RESOURCE
https://doi.org/10.1038/s41594-019-0358-z

nature structural & molecular biology

Defining the landscape of ATP-competitive inhibitor resistance residues in protein kinases

N. S. Persky[1], D. Hernandez[1,9], M. Do Carmo[1,9], L. Brenan[1,9], O. Cohen[1,2,3,4], S. Kitajima[2], U. Nayar[1,2,3,4], A. Walker[1], S. Pantel[1], Y. Lee[1], J. Cordova[5], M. Sathappa[1], C. Zhu[1], T. K. Hayes[1,2], P. Ram[1,2,3,4], P. Pancholi[2], T. S. Mikkelsen[1,6,7], D. A. Barbie[2], X. Yang[1], R. Haq[2], F. Piccioni[1], D. E. Root[1] and C. M. Johannessen[1,8*]

Kinases are involved in disease development and modulation of their activity can be therapeutically beneficial. Drug-resistant mutant kinases are valuable tools in drug discovery efforts, but the prediction of mutants across the kinome is challenging. Here, we generate deep mutational scanning data to identify mutant mammalian kinases that drive resistance to clinically relevant inhibitors. We aggregate these data with subsaturation mutagenesis data and use it to develop, test and validate a framework to prospectively identify residues that mediate kinase activity and drug resistance across the kinome. We validate predicted resistance mutations in CDK4, CDK6, ERK2, EGFR and HER2. Capitalizing on a highly predictable residue, we generate resistance mutations in TBK1, CSNK2A1 and BRAF. Unexpectedly, we uncover a potentially generalizable activation site that mediates drug resistance and confirm its impact in BRAF, EGFR, HER2 and MEK1. We anticipate that the identification of these residues will enable the broad interrogation of the kinome and its inhibitors.

Protein kinases are effective therapeutic targets in cancer and other diseases. Knowledge of generalizable drug-resistant 'Gatekeeper' mutations in kinases has helped confirm the on-target nature of small molecule inhibitors, aided development of more efficacious drugs[1–7] and deepened our understanding of kinase substrates and signaling networks[8–10]. As a tool, however, the utility of the Gatekeeper is limited; many kinases diverge at this position, making it difficult to predict drug-resistance mutations that preserve kinase function[6].

To enable studies of drug-resistant mutants for all disease-related kinases it would, therefore, be useful to identify additional residue locations where a mutation could confer resistance to ATP-competitive inhibitors of many members of the kinome. However, discovering residues that confer drug resistance broadly across kinases has historically been difficult and time-consuming—it took over a decade to identify the Gatekeeper as a common site of ATP-competitive inhibitor resistance in kinases[11–19]. Scaling this discovery process across the kinome is a challenge, especially when clinical samples are difficult to obtain, as is the case for many drug-resistant tumors. Repair-deficient bacteria, error-prone PCR and other subsaturation mutagenesis techniques have helped identify drug-resistant mutants in individual kinases[20–30] by prospectively testing the drug sensitivity of hundreds of mutations in parallel, but these approaches are limited by the lack of mutational diversity inherent to the techniques. Recent advances in DNA synthesis-based mutagenesis have made it possible to create mutant libraries containing all possible amino acid substitutions, allowing for the comprehensive mapping of coding variants[31–41]. Such saturating mutagenesis approaches are often inclusive of mutations unlikely to occur in nature, such as those caused by substitutions requiring di- or tri-nucleotide changes, providing an opportunity for thorough mapping of structure/function relationships that may not be otherwise detectable.

Here, we used the structural conservation of protein kinase domains, the similar mechanism of action for ATP-competitive kinase inhibitors and a systematic comparison of mutagenesis assays to prospectively identify common sites of drug resistance in kinases. Within this analytical framework, we identify and validate multiple sites that drive drug resistance when mutated in clinically relevant kinases. Unexpectedly, we uncovered one position in which mutation leads to activation, and subsequently drug resistance, in protein kinases. The end result of this effort is a toolkit that may enable the scientific community to identify sites that when mutated impact drug resistance and/or basal activity in any kinase.

Results

The landscape of drug-resistant CDK4/6 mutants. We generated complementary DNA libraries that contain 99.98% of all possible single amino acid variants of human CDK6 and CDK4 using mutagenesis by integrated tiles (MITE)[32,34,38,41]. Mutant libraries were transduced into the Meljuso melanoma cell line (Fig. 1a), verified to have even representation of all mutants (Extended Data Fig. 1a–e) and treated with the ATP-competitive CDK4/6 inhibitor palbociclib (Fig. 1a and Extended Data Fig. 1a,f,g). Mutants in surviving cells were quantified via massively parallel sequencing and enrichment or depletion quantified by calculating the *z*-score normalized $\log_2$(fold change) (LFC) relative to pretreatment reference abundance (Methods and Extended Data Fig. 1a). We identified 177 CDK6 mutants in 66 amino acids and 103 CDK4 mutants

[1]The Broad Institute of MIT and Harvard, Cambridge, MA, USA. [2]Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA. [3]Center for Cancer Precision Medicine, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA. [4]Harvard Medical School, Boston, MA, USA. [5]University of New Mexico, Albuquerque, NM, USA. [6]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA, USA. [7]Present address: 10x Genomics, Inc., Pleasanton, CA, USA. [8]Present address: Oncology, Novartis Institutes for BioMedical Research, Cambridge, MA, USA. [9]These authors contributed equally: D. Hernandez, M. Do Carmo, L. Brenan. *e-mail: johannes@broadinstitute.org

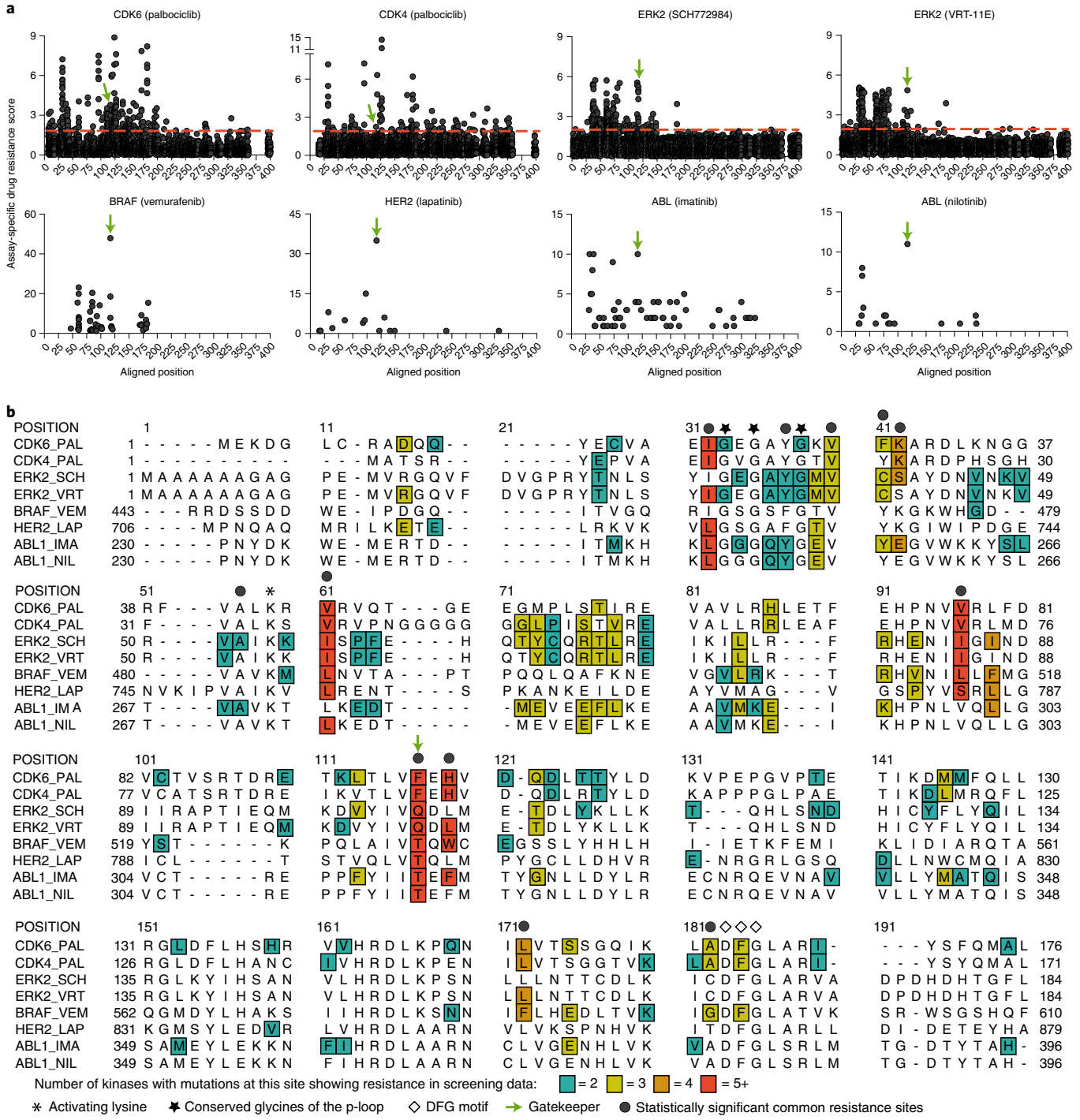


Fig. 1 | Hotspot analysis of saturating and subsaturating mutagenesis screens identifies common sites that mediate ATP-competitive kinase inhibitor resistance. a, Scatter plots showing the average assay-specific resistance score achieved for each substitution (*y* axis, black dots, an average of $n = 3$ replicate measurements for each of 19 substitutions for CDK4, CDK6 and ERK2) aligned to a common kinase domain (*x* axis) for each kinase and drug. For all other kinases, only scoring substitutions are shown as determined in the original datasets. Red lines indicate a *z*-score of 1.95, the threshold of significance for resistance phenotypes in saturation screens (see Methods for details). **b**, Identification of common drug-resistant sites visualized on the protein alignment of kinases included in this analysis. Box colors indicate the number of kinases in which site mutation mediated drug resistance. Positions 200–419 (not shown) contained two positions (at 261 and 271) that shared resistance in two kinases (CDK4 and ABL1).

in 56 amino acids above a threshold *z*-score of >1.95 ($\approx P < 0.05$), including Gatekeeper mutants (CDK6$^{F98Q/E}$ and CDK4^{F93E}) (Fig. 1a, Extended Data Fig. 2a–d and Supplementary Tables 1 and 2) and 10 (of 15) residues that physically contact palbociclib (Extended Data Fig. 3a–c).

'Hotspot' analysis reveals generalizable ATP-competitive kinase inhibitor resistance residues. We employed 'hotspot' analysis, reminiscent of approaches used in genomic characterization efforts, to identify residues that when mutated cause drug resistance across kinases. To do this, we consolidated deep mutational scanning

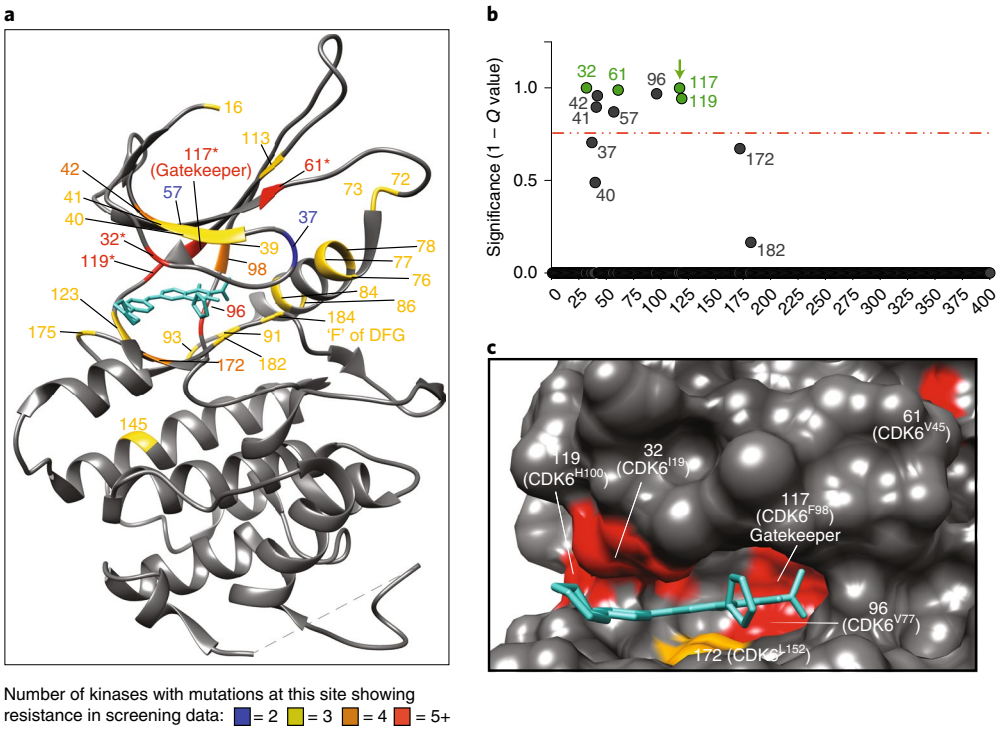


Fig. 2 | Key common drug-resistant positions mapped onto a ribbon structure of CDK6. a, The structure of CDK6 bound to palbociclib (PDB 2EUF) highlighting the 29 residues that cause resistance when mutated in three kinases (27 residues) or achieved statistical significance in only two (two residues) kinases. Residues are identified using the common numbering scheme and colored by the number of kinases with resistant mutations at this position in our analysis. **b**, Assessment of the statistical significance of common drug-resistant positions using a robust-rank aggregation that incorporates recurrence of observation and strength of resistance phenotypes across the kinases tested. The *x* axis shows the aligned kinase position and the *y* axis shows the strength of significance as indicated by 1 – *Q* value. The significance of the observed recurrence was computed as compared to a null model of uncorrelated ranking. The resulting *P* values of all paralogous positions were corrected for multiple hypotheses testing using the Benjamini–Hochberg method of FDR. Green circles indicate the positions that were prioritized and validated in this study, gray circles represent all other residues with non-zero 1 – *Q* values. **c**, Positions of ATP-competitive-inhibitor resistance sites common to four or more kinases mapped onto the structure of CDK6 bound to palbociclib (turquoise) (PDB 2EUF) and colored as indicated in **a**.

drug-resistance data from human CDK4, CDK6 and ERK2 (ref. [32]) with data from subsaturation drug-resistance mutagenesis screens in four human kinases, including BRAF (vemurafenib), HER2 (lapatinib) and BCR-ABL (imatinib and nilotinib)[20,27,28,42] (Fig. 1a).

We next aligned the kinase domains of these six proteins using an iterative process of automated two-dimensional (2D) alignments, three-dimensional (3D) alignments and manual spot-checking of well-known and structurally conserved residues including the activating lysine, Gatekeeper residue, DFG motif, and p-loop glycines (Fig. 1b). To facilitate cross-kinase comparisons of individual residues, a common kinase domain numbering system was used. Residues containing drug-resistant mutants were mapped onto this alignment, revealing the Gatekeeper (position 117) and 28 additional positions that, when mutated, yielded resistance phenotypes in at least half of the six kinases screened (Fig. 1b) and were distributed across the N-lobe of a representative kinase, CDK6 (Fig. 2a). Notably, mutations at position 61 (for example, CDK6^{V45}) were as prevalent across kinases as the Gatekeeper mutation, showing a drug-resistance phenotype in every protein for which data were available, despite being outside the drug-binding pocket (Fig. 2a).

To evaluate the statistical significance of drug-resistance hotspots, we used a robust-rank aggregation[43] that incorporated recurrence of observation and strength of resistance phenotypes across kinases. Eight resistance positions had a false discovery rate (FDR) of 15%, including position 32, 117 (Gatekeeper), 61, 96, 41, 42, 119 and position 57 (Fig. 2b). We focused our characterization efforts on positions 32, 61 and 119, all of which showed drug-resistance phenotypes in at least five of the six kinases analyzed and were found in spatially distinct locations (Figs. 1b and 2a,c).

Validation of resistance phenotypes for common CDK6, CDK4 and ERK2 mutants. To experimentally validate our hotspot analysis, we generated CDK6 mutants at positions 32 (CDK6^{I19W}), 61 (CDK6^{V45M}), 117 (CDK6^{F98E}) and 119 (CDK6^{H100F}) containing the strongest substitutions at each site observed in primary screening data. Mutant CDK6 were transduced into Meljuso cells and population doublings were calculated with and without palbociclib. Relative to wild-type (WT) or kinase-dead CDK6^{K43M}, we observed resistance phenotypes with CDK6^{I19W} (position 32) and CDK6^{F98E} (position 117) comparable to cells with single guide RNA (sgRNA) mediated knockout of *RB1*, our positive control for this assay (7.5 doublings > WT), whereas CDK6^{H100F} (position 119) and CDK6^{V45M} (position 61) provided a more modest resistance phenotype (4.4 and 3.4 doublings > WT, respectively) (Fig. 3a). Similar patterns of resistance were observed for a subset of CDK6 mutants in the presence of equipotent concentrations of two additional ATP-competitive CDK6 inhibitors, abemaciclib and ribociclib (Extended Data Fig. 4a). We further observed resistance phenotypes consistent with the pooled primary screening data for a range of CDK6 mutants identified in the CDK6 screen, but for which the analogous mutants did not exhibit broad resistance in the screening data of other kinases (Fig. 3a and Extended Data Fig. 4b). For all resistant CDK6 mutants, expression at least partially restored phosphorylation of Rb in the presence of palbociclib (Fig. 3b).

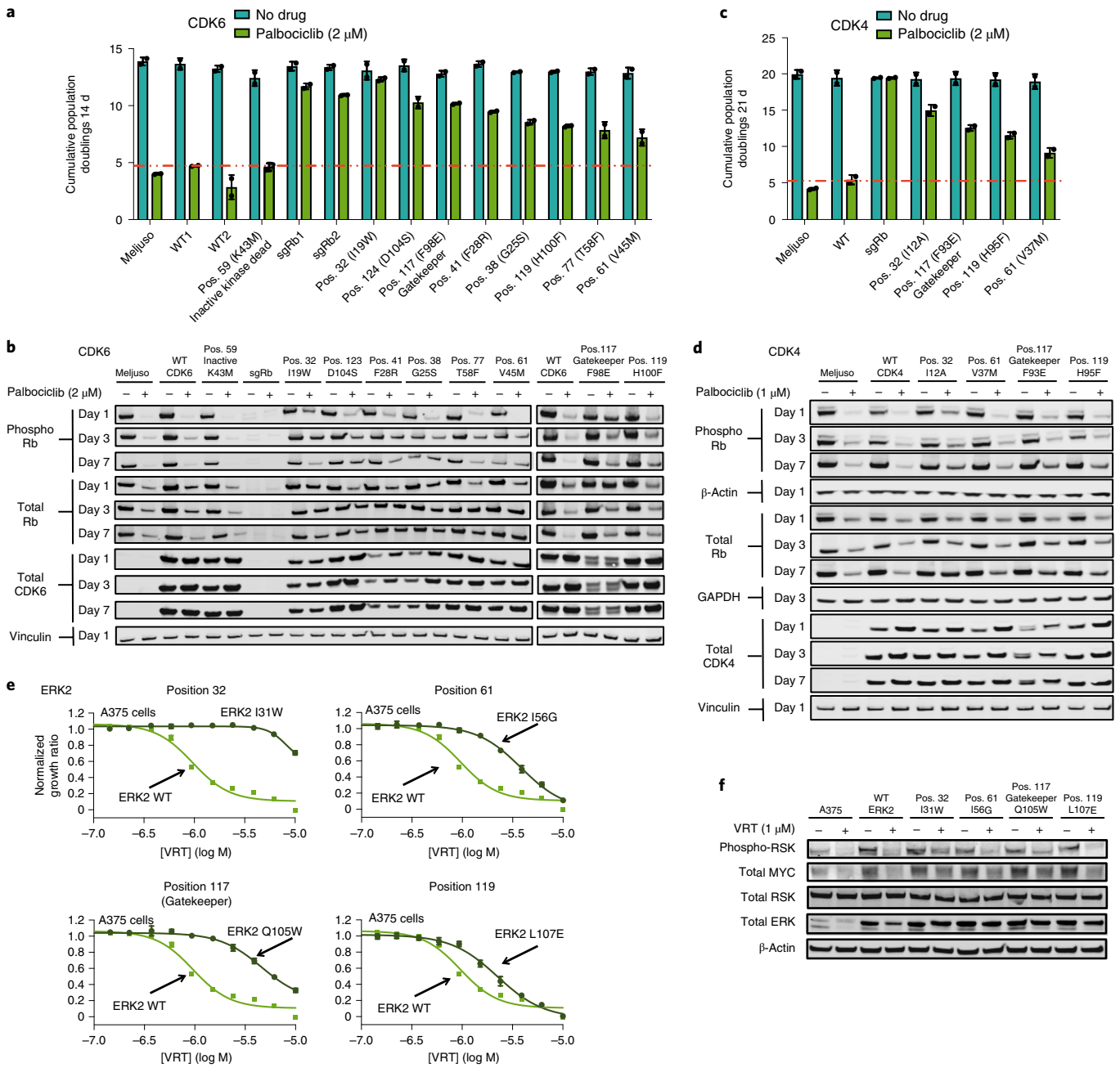


Fig. 3 | Mutation of common drug-resistant positions in CDK6, CDK4 and ERK2 causes drug resistance. a, Average cumulative population doublings of Meljuso cells transduced with WT or mutant CDK6 in the presence (green) or absence (turquoise) of 2 μM palbociclib ($n=2$ independent experiments), error bars represent s.d. Pos., position. **b**, Western blot assessment of CDK6-mediated signaling in Meljuso cells, a representative example of two experiments. **c**, Average cumulative population doublings of Meljuso cells transduced with WT or mutant CDK4, as in **a** ($n=2$ independent experiments), error bars represent s.d. **d**, Western blot assessment of CDK4-mediated signaling in Meljuso cells, a representative example of two experiments. **e**, Proliferation assay measuring ERK inhibitor (VRT, VRT-11E) sensitivity of A375 cells transduced with WT or mutant ERK2 ($n=3$ replicate measurements for each mutant at each concentration). **f**, Western blot assessment of ERK2-mediated signaling in the presence or absence of an ERK inhibitor, a representative example of two experiments. The uncropped westerns blots for **b**, **d** and **f** are shown in the source data. Statistical source data for **a**, **c** and **e** are available online.

We next tested mutations of these four residues that also scored in CDK4, using the maximally resistant substitutions seen in primary CDK4 screening for each of position 32 ($CDK4^{I12A}$), 117 ($CDK4^{F93E}$), 61 ($CDK4^{V37M}$) and 119 ($CDK4^{H95F}$). We found phenotypes of these CDK4 mutants, such as CDK6, were consistent with our pooled primary screening data (Fig. 3c and Extended Data Fig. 4c) and partially restored Rb phosphorylation in the presence of palbociclib (Fig. 3d).

We broadened our validation efforts by introducing mutations in ERK2 at positions 32 ($ERK2^{I31W}$), 61 ($ERK2^{I56G}$), 117 (Gatekeeper, $ERK2^{Q105W}$) and 119 ($ERK2^{L107E}$) using maximally resistant substitutions observed at each site in the ERK2 screening data. We transduced these mutant ERK2 into $BRAF^{V600E}$-mutant A375 cells, induced expression with doxycycline[32] and measured proliferative responses to VRT-11E, an ATP-competitive ERK1/2 inhibitor[44]. Expression of $ERK2^{I56G}$ (position 61) generated resistance

to VRT-11E to a degree comparable to the Gatekeeper mutant ERK2^{Q105W} (position 117) (four- and fivefold shift in half-maximum inhibitory concentration (IC_{50}), respectively), whereas ERK2^{I31W} (position 32) produced a more robust resistance phenotype, resulting in an eightfold shift in IC_{50} (Fig. 3e). Resistance produced by ERK2^{L107E} (position 119) was more modest (twofold IC_{50} shift). Expression of drug-resistant mutants of ERK2 was sufficient to restore phosphorylated RSK and total MYC levels, markers of ERK activity (Fig. 3f). These data confirm the drug-resistance phenotypes observed in the primary screening data for CDK6, CDK4 and ERK2, supporting the idea that mutations at these sites may be generalizable across kinases.

Common drug-resistance mutations differentially impact drug binding. To probe the mechanism of resistance for position-32 (CDK6^{I19W}), -61 (CDK6^{V45M}), -117 (CDK6^{F98E}) and -119 (CDK6^{H100F}) mutants, we indirectly assessed drug binding using a cellular thermal shift assay (CETSA)[45]. Consistent with the impaired drug binding of Gatekeeper mutations in other kinases[46], the stability of the CDK6 Gatekeeper mutant (CDK6^{F98E}) was unaffected by palbociclib (Fig. 4a,b) and was comparable to that of the position-32 mutant (CDK6^{I19W}). In contrast, CDK6^{V45M} and CDK6^{H100F} both showed thermal stability profiles similar to that of WT CDK6, indicating drug-binding proficiency (Fig. 4a,b). Given the qualitative nature of the CETSA assay it is formally possible that CDK6^{V45M} and CDK6^{H100F} may decrease the affinity for palbociclib but in amounts below the detection limit of the assay.

To substantiate CETSA results more directly, we affinity-purified Flag-tagged CDK6 mutants from transfected 293T cells and performed differential scanning fluorescence (DSF) in vitro in the presence and absence of palbociclib (Extended Data Fig. 5a,b). The results obtained from this orthogonal approach were consistent with what we observed in lysates using CETSA (Fig. 4b), suggesting that CDK6^{V45M} (position 61) and CDK6^{H100F} (position 119) are drug-binding proficient, whereas CDK6^{F98E} (position 117) and CDK6^{I19W} (position 32) are drug-binding deficient (Fig. 4c).

To test if drug-binding proficient, but drug-resistant, CDK6 mutants (CDK6^{V45M} and CDK6^{H100F}) might have greater affinity for drug, serving as sponges for palbociclib, reducing drug availability and limiting potency, we introduced a kinase-dead point mutant into the activating lysine (CDK6^{K43M}) of WT CDK6, CDK6^{F98E}, CDK6^{I19W}, CDK6^{H100F} and CDK6^{V45M}. Kinase-dead CDK6 (CDK6^{K43}) showed stabilization equivalent to WT CDK6 (melting temperatures of 56.7 ± 0.2 versus 55.3 ± 0.6, respectively) in CETSA, suggesting that drug binding is preserved in this mutant (Extended Data Fig. 5c) and allowing an assessment of the drug independent of kinase activity via population doubling assays (Fig. 4d). The introduction of CDK6^{K43M} was sufficient to abrogate resistance for all common resistance sites tested, implying kinase activity is required for drug resistance and that drug binding alone is not sufficient to cause resistance phenotypes (Fig. 4d). Overall, these data suggest mutation of position 32 mediates resistance by impairing drug binding, whereas mutations at positions 61 and 119 cause resistance through alternative mechanisms.

Common resistance positions validate in additional kinases and are found in drug-resistant patient tumors. Having validated resistance phenotypes for positions 32, 61 and 119, in CDK4, CDK6 and ERK2, respectively, we asked if alterations at these positions have been reported in drug-resistant human tumors. Although genomic profiling of drug-resistant tumors for many kinase inhibitors remains rare (as is the case for ERK2 and CDK4/6), multiple generations of EGFR inhibitors, including erlotinib and osimertinib, have clinically targeted activating mutants of EGFR (for example, EGFR^{L858R}) in non-small-cell lung cancer. Mutations in the Gatekeeper residue of EGFR (EGFR^{T790M}) are known to occur in about 50% of patients who are erlotinib resistant[5]. We additionally found literature reports of mutations in EGFR at position 32 (EGFR^{L718}), 61 (EGFR^{L747}) and 119 (EGFR^{T792}) in erlotinib-resistant tumors, albeit at lower frequencies than observed for Gatekeeper mutations[5,47–55]. Mutations in EGFR position 32 (EGFR^{L718}) and 119 (EGFR^{T792}) have also been reported in relapsed samples of patients treated with osimertinib[56,57].

To further confirm these observations, we tested the resistance phenotypes of EGFR mutants at positions 32 (EGFR^{L718W}), 61 (EGFR^{L747P}), 117 (Gatekeeper, EGFR^{T790M}) and 119 (EGFR^{L792E}). EGFR mutants were expressed in the EGFR mutant cell line PC9 exposed to erlotinib, an ATP-competitive EGFR inhibitor. EGFR^{L718W} (position 32) and the Gatekeeper mutant EGFR^{T790M} (position 117) produced more than a tenfold shift in the IC_{50} for erlotinib, whereas EGFR^{L747P} (position 61) and EGFR^{T792E} (position 119) shifted the erlotinib IC_{50} by two- to threefold (Fig. 5a). These mutants were additionally evaluated for resistance to the covalent ATP-competitive EGFR inhibitor osimertinib that retains activity against the EGFR Gatekeeper (EGFR^{T790M}). Osimertinib resistance was observed for all mutations except the Gatekeeper, with EGFR^{T792E} (position 119) producing an enhanced osimertinib resistance phenotype relative to erlotinib (Fig. 5b). To model second-site resistance mutations found in EGFR inhibitor resistant tumors, we introduced these mutants into a background of EGFR^{L858R}. Mutations of EGFR at these positions in *cis* with EGFR^{L858R} were sufficient to cause drug resistance to erlotinib and osimertinib (Fig. 5c,d). In both *cis* and *trans*, resistance for each mutant was associated with sustained expression of phosphorylated EGFR, a marker of protein activity, even in cases where lower protein levels of EGFR were observed, as was the case for EGFR^{L747P} (Fig. 5e,f).

Beyond EGFR, we found reports of mutations in each of these three positions in clinical samples of patients who relapsed while being treated with inhibitors of BCR-ABL (L248V[58], L273M[59] and F317I/L[60]), EML4-ALK (L1122V[61], L1198F[61] and L1152R/P[62]) and for two positions (32 and 61) in HER2 (L726I[63] and L755S[64], respectively). We independently validated a subset of these alterations as causal of resistance to ceritinib when introduced into EML4-ALK (Fig. 5g) and to lapatinib when introduced into HER2, despite lower expression levels compared to WT HER2 (Fig. 5h,i and Extended Data Fig. 6a,b). BRAF-directed inhibitors are effective in malignant melanoma harboring activating mutations in the oncogenic kinase BRAF (BRAF^{V600E})[32]. While second-site resistance mutations in BRAF are rare, we introduced a mutation at position 61 (BRAF^{L485S}) that has been reported in drug-naive tumor samples[65] and a BRAF subsaturation mutagenesis screen[28] and confirmed that this mutant caused resistance to the BRAF inhibitor dabrafenib (Fig. 5j and Extended Data Fig. 6c). These data support the notion that resistance residues identified using this approach can be extended to other kinases and are clinically relevant.

Resistance mutations at position 32 are highly predictable. A main goal of our approach was to facilitate the prediction of drug-resistance mutations across the kinome. Analysis of CDK6, CDK4 and ERK2 saturating mutagenesis data revealed a clear trait of position 32: resistance could be caused by an unusually high number of substitutions (Table 1). Position 32 is also conserved in the kinome, with the vast majority of kinases endogenously containing an I, L or V (Extended Data Fig. 6d), in contrast to resistance-causing substitutions at the Gatekeeper residue, which vary dramatically across kinases and drugs[66] (Extended Data Fig. 6e), our data suggested that position 32 could be uniquely harnessed as a site where resistance mutants could be predicted.

We therefore asked if our analysis would enable the prediction of drug-resistant mutants in kinases with no known resistance alleles. We first focused on TBK1, a kinase that serves a key role in regulating expression of inflammatory cytokine expression. MRT67307

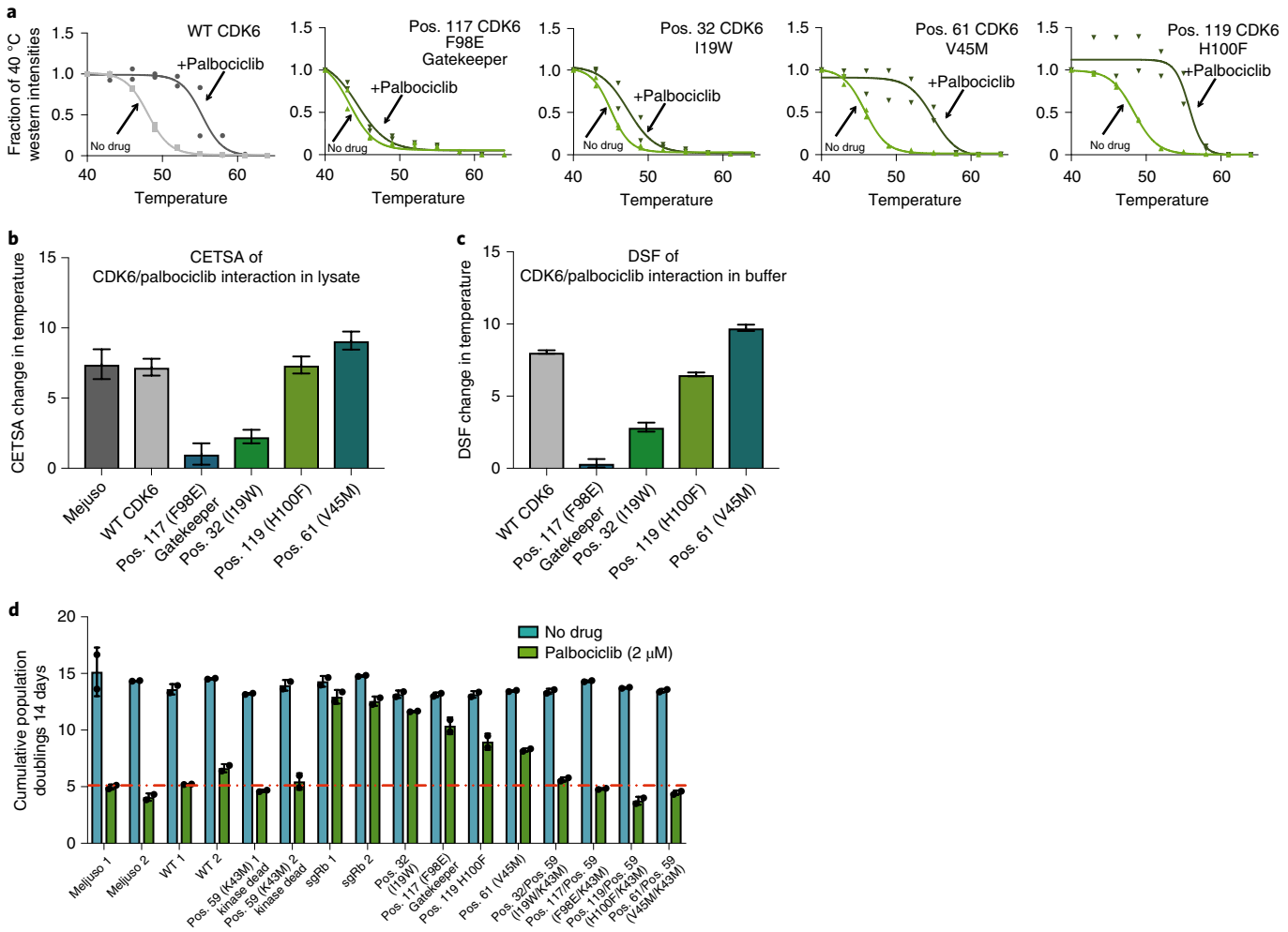


Fig. 4 | Mutation of common drug-resistant sites differentially impacts drug binding. a, CETSA was performed on Meljuso cells transduced with WT or mutant CDK6. Protein melting curves were calculated in the absence or presence of palbociclib, relative to protein abundance at 40 °C ($n = 2$ replicate measurements). **b**, Summary of changes in melting temperature of CDK6 mutants profiled in **a** using CETSA, calculated as the mean difference in melting temperature between the absence and the presence of palbociclib ($n = 2$ replicate measurements). Error bars represent propagated s.e.m. associated with curve fit. **c**, Summary of changes in melting temperature of CDK6 mutants profiled using DSF, calculated as the mean difference between the absence and the presence of palbociclib for $n = 5$ replicate measurements. Error bars represent propagated s.e.m. associated with curve fit. **d**, Average cumulative population doublings of Meljuso cells transduced with WT or mutant CDK6 ($n = 2$ independent experiments), error bars represent s.d. Statistical source data are available online.

(MRT) and momelotinib (MMB) target TBK1 (refs. [67,68]), however drug-resistant TBK1 mutants have not been described. To assess the impact of position-32 mutants on TBK1 drug sensitivity, we used a TBK1-dependent cytokine release assay that provides a surrogate for TBK1 activity. In this assay, introduction of the maximally effective substitution observed for CDK6 (W) at position 32 in TBK1 ($TBK1^{L15W}$), but not WT TBK1, restored CCL5 and CXCL10 protein levels in the presence of MRT (5 μM) or MMB (5 μM) to untreated levels (Fig. 6a,b). Similar results were achieved when using phosphorylated IRF3, a TBK1 substrate, as a readout for TBK1 activity (Fig. 6c).

We next asked if drug-resistant position-32 mutants in CK2A1 could be predicted. As with TBK1, the choice of the CK2A1 substitution was based on the maximally effective substitution observed for CDK6 (W) at position 32. WT or the CK2A1 position-32 mutant ($CK2A1^{L45W}$) were transduced into two melanoma cell lines (SKMEL5 and WM266.4) and viability assessed in the presence of the CK2 inhibitor silmitasertib (CX4945) (Fig. 6d,e and Extended Data Fig. 6f–h). As with the Gatekeeper mutation ($CK2A1^{F113A}$), cell lines expressing $CK2A1^{L45W}$ were profoundly resistant to silmitasertib, even at concentrations that eliminated WT CK2A1-expressing cells, and were sufficient to restore downstream CK2A1 signaling (Fig. 6d,e and Extended Data Fig. 6f–h).

Although subsaturating mutagenesis libraries of BRAF have been screened against ATP-competitive BRAF inhibitors[28], position-32 mutants were not detected, nor to our knowledge have they been described elsewhere. To measure the effect of BRAF position-32 mutants on drug sensitivity, we transduced A375, a $BRAF^{V600E}$-mutant melanoma cell line that is sensitive to the small molecule RAF-inhibitor dabrafenib, with WT or a position-32 mutant of BRAF and assessed their impact on drug sensitivity (Fig. 6f). Expression of $BRAF^{I463W}$ led to a tenfold shift in the dabrafenib IC_{50} relative to WT BRAF. Notably, the $BRAF^{I463W}$ mutation did not impact the activity of $BRAF^{V600E}$ when made in *cis*, as has been reported for the BRAF Gatekeeper mutation[6] (Fig. 6f). In these assays, RAF-inhibitor resistance was correlated with restoration of downstream MEK phosphorylation (Fig. 6g). Together these data suggest that mutations at position 32 may be a uniquely predictable means of causing

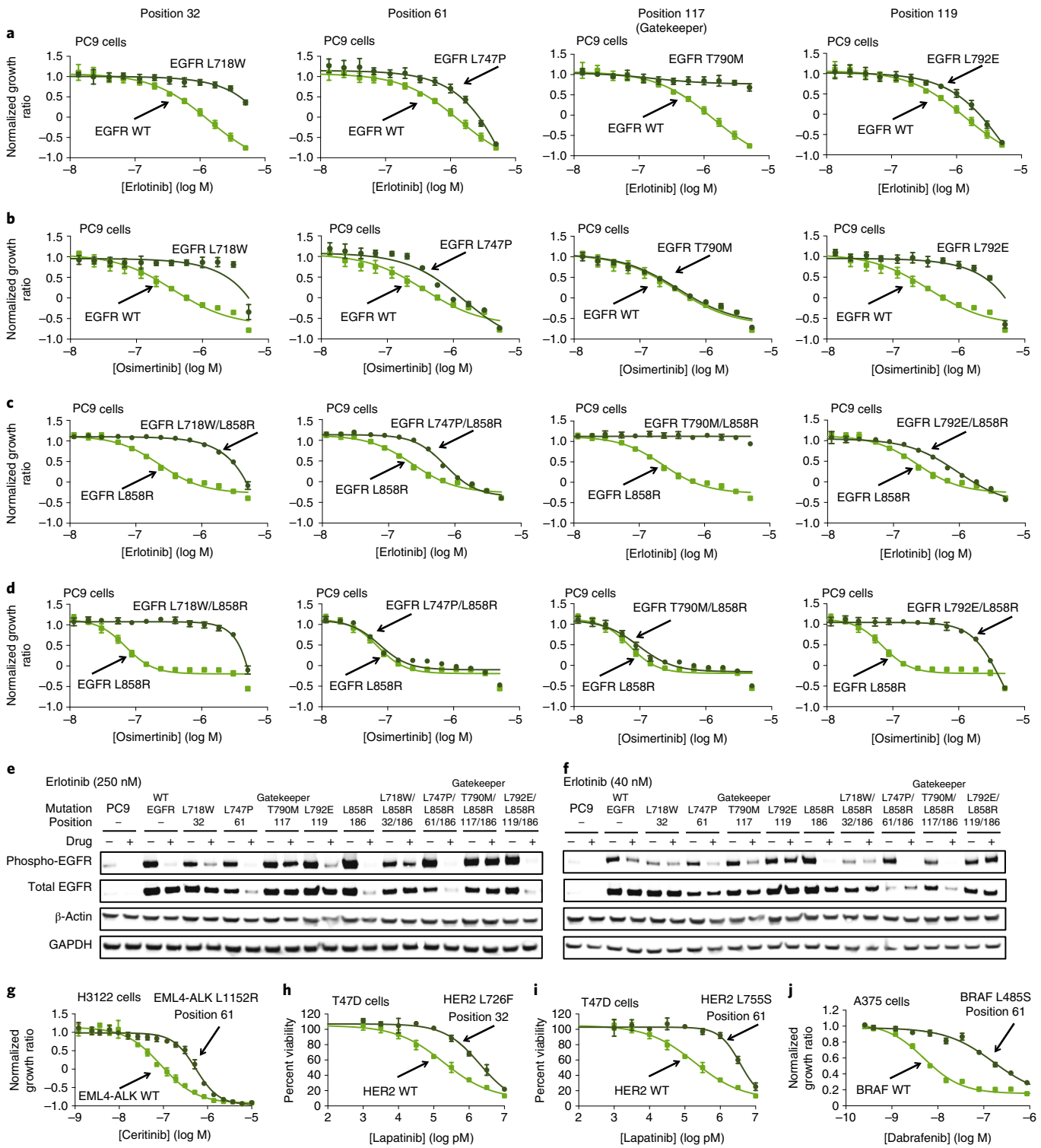


Fig. 5 | Validation of common drug-resistant positions in EGFR, EML4-ALK, HER2 and BRAF that are associated with drug resistance in patient samples. a,**b**, Proliferation assay measuring EGFR inhibitor sensitivity (erlotinib (**a**) or osimertinib (**b**)) in PC9 cells transduced with WT or single mutants of EGFR ($n = 7$ replicate measurements for each mutant at each concentration). **c**,**d**, Proliferation assay measuring EGFR inhibitor sensitivity to erlotinib (**c**) or osimertinib (**d**) in PC9 cells transduced with $EGFR^{L858R}$ single or compound (*cis*) mutants of EGFR ($n = 7$ replicate measurements for each mutant at each concentration). **e**, Western blot assessment of EGFR-mediated signaling in PC9 cells treated with erlotinib, corresponding to mutants used in **c**, representative example of two experiments. **f**, Western blot assessment of EGFR-mediated signaling in PC9 cells treated with osimertinib, corresponding to mutants used in **d**, representative example of two experiments. **g**, Proliferation assays measuring ALK inhibitor sensitivity (ceritinib) in H3122 cells transduced with WT or mutant EML4-ALK. **h**,**i**, HER2-inhibitor sensitivity (lapatinib) in T47D cells transduced with WT or mutant HER2. **j**, RAF-inhibitor sensitivity (dabrafenib) in A375 cells transduced with WT or mutant BRAF ($n = 4$ replicate measurements for each mutant at each concentration for panels **g** and **j**, $n = 3$ replicate measurements for each mutant at each concentration for panels **h** and **i**). The uncropped western blots for **e** and **f** are shown in the source data. Statistical source data for **a**–**d** and **g**–**j** are available online.

Table 1 | Summary of parameters for the prediction of functional resistance mutations in mammalian kinases

Common residue no.	Kinases hit	Starts as	Changes to	Hit kinase identity	Statistically significant	In this study
32 (Pocket Protector)	5	I	W,Q,A,S,K,E,R,H,T,M,N,G,F,C,D,Y,V,L	CDK6, CDK4, ERK2	Yes	Novel resistance mutations predicted and verified in TBK1 and CSNK2 (I,L > W). Resistance mutations validated in CDK6, CDK4, ERK2, EGFR, BRAF and HER2
		L	F,R,V	HER2, ABL1		
61 (Keymaster)	6	V	M,R,Q,Y,N,W	CDK6,CDK4	Yes	Resistance mutations validated in CDK6, CDK4, ERK2, EGFR, EML-ALK HER2, MEK1 and BRAF
		I	G,P,S,Q,D,Y,K,W,E,H,A,F,V,M,R,N	ERK2		
		L	S,F	HER2, ABL1, BRAF		
117 (Gatekeeper)	6	F	E,Q	CDK6, CDK4	Yes	As a positive control resistance mutations were validated in CDK6, CDK4, ERK2 and EGFR
		Q	W,Y,N,V,M,F,T,E,V,I	ERK2		
		T	M,L,F,I,S,G	BRAF, HER2, ABL1		
119	5	H	F,W,M,Q,Y,K	CDK6, CDK4,	Yes	Resistance mutations validated in CDK6, CDK4, ERK2 and EGFR
		L	E	ERK2		
		W	C,F,I,R	BRAF		
		F	L	ABL1		
41	3	F	R,T,K,A,Q,V,C,G	CDK6	Yes	Resistance mutation validated in CDK6
		C	P,E,I	ERK2		
		Y	C	ABL1		
42	4	K	L,F	CDK6, CDK4	Yes	–
		S	M,N	ERK2		
		E	D	ABL1		
96	5	V	L,R,Q,M,Y,E,H,N,	CDK6, CDK4	Yes	–
		I	R,C,W	ERK2		
		L	E	BRAF		
		S	P	HER2		
16	3	D	F	CDK6	–	–
		R	P	ERK2		
		E	K	HER2		
39	3	M	P,R,Q	ERK2	–	–
		T	I	HER2		
		E	K	ABL1		
40	3	V	A,D,T,F	CDK6, CDK4, ERK2	–	–
72	3	G	S	CDK4	–	–
		T	R,H,K,D,Q	ERK2		
		M	L	ABL1		
73	3	L	R,K	CDK4	–	–
		Y	V,E,G,Q,R,S,I,C,M,T,K,A,D,L,P,H,N	ERK2		
		E	K	ABL1		
76	3	S	T	CDK4	–	–
		R	I,V,H,G	ERK2		
		E	K,D	ABL1		
77	3	T	F,Y,D,M,H,Q,N,I,L,P,A,E,S,R	CDK6, ERK2	–	Resistance mutation validated in CDK6
		F	L	ABL1		
78	3	V	T	CDK4	–	–
		L	V,F	ERK2, ABL1		
84	3	L	W,R,H,P,E,Y,G,S	ERK2, BRAF	–	–
		M	L,T	ABL1		

Continued

Table 1 | Summary of parameters for the prediction of functional resistance mutations in mammalian kinases (continued)

Common residue no.	Kinases hit	Starts as	Changes to	Hit kinase identity	Statistically significant	In this study
86	3	H	F	CDK6	–	–
		R	Y	CDK4		
		E	Q,K	ABL1		
91	3	R	V,L,M,A,G	ERK2, BRAF	–	–
		K	R	ABL1		
93	3	E	N	ERK2	–	–
		V	M	BRAF		
		P	L	HER2		
98	4	I	W	ERK2	–	–
		F	G	BRAF		
		L	F	HER2, ABL1		
113	3	L	V,I	CDK6	–	–
		V	M,L	ERK2		
		F	L,V	ABL1		
123	3	Q	V	CDK6	–	Resistance mutation validated in CDK6
		T	P	ERK2		
		G	W	ABL1		
145	3	M	Y,T	CDK6, ABL1	–	–
		L	Y,F	CDK4		
172	4	L	F,Y,M,G,Q,H,N	CDK6, CDK4, ERK	–	–
		F	M,L,Y	BRAF		
175	3	S	I,R	CDK6	–	–
		E	G,K	BRAF, ABL1		
182 (DFG-1)	3	A	N,H,Q,I,T,M,P,Y,D	CDK6, CDK4	–	–
		G	I	BRAF		
184 ('F' of DFG)	3	F	C,K,M	CDK6, CDK4, BRAF	–	–
37	2	Y	R,L,V,A,H,M,T,E,G,S,D,I,N,Q,C,F	ERK2, ABL1	Yes	–
57 (activating lysine-2)	2	A	G,V	ERK2, ABL1	Yes	–

resistance to ATP-competitive inhibitors (Table 1). Given its location at the mouth of the drug-binding pocket (Fig. 2c) and that mutations at this site appear to limit inhibitor access to the drug-binding pocket, we propose calling this residue the 'Pocket Protector'.

Position 61 is a potentially generalizable activation site that mediates drug resistance. A number of clinical and experimental observations led us to hypothesize that position-61 mutants possess elevated kinase activity and require higher drug concentrations for complete inhibition. First, EGFR position-61 mutations ($EGFR^{L747P/S}$) have been reported as oncogenic in samples from patients with non-small-cell lung carcinoma[69–71]. Second, HER2 position-61 mutants ($HER2^{L755S/P/M/A/W}$) are recurrent in breast cancer, and these alterations are activating and transforming[27,72–74]. Third, BRAF mutated at position 61 ($BRAF^{L485S}$) has been reported as a rare oncogenic driver in lung cancer[65]. Moreover, in EGFR, HER2, MEK1 and BRAF, position 61 marks the start of a series of indels with gain-of-function activities[75–77]. Fourth, among the positions identified in our analysis, position 61 was uniquely identified in the kinome as significantly mutated, indicative of driver activity at this position[78]. Fifth, crystal structures of cyclin-dependent kinases (and others), have revealed an auto-inhibitory mechanism mediated by interactions between the activation loop and the hydrophobic pocket surrounding position 61 (refs. [79–81]). Sixth, a position-61 CDK6 mutant ($CDK6^{V45M}$) was enriched in the control-treatment arm of our mutagenesis screen, potentially indicative of increased activity and proliferation. This notion is consistent with the observation that, in this assay, other plausibly gain-of-function mutants predicted to affect binding to the cyclin-dependent kinase inhibitor p16 are enriched (Extended Data Fig. 7a,b). Lastly, position-61 mutants in SRC ($SRC^{L300R/N/P/K}$) have been identified as sufficient to confer gain-of-function activities in an unbiased screen[31]. With the exception of $HER2^{L755S}$—an appreciated mutational hotspot in breast cancer—position-61 alterations are rare in individual genes, but are more common when aggregated across the kinome[78].

To test the hypothesis that position-61 mutants are activated, we first confirmed the activity of tumor-associated position-61 mutants of EGFR and HER2 in a drug complementation assay (Fig. 7a). Specifically, we transduced the EML4-ALK mutant, ALK-dependent, cell line H3122 with WT EGFR, the EGFR position-61 mutant $EGFR^{L747P}$ or the constitutively-active $EGFR^{L858R}$ mutant. Here, EGFR activity is sufficient to rescue the antiproliferative

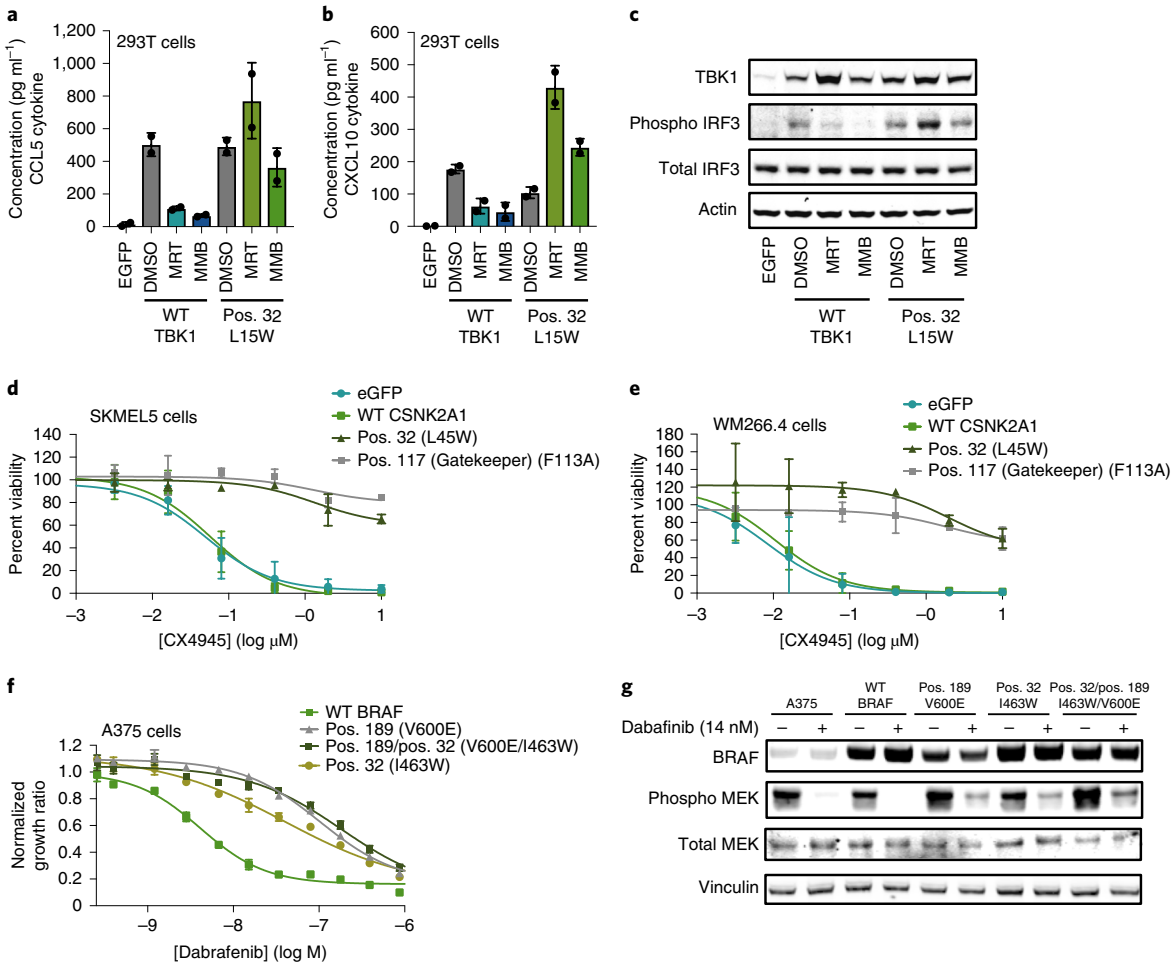


Fig. 6 | Validation of drug-resistance phenotypes predicted for the Pocket Protector residue (position 32) in TBK1, CSNK2A1 and BRAF. a,**b**, Assessment of WT and mutant TBK1-sensitivity to two inhibitors, MRT and MMB, in cytokine release assays measuring the average abundance of CCL5 (**a**) and CCL10 (**b**) ($n = 2$ replicate measurements for each drug), error bars represent s.d. **c**, Western blot assessment of WT- and mutant TBK1-function as in **a** and **b**, measured by downstream signaling (representative example of two experiments). **d**,**e**, Proliferation assay measuring the resistance of SKMEL5 (**d**) or WM266.4 (**e**) cells exogenously expressing WT or mutant CSNK2A1 to CX4945 ($n = 2$ replicate measurements for each mutant at each concentration). **f**, Proliferation assays measuring RAF-inhibitor sensitivity (dabrafenib) in A375 cells transduced with WT or mutant BRAF ($n = 4$ replicate measurements for each mutant at each concentration). **g**, Western blot assessment of WT- and mutant BRAF function measured by downstream signaling, representative example of two experiments. The uncropped westerns blots for **c** and **g** are shown in the source data. Statistical source data for **a**, **b** and **d**–**f** are available online.

effects of ALK inhibition (Fig. 7b)[82,83]. In this assay, expression of position-61 mutants in EGFR (EGFR^{L747P}) alone or in *cis* with EGFR^{L858R} (EGFR$^{L747P/L858}$) was sufficient to restore proliferation in the presence of ceritinib to a degree that approximated an EGFR^{L858R} mutant (Fig. 7b). Moreover, these mutants increased basal levels of phosphorylated EGFR, consistent with elevated activity (Fig. 7c). Similar results were observed with the position-61 mutant (HER2^{L755S}), whose expression is sufficient to rescue the antiproliferative effects of ER-inhibition in ER-dependent breast cancer cells (Fig. 7d,e, Extended Data Fig. 6b and ref. [84]). We (and others) have shown that position-61 mutants of EGFR and HER2 are drug resistant[72] (Fig. 5a–g).

We next asked if it was possible to predict and engineer an activating mutation at position 61 in a kinase where such an event had never been observed. MEK1 represented a reasonable candidate to test, as MEK1 marks the start of a recently characterized activating indel[76] and a position-61 mutant of MEK1 was identified in a mutagenesis screen for drug-resistant MEK1 mutants[21]. Thus, we leveraged our saturation mutagenesis data and previous subsaturation screens[21] to guide a series of position-61 mutations in MEK1, including MEK1^{I99T}, MEK1^{I99G} and MEK1^{I99M} (Table 1). Consistent with position-61 mediating resistance, MEK1^{I99T} and MEK1^{I99G} shifted the IC_{50} of trametinib more than tenfold over WT MEK1, whereas MEK1^{I99M} behaved like WT MEK1 (Extended Data Fig. 7c–e). To assess the activity of MEK1 mutants, we performed a pharmacological complementation assay in the BRAF-mutant, BRAF-dependent A375 cell line[32]. In these cells, growth suppression by BRAF inhibitors can be rescued by activated mutants of MEK1, such as MEK1$^{S218/222D}$ (MEK^{DD}) (Fig. 7f). Like MEK^{DD}, MEK1^{I99G} and MEK1^{I99T} rescued A375 proliferation in the presence of a BRAF inhibitor and restored downstream signaling to ERK, consistent with gain-of-activity properties (Fig. 7f,g and Extended Data Fig. 7d,e), whereas MEK1^{I99M} could not. To complement these experiments, we transiently transfected epitope-tagged WT MEK1, MEK^{DD}, MEK1^{I99G} or MEK1^{I99T} into 293T cells and assessed their ability to phosphorylate ERK2, a MEK1 substrate. Consistent with increased basal activity, expression of MEK1^{I99G} and MEK1^{I99T} led to elevated phosphorylation of ERK2 relative to WT MEK1

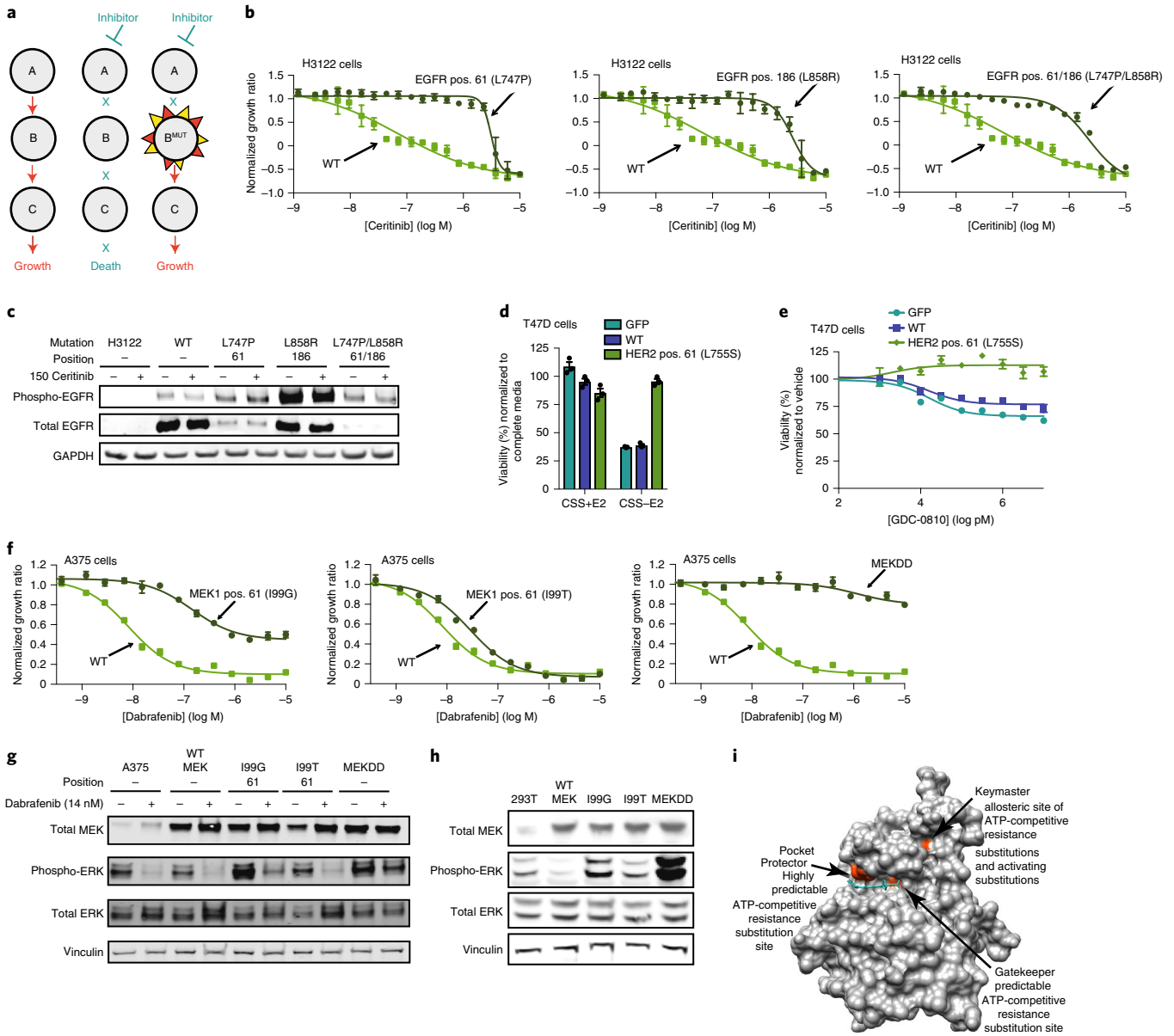


Fig. 7 | Mutation of the Keymaster (position-61) residue leads to gain-of-function activity in protein kinases. a, Schematic of the drug complementation assays used to assess protein activation in **b–f**. **b**, Complementation assays measuring proliferation of H3122 transduced with WT or mutant EGFR in the presence of the ALK inhibitor ceritinib ($n = 4$ replicate measurements for each mutant at each concentration). **c**, Western blot assessment of WT- and mutant EGFR function measured by downstream signaling, representative example of two experiments. **d**,**e**, Complementation assays measuring average proliferation of T47D cells transduced with WT or mutant HER2 in CSS ± E2 (**d**) or in the presence of the HER2-selective degrader GDC-0810 (**e**) ($n = 3$ replicate measurements for each mutant at each concentration), error bars represent s.d. **f**, Complementation assay measuring the proliferation of A375 cells transduced with WT or mutant MEK1 in the presence of the BRAF inhibitor dabrafenib ($n = 4$ replicate measurements for each mutant at each concentration). **g**, Western blot assessment of WT- and mutant-BRAF function measured by downstream signaling, representative example of two experiments. **h**, Assessment of WT or mutant MEK1 phosphotransferase activity as assessed through transfection in 293T cells and detection of phosphorylated ERK levels via western blotting, representative example of two experiments. **i**, Overall summary of Pocket Protector and Keymaster sites mapped onto a canonical kinase domain (CDK6; PDB 2EUF). The uncropped westerns blots for **c**, **g** and **h** are shown in the source data. Statistical source data for **b** and **d**–**f** are available online.

(Fig. 7h). For these mutants, MEK-inhibitor resistance phenotypes mirrored those observed in pharmacological complementation assays, suggesting that drug resistance and activity may be coupled. As position 61 is a central residue within a hydrophobic pocket ('lock') that sequesters the activation loop (the 'key'), we designate it the Keymaster, a functionally distinct complement to the Gatekeeper and Pocket Protector residues (Table 1, Fig. 7i and Extended Data Fig. 7f).

Discussion

A pressing challenge in kinase inhibitor development is the confirmation of on-target activity. Gold standard approaches to this issue involve generating drug-resistant mutants and showing they can rescue drug-mediated phenotypes. Gatekeeper mutations have been useful for this approach, but are limited by the challenges of predicting functional Gatekeeper mutations. Our study uncovered a suite of over 28 residues that mediate resistance, including three broadly

validated and mechanistically distinct residues, the Pocket Protector (position 32), Keymaster (position 61) and so-called Gatekeeper + 2 (position 119), generating a toolkit of resistance residues that, along with the Gatekeeper, can support the prediction of drug-resistant mutants for biological and therapeutic discovery (Table 1).

Additionally, we uncovered the Keymaster as a potentially generalizable activating mutational site. This knowledge may help identify rare, but potentially-actionable, oncogenic driver mutations in tumors, complementing genetic 'hotspot' analysis of kinase mutations in cancer[85]. It also may support the identification of kinases that exist in monomeric auto-inhibited states and dimerized activated states, which are difficult to discover in crystallization studies due to the requirement of high protein concentrations, and subsequent bias towards dimerization, as initially suggested by Kuryian and colleagues[81]. In this way, the identification of the Keymaster residue highlights how saturation mutagenesis can orthogonally support structural studies to help advance biological discovery[86].

The comparative approach taken here complements recent advances in mutagenesis, methods for sequence conservation and structural inference, but need not be restricted to kinases[87]. We envision that similar types of comparative phenotypic studies could be used for additional protein or domain families to support the discovery of unappreciated regulatory regions, increasing the efficacy of drug discovery and improving the clinical interpretation of disease-associated mutants.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41594-019-0358-z.

Received: 1 May 2019; Accepted: 27 November 2019;
Published online: 10 January 2020

References

1. Daub, H., Specht, K. & Ullrich, A. Strategies to overcome resistance to targeted protein kinase inhibitors. *Nat. Rev. Drug Discov.* **3**, 1001–1010 (2004).
2. Eyers, P. A., Craxton, M., Morrice, N., Cohen, P. & Goedert, M. Conversion of SB 203580-insensitive MAP kinase family members to drug-sensitive forms by a single amino-acid substitution. *Chem. Biol.* **5**, 321–328 (1998).
3. Fox, T. et al. A single amino acid substitution makes ERK2 susceptible to pyridinyl imidazole inhibitors of p38 MAP kinase. *Protein Sci.* **7**, 2249–2255 (1998).
4. Liu, Y. et al. Structural basis for selective inhibition of Src family kinases by PP1. *Chem. Biol.* **6**, 671–678 (1999).
5. Rotow, J. & Bivona, T. G. Understanding and targeting resistance mechanisms in NSCLC. *Nat. Rev. Cancer* **17**, 637–658 (2017).
6. Whittaker, S. et al. Gatekeeper mutations mediate resistance to BRAF-targeted therapies. *Sci. Transl. Med.* **2**, 35ra41 (2010).
7. Wylie, A. A. et al. The allosteric inhibitor ABL001 enables dual targeting of BCR-ABL1. *Nature* **543**, 733–737 (2017).
8. Azam, M., Seeliger, M. A., Gray, N. S., Kuriyan, J. & Daley, G. Q. Activation of tyrosine kinases by mutation of the gatekeeper threonine. *Nat. Struct. Mol. Biol.* **15**, 1109–1118 (2008).
9. Carlson, S. M. et al. Large-scale discovery of ERK2 substrates identifies ERK-mediated transcriptional regulation by ETV3. *Sci. Signal.* **4**, rs11 (2011).
10. Shah, K., Liu, Y., Deirmengian, C. & Shokat, K. M. Engineering unnatural nucleotide specificity for Rous sarcoma virus tyrosine kinase to uniquely label its direct substrates. *Proc. Natl Acad. Sci. USA* **94**, 3565–3570 (1997).
11. Blencke, S., Ullrich, A. & Daub, H. Mutation of threonine 766 in the epidermal growth factor receptor reveals a hotspot for resistance formation against selective tyrosine kinase inhibitors. *J. Biol. Chem.* **278**, 15435–15440 (2003).
12. Blencke, S. et al. Characterization of a conserved structural determinant controlling protein kinase sensitivity to selective inhibitors. *Chem. Biol.* **11**, 691–701 (2004).
13. Carter, T. A. et al. Inhibition of drug-resistant mutants of ABL, KIT, and EGF receptor kinases. *Proc. Natl Acad. Sci. USA* **102**, 11011–11016 (2005).
14. Gorre, M. E. et al. Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. *Science* **293**, 876–880 (2001).
15. Kobayashi, S. et al. EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* **352**, 786–792 (2005).
16. Pao, W. et al. Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. *PLoS Med.* **2**, e73 (2005).
17. Shah, N. P. et al. Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. *Cancer Cell* **2**, 117–125 (2002).
18. Tamborini, E. et al. A new mutation in the KIT ATP pocket causes acquired resistance to imatinib in a gastrointestinal stromal tumor patient. *Gastