Molecular Modeling and Functional Analysis of Exome Sequencing-Derived Variants of Unknown Significance Identifies a Novel Constitutively Active FGFR2 Mutant in Cholangiocarcinoma

Egan, et al

SUPPLEMENTARY MATERIALS AND METHODS

Oncology Genomic Testing

Patients underwent clinical genomic testing including next-generation sequencing panels, and/or whole exome sequencing conducted in Clinical Laboratory Improvement Amendments (CLIA) certified laboratories. These laboratories included: Foundation Medicine (Cambridge, MA), Baylor College of Medicine (Houston, TX), Caris Life Sciences (Phoenix, AZ), Genoptix (Carlsbad, CA) and Mayo Clinic (Rochester, MN). Genetic counseling was provided to patients undergoing whole exome genomic testing to discuss the benefits and risks of genetic testing.

Fresh frozen tissue specimens were collected during surgical resection or biopsy, and maintained at -80°C until nucleic acid extraction. A pathologist evaluated a portion of each specimen that was processed through standard formalin fixation and paraffin embedding, plus touch preparations for the biopsies, to confirm the presence of tumor, degree of necrosis, percent cellularity and percent of tumor nuclei. The remaining portion was processed through standard preservation methods in paraffin. Testing that only required paraffin embedded samples was performed either on newly acquired or clinically

acquired archival tissue. Solid tumors represented 57% and hematological malignancies 43% of the cancers tested (Supplementary Table S1).

Mutation Evaluation

In silico VUS Prioritization. Upon receipt of findings an analysis team consisting of bioinformaticians and cancer biologists provided additional biological context utilizing publically available databases and functional prediction algorithms (Supplementary Table S7). In order to identify VUS of potential functional significance a mutation filtering process was employed. Only single nucleotide variants reported to be VUS by the reporting laboratory were then evaluated to determine if FDA approved drugs or drugs in clinical trials were potentially available as treatment options to target these genes. Once the VUS had been reduced to all therapeutically targetable VUS, the function of each protein was annotated utilizing the Uniprot database [1]. This filtered list of genes and mutations was then subject to *in silico* functional prediction tools, Mutation Assessor [2] and Polyphen-2 [3] in order to determine the potential effect of the mutation on protein function. Mutations for which the prediction tools did not provide a prediction were evaluated manually by determining the location of the mutation in the protein [1]. Mutations predicted to be potentially damaging by at least one tool or falling in domains of known functional significance (e.g. kinase domain) were considered potentially deleterious and therefore therapeutic targets of greatest interest. Next, potentially deleterious VUS were screened for the availability of experimental structural biology data. Due to the availability of numerous therapies targeting kinases the list was further reduced to only kinases with experimental structural biology data. VUS present in kinases with this structural data were evaluated more closely to determine the location of the VUS

within the protein's functional domains. From these kinases, mutations falling within the kinase domain of five genes were selected for 3D modeling and final *in silico* functional prediction. Representative models from this filtering process are in Supplementary Figure S1-5. For *in vitro* studies, 10 VUS were selected for further study. Several of these mutations fell outside the filtering process as they were not predicted deleterious or they lacked experimental structural data, but were located in regions of functional interest within their respective proteins. Consequently, they were also selected for further *in vitro* study.

Structural modeling. FGFR2 F276C was modeled using PyMOL version 1.8.0.4. The FGFR4 R78H and FGFR2 F276C point mutations were introduced by applying the PyMOL Mutagenesis function using PDB files 1QCT and 1EV2, respectively. For each structure, favorable side chain rotamers were applied to prevent steric clash with surrounding side chains.

Cloning and Expression Constructs. FGFR2 in the vector p3XFLAG-CMV-13 (Sigma) was a gift from Dr. Moriyama (Tokyo Medical and Dental University, Tokyo, Japan). FGFR2 K41E and F276C mutations were prepared using the QuikChange II XL site-directed mutagenesis kit from Agilent Technologies (Cat. # 200522, Santa Clara, CA). FGFR4 in the vector pCMV6-Entry was purchased from OriGene Technologies, Inc (Cat. # RC204230, Rockville, MD). KDR and PDGFRA in the pDONR223 vector were purchased from Addgene (Cat. #s 23892 and 23925, Cambridge, MA). PDGFRB in the pCDNA3 vector was a gift from Dr. Carl-Henrik Heldin (Uppsala University, Sweden). KDR, PDGFRA and PDGFRB coding regions were cloned into the pCMV14-FLAG vector (Sigma Cat. # E7908)

using standard DNA recombinant protocols. Mutations of FGFR4 (R78H), KDR (G55E, G539R), PDGFRA (G251E, V484M, T632M) and PDGFRB (V258L, V316M) were prepared using the Q5 Site-Directed Mutagenesis Kit from New England Biolabs (Cat. # E0554S, Ipswich, MA). Details of cloning and mutagenesis primers are available upon request.

