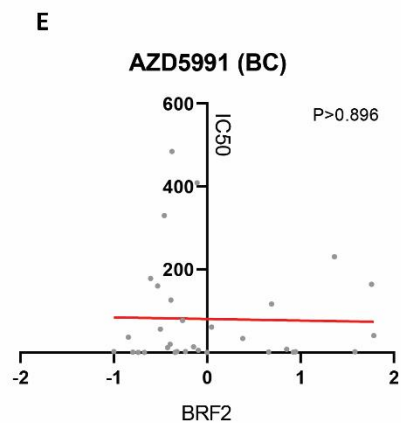
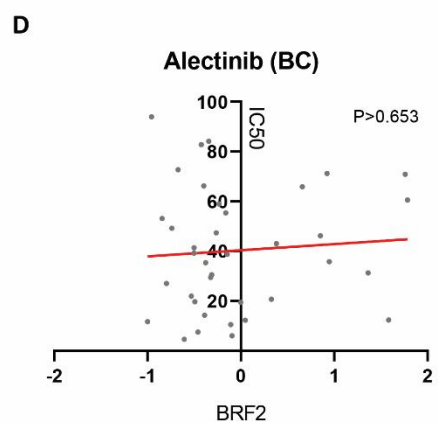
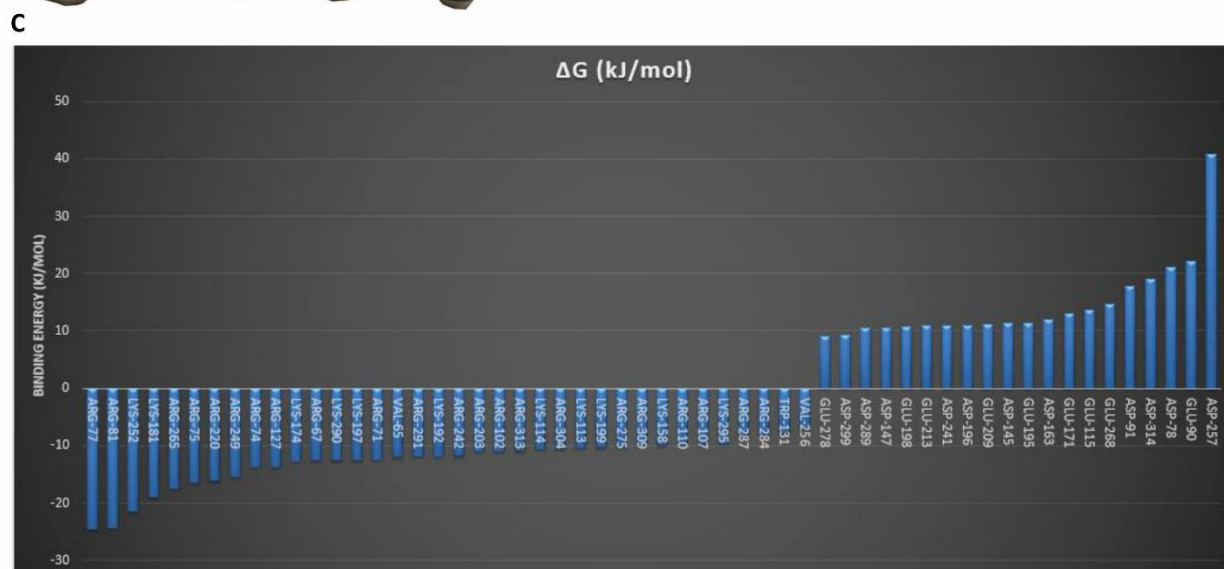
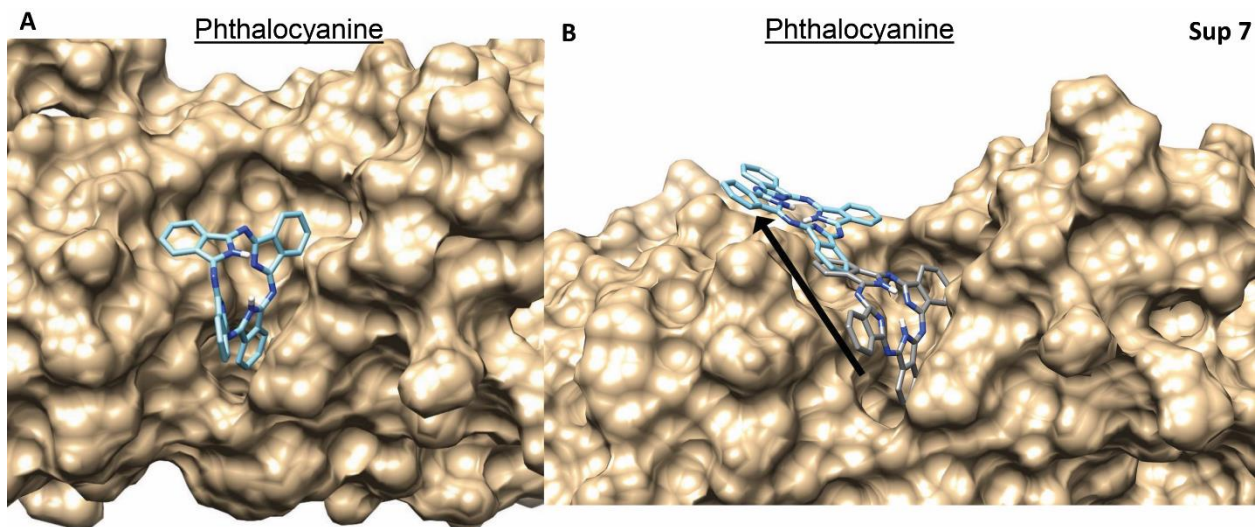
Sup 4: Immune cell infiltration: The levels of indicated tumour-infiltrated immune cell types were estimated using the tumor immune estimation resource, for more details see the methods and for statistics methodology see Sup 3.



Sup 5A-L: Docking shows architecture of binding of indicated drugs to C361 amino acid of BRF2 (black arrowed).



Sup 6A-E. Molecular docking and interaction between amino acids of BRF2 with ligands as indicated in figure (compounds listed 8th -15th in ranking).



Sup 7A-D. The surface view of docked BRF2-Phthalocyanine complex at (A) starting point (B) 8 ns after simulation. The arrow shows the direction of ligand in moving out of the cavity. (C) The

Electrostatic energy calculated for all residues. **(D-E)** Correlation between BRF2 expression in breast cancer cell lines (threshold of expression is set between -2 to 2) and IC50 of **(D)** Alectinib and **(E)** AZD5991.