Cell lines, transfections and treatments. KMCH-1, KMBC and Hucct-1 cholangiocarcinoma cell lines were a gift from Dr. Gregory Gores (Mayo Clinic) and were derived as described [4-6]. Panc1 and HepG2 cell lines were obtained from the American Type Culture Collection. Cell lines were not authenticated. All cells were cultured in DMEM (Invitrogen) supplemented with 10% fetal bovine serum (FBS; SAFC BioScience). Cell lines in normal growth media were transiently transfected using X-tremeGENE HP (Sigma Cat. # 6366546001) using a ratio of 1 µg plasmid DNA/ 3 µl X-tremeGENE HP /ml of growth medium. In some experiments cells transfected the previous day were cultured at 37°C in DMEM/0.1% BSA overnight with or without 20 ng/ml FGF2, 25 ng/ml $VEGF_{165}$ (referred to throughout as VEGF), or 50 ng/ml PDGF-BB (referred to throughout as PDGF) (Cat. #s 233-B-025, 293-VE and 220-BB-010, R&D Systems, Minneapolis MN). For short-term ligand treatments, transfected cells were serum-starved overnight in DMEM/0.1% BSA. The following day, cells were treated at 37°C ± 10 ng/ml FGF2 or 50 ng/ml PDGF-BB. For experiments with the FGFR inhibitor, BGJ398 (Cat. # S2183, Selleck Chemicals, Houston TX), cells transfected with FGFR2 constructs the previous day were cultured overnight with 20 ng/ml FGF2 overnight in DMEM/0.1% BSA, and then treated for 3 h with

various concentration of BGJ398 or equivalent DMSO in DMEM/0.1% BSA with 20 ng/ml FGF2.

Western blotting. Cells in 10 cm dishes were transfected with receptor tyrosine kinases (RTKs) or control vector (pCMV) and treated with ligand as described above. After 2 days, cells were washed 3 times with ice-cold PBS and then lysed in 50 mM Tris, pH 7.5, 1% NP-40, 0.1% SDS, 300 mM NaCl, 2.5% glycerol, supplemented with 10 µg/ml leupeptin, 10 µg/ml aprotinin, 1 µg/ml pepstatin, 1 mM PMSF, Complete protease inhibitor cocktail (Roche), 10 mM NaF, 2 mM Na_3VO_4, 20 mM β-glycerophosphate, and 4 mM NaP_7O_4. Protein concentrations of lysates were determined using the BCA protein reagent (Thermo). Lysates were then diluted in Laemmli buffer with 20 mg/ml DTT and run on 7.5% Bio-Rad Min-Protean TGX or Criterion Tris-HCl precast gels. Gels were transferred to PVDF membranes in Towbin transfer buffer with 0.01% SDS. The transferred membranes were blocked in 3% BSA in PBS with 0.3% Tween-20 (PBS-T) for 1 h. Membranes were then incubated at room temperature for 2 h with primary antibodies in PBS-T with 3% BSA. The following primary antibodies were used: M2-FLAG, β-actin, α-tubulin (Cat. #s F1804, A5441, T9026, Sigma), Pan AKT, phospho-S473-AKT, ERK 1/2, phospho-ERK (phospho-T202/Y204 ERK1, phospho-T185/Y187-ERK2) (Cat. #s MAB2055, MAB887, MAB18251 and MAB15761, R&D systems), and vinculin (Cat. # 2669856, Millipore). After washing, blots were incubated for 1 h with HRP-goat anti-mouse or anti-rabbit antibody (Cat. # AP124P and AP132P, Millipore) in PBS-T with 3% BSA. After further washes, immunoreactive signals on blots were detected with SuperSignal

West Pico Chemiluminescent Substrate (Cat. # 34080, Thermo) and captured on X-Ray film. Films were converted to digital images by scanning. For quantitative analysis, immunoblots of replicate experiments each included the same set of samples (e.g., control vector, WT and mutant proteins, ± ligand). Digital images were quantified using Image J 1.49v gel analysis tools. To combine multiple replicate analyses, blot signal values for individual proteins (e.g., FGFR2, pERK) on each blot were divided by the average of all signals for that protein on the blot. The normalized values of results from several blots were then averaged and standard errors calculated. For presentation in figures, these relative values were expressed in relation to the value for the WT protein (either with or without ligand) as defined in the figure legends. Two-tailed t-tests were performed using Microsoft Excel 2011.

Immunofluorescence—Hucct-1 cells were grown on glass cover slips in 3.5 cm dishes and transfected with RTKs. After 2 days, the cells were fixed in 4% formaldehyde for 40 min and permeabilized in 0.1% Triton-X-100 at 4°C for 5 min. Cells were blocked in immunofluorescence blocking buffer (IFBB; PBS, 5% goat serum, 5% glycerol) and then incubated for 2 h at room temperature with 1:400 M2-FLAG antibody in IFBB. After washing cells were incubated with Alexafluor 598 goat anti-mouse antibody (Cat. # A-11001, Thermo) used at 1:200. After staining, cells were mounted in Prolong Gold with DAPI (Cat. # P36931, Thermo). Microscopy was performed using an Olympus AX70 equipped with a Hamamatsu C4742-95 camera. Images were captured using MetaMorph (Universal Imaging Corp) and processed using Adobe Photoshop.

REFERENCES

1 UniProt, C. UniProt: a hub for protein information. *Nucleic Acids Res* **43**, D204-212, doi:10.1093/nar/gku989 (2015).

2 Reva, B., Antipin, Y. & Sander, C. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res* **39**, e118, doi:10.1093/nar/gkr407 (2011).

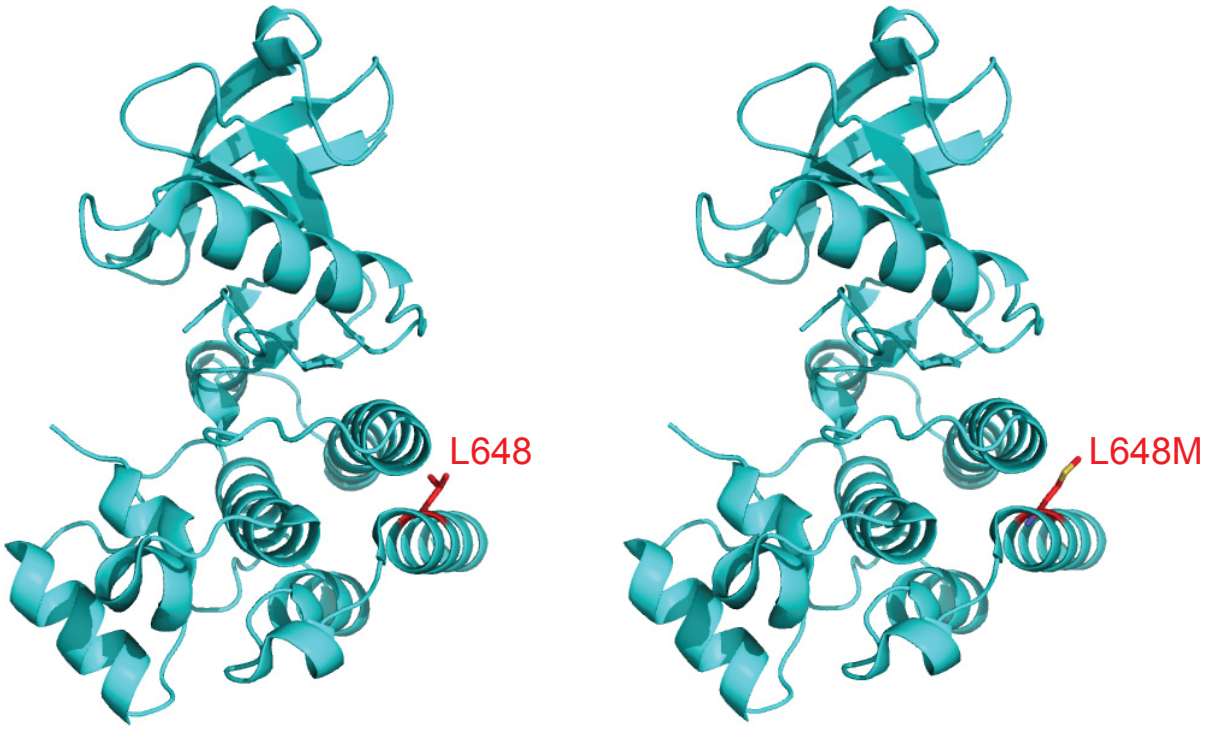
3 Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations. *Nat Methods* **7**, 248-249, doi:10.1038/nmeth0410-248 (2010).

4 Murakami, T., Yano, H., Maruiwa, M., Sugihara, S. & Kojiro, M. Establishment and characterization of a human combined hepatocholangiocarcinoma cell line and its heterologous transplantation in nude mice. *Hepatology* **7**, 551-556 (1987).

5 Yano, H., Maruiwa, M., Iemura, A., Mizoguchi, A. & Kojiro, M. Establishment and characterization of a new human extrahepatic bile duct carcinoma cell line (KMBC). *Cancer* **69**, 1664-1673 (1992).

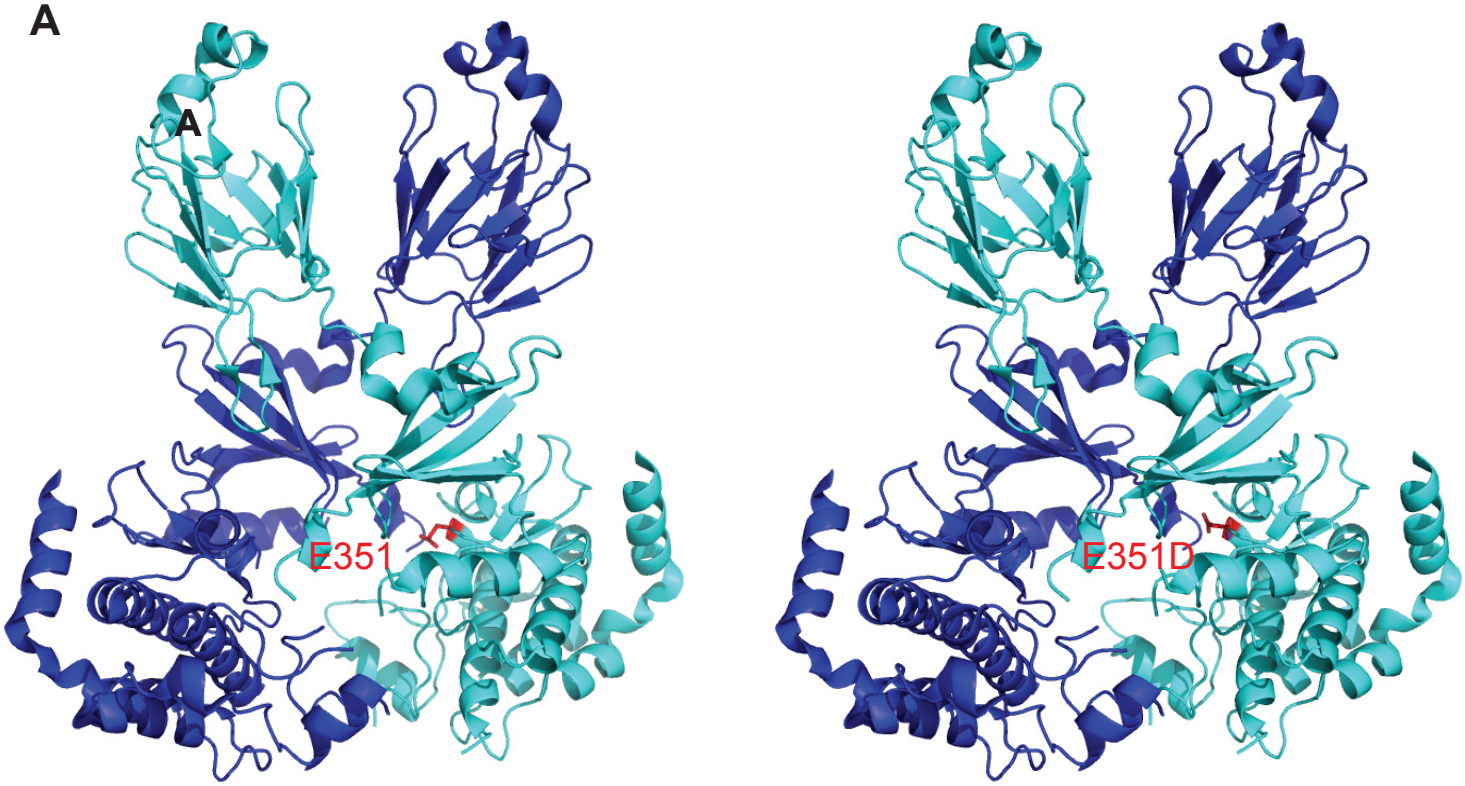
6 Miyagiwa, M., Ichida, T., Tokiwa, T., Sato, J. & Sasaki, H. A new human cholangiocellular carcinoma cell line (HuCC-T1) producing carbohydrate antigen 19/9 in serum-free medium. *In Vitro Cell Dev Biol* **25**, 503-510 (1989).

BTK

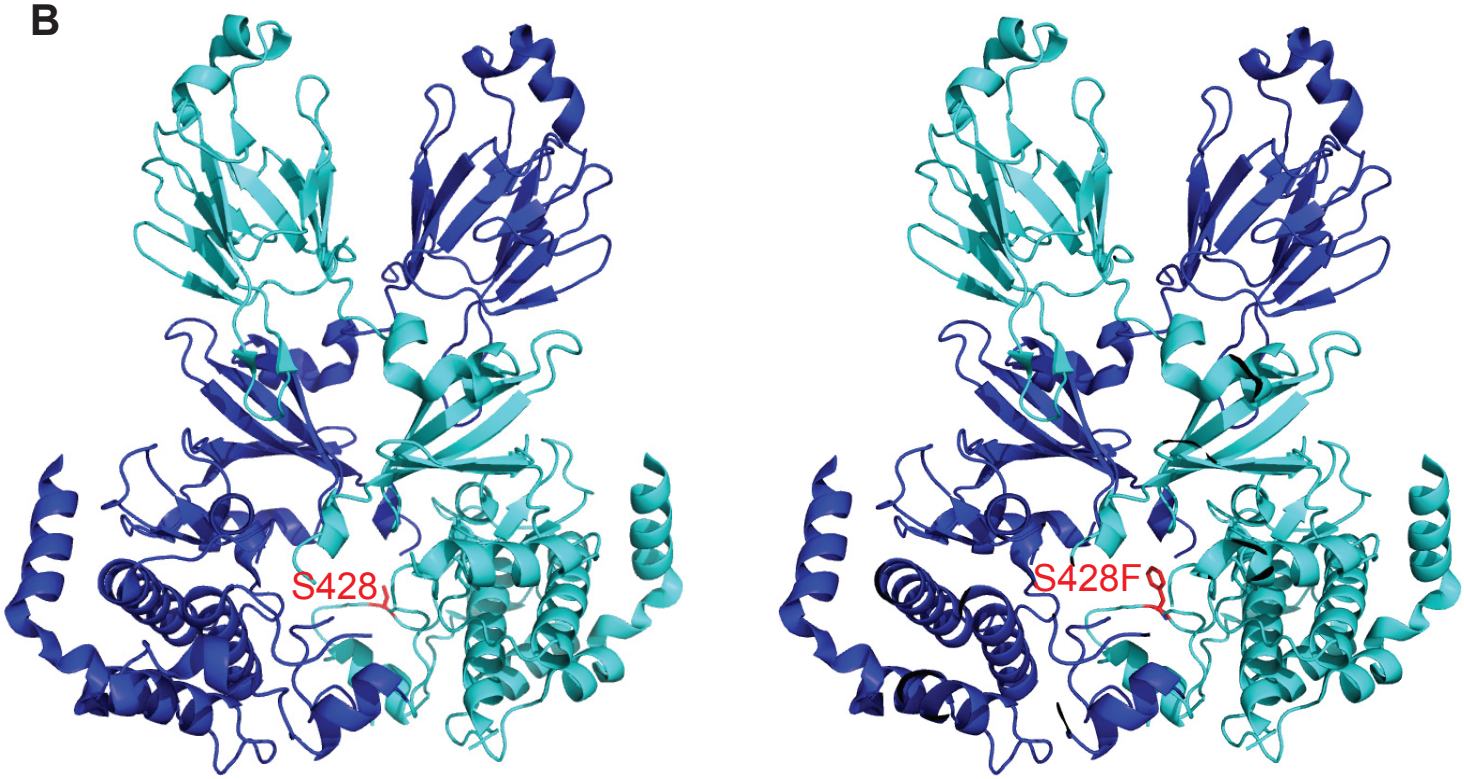


CHEK2

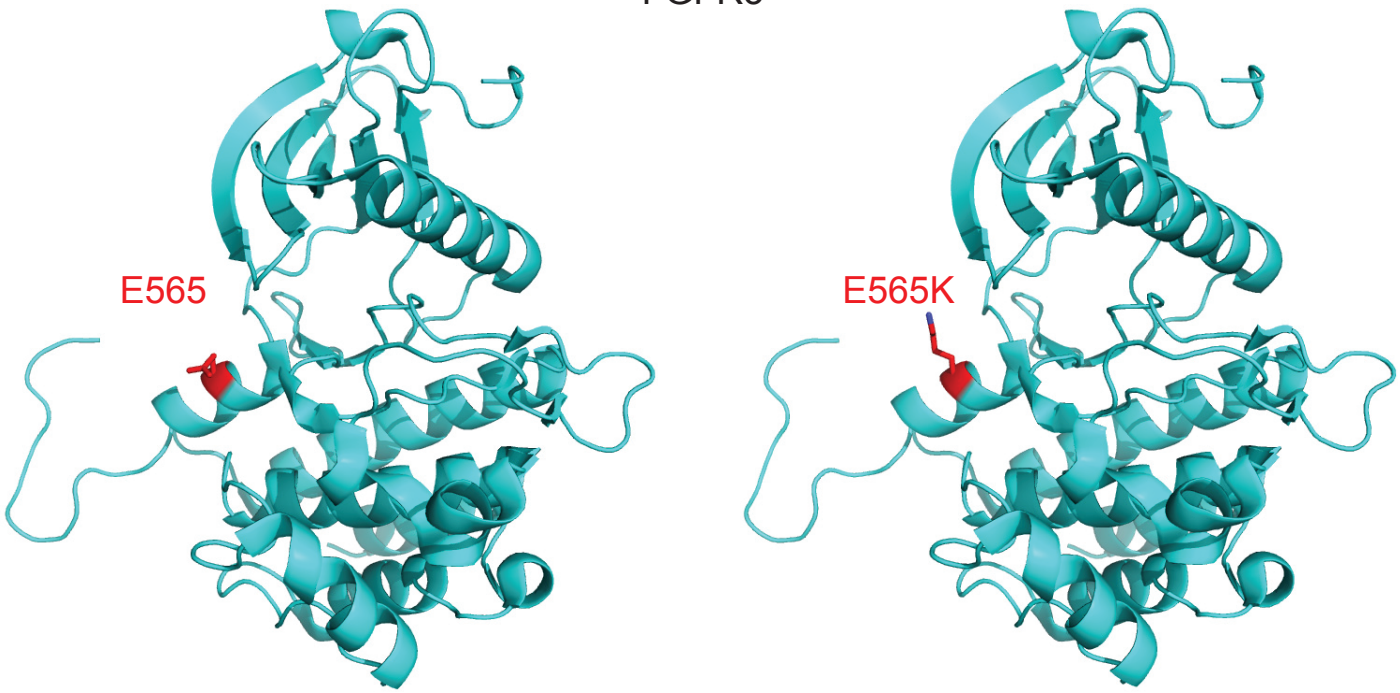
A



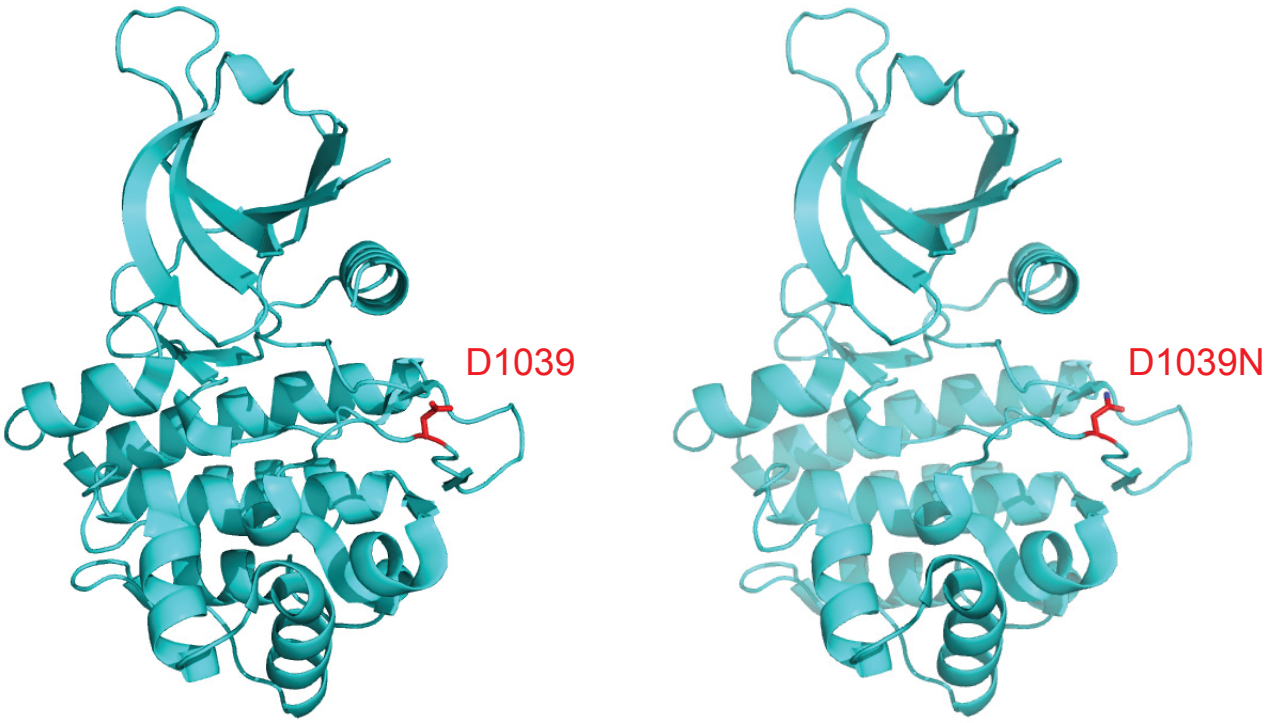
B

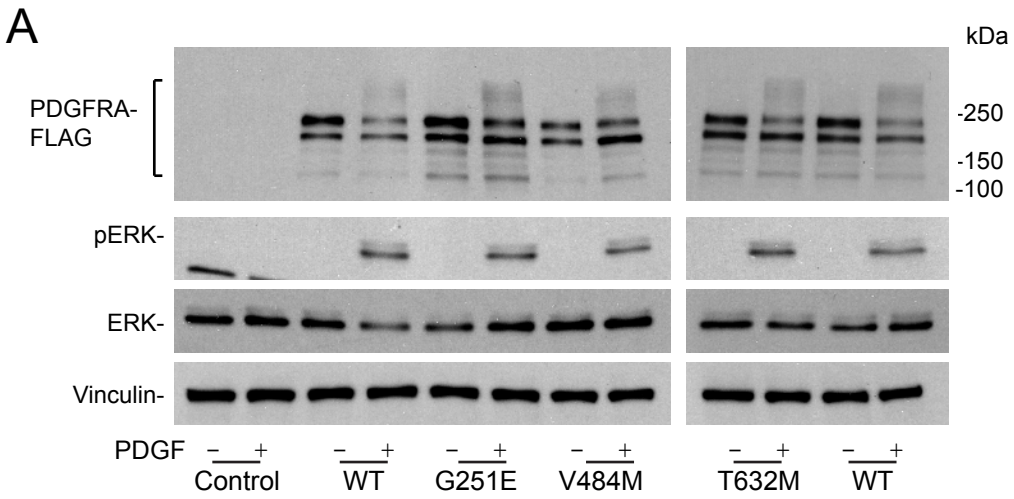


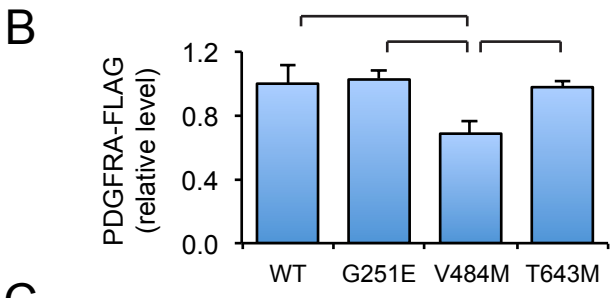
FGFR3

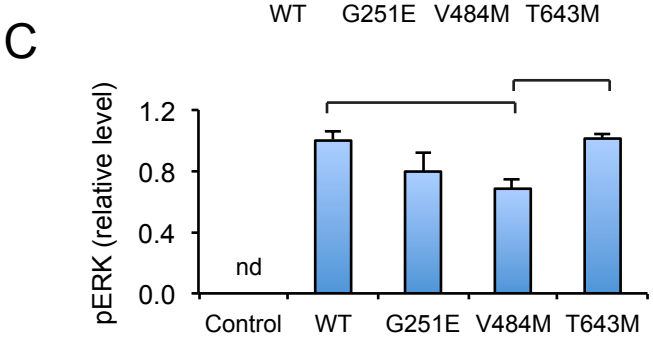


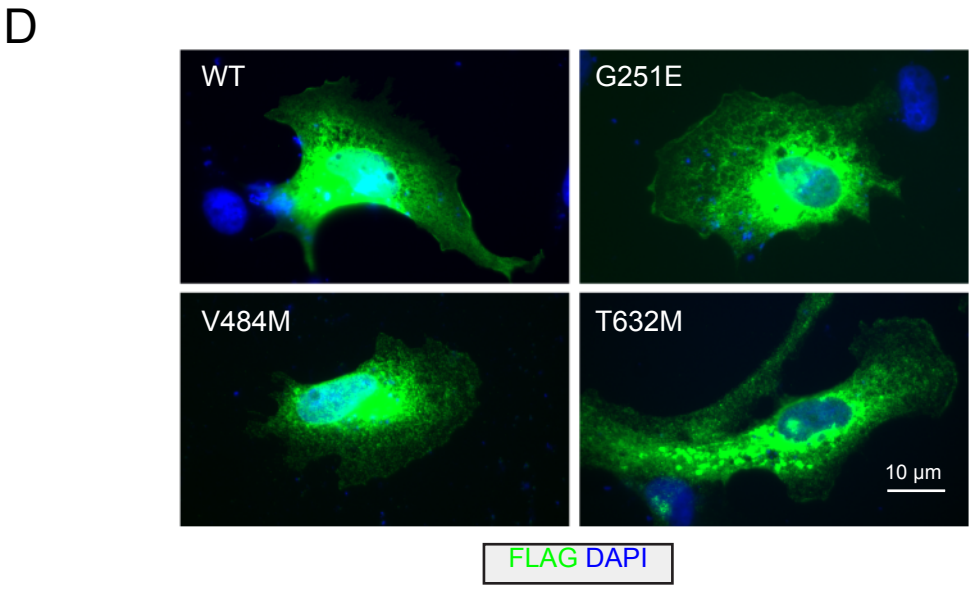
JAK1

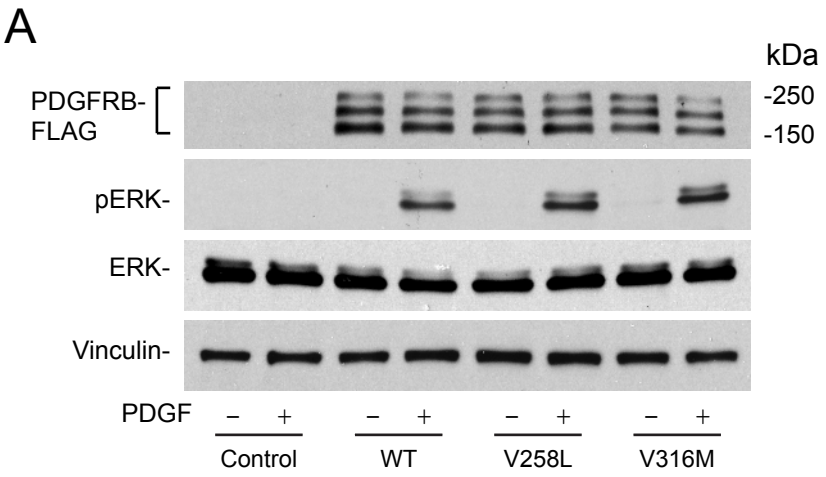


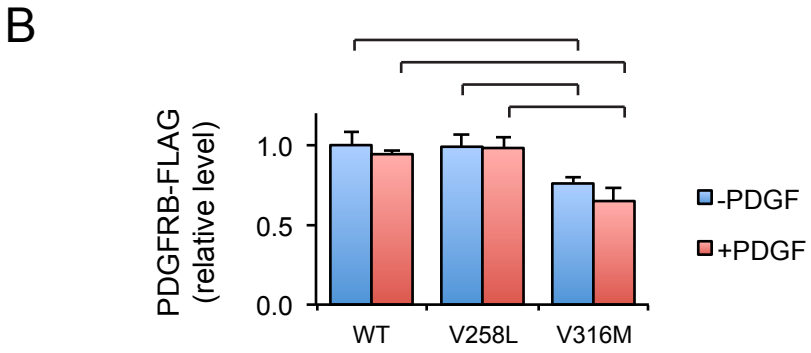

A
kDa
PDGFRA-
FLAG
-250
-150
-100
pERK-
ERK-
Vinculin-
PDGF
Control
WT
G251E
V484M
T632M
WT

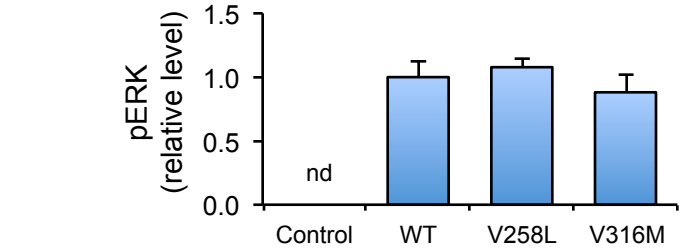

B
PDGFRA-FLAG
(relative level)
1.2
0.8
0.4
0.0
WT
G251E
V484M
T643M

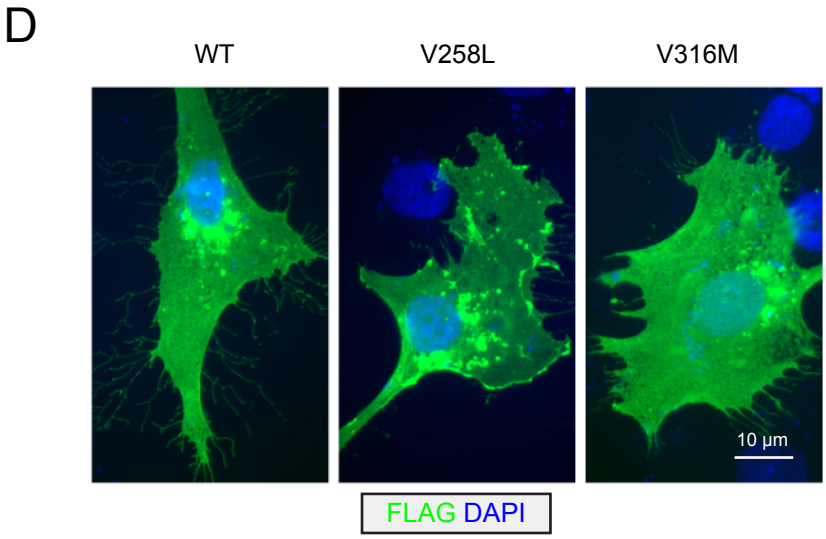

C
pERK (relative level)
1.2
0.8
0.4
0.0
nd
Control
WT
G251E
V484M
T643M


D
WT
G251E
V484M
T632M
10 μm
FLAG DAPI


A
kDa
PDGFRB-FLAG
-250
-150
pERK-
ERK-
Vinculin-
PDGF
– + – + – + – +
Control
WT
V258L
V316M


B
PDGFRB-FLAG (relative level)
1.0
0.5
0.0
WT
V258L
V316M
-PDGF
+PDGF


C
pERK (relative level)
1.5
1.0
0.5
0.0
nd
Control
WT
V258L
V316M


D
WT
V258L
V316M
10 μm
FLAG DAPI

Supplementary Table S1. Tumor demographics

Solid tumors (N=174)	**N (%)**
Hepatobiliary/pancreas	32 (18)
Gynecologic	28 (16)
Genitourinary	27 (16)
Gastrointestinal	22 (13)
Breast	18 (10)
Sarcoma	17 (10)
Brain	7 (4)
Endocrine	6 (3)
Lung/thoracic	6 (3)
Skin	5 (3)
Head/neck	4 (2)
Carcinoma unknown primary	2 (1)
Hematological maligancies (N=134)	
Lymphoma	46 (34)
Acute Leukemia	41 (31)
Myeloma	26 (19)
Chronic myeloproliferative/myelodysplastic neoplasms	21 (16)

See Data Supplement 2 (Supplemental Table 2)

See Data Supplement 3 (Supplemental Table 3)

See Data Supplement 4 (Supplemental Table 4)

Supplementary Table S5. Modeling predicted effects

Gene	Mutation	Modeling prediction
BTK	L648M	Large lobe surface mutation. May possibly serve as allosteric modulator or mediator of protein-protein interactions
CHEK2	E351D	Mutation in activation loop
CHEK2	S428F	Resides in loop preceding the aG loop of the large lobe in the kinase domain (the aEF and aG loops have different conformers between active and inactive states) [1]
FGFR3	E565K	Mutation in large lobe near ATP binding pocket
JAK1	D1039N	Mutation in activation loop
SRC	P491R	Surface of the kinase domain. May possibly serve as allosteric modulator or mediator of protein-protein interactions

[1] Cai Z, Chehab NH, Pavletich NP. Structure and activation mechanism of the CHK2 DNA damage checkpoint kinase. Mol Cell. 2009 Sep 24;35(6):818-29

Supplementary Table S6. Receptor tyrosine kinases selected for *in vitro* study

Gene	Mutation	Mutation location in protein	Cancer type	No. of occurrences[1]	Other reports[2]
FGFR2	K41E	Ig-like C2-type 1 domain	Acute Leukemia	1	
FGFR2	F276C	Extracellular, Ig-like C2-type 3 domain	Cholangiocarcinoma	1	Cholangiocarcinoma (COSM1743352)
FGFR4	R78H	Extracellular, Ig-like C2-type 1	Acute Leukemia	1	
KDR	G55E	Extracellular, Ig-like C2-type 1 domain	Gynecologic	1	G55R Melanoma (COSM540133)
KDR	G539R	Extracellular, Ig-like C2-type 5 domain	Genitourinary	1	Gastric Adenocarcinoma (COSM4125180)
PDGFRA	G251E	Extracellular, Ig-like C2-type 3	Head/neck	1	G251V Esophageal Adenocarcinoma (COSM5034387)
PDGFRA	V484M	Extracellular, Ig-like C2-type 5 domain	Gynecologic	1	Gastrointestinal stromal tumor (Clinvar variation ID 240309)
PDGFRA	T632M	Protein kinase domain	Acute Leukemia	1	T632K Small cell lung cancer (COSM5675361)
PDGFRB	V258L	Extracellular, Ig-like C2 type 3	Lung/thoracic	1	
PDGFRB	V316M	Extracellular	Acute Leukemia, Head/neck, Myeloma	3	

[1] Number of occurrences within the current data set of 4328 single nucleotide variants from 308 tumors.
[2] Reports of the same VUS or alternate substitutions at the same amino acid site from publically available data bases. Source IDs are listed in parentheses.

Supplementary Table S7. Publicly available databases utilized in mutation filtering algorithm

Database	URL	Reference
Uniprot	http://www.uniprot.org	[1]
COSMIC	http://cancer.sanger.ac.uk/cosmic	[2]
Mutation Assessor	http://mutationassessor.org/r3/	[3]
Polyphen-2	http://genetics.bwh.harvard.edu/pph2/	[4]
DGIdb	http://dgidb.genome.wustl.edu	[5]
Food and Drug Administration	http://www.accessdata.fda.gov/scripts/cder/drugsatfda/	
Clinicaltrials.gov	https://clinicaltrials.gov	

REFERENCES

1 UniProt, C. UniProt: a hub for protein information. *Nucleic Acids Res* **43**, D204-212, doi:10.1093/nar/gku989 (2015).

2 Forbes, S. A. *et al.* COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res* **43**, D805-811, doi:10.1093/nar/gku1075 (2015).

3 Reva, B., Antipin, Y. & Sander, C. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res* **39**, e118, doi:10.1093/nar/gkr407 (2011).

4 Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations. *Nat Methods* **7**, 248-249, doi:10.1038/nmeth0410-248 (2010).

5 Wagner, A. H. *et al.* DGIdb 2.0: mining clinically relevant drug-gene interactions. *Nucleic Acids Res* **44**, D1036-1044, doi:10.1093/nar/gkv1165 (2016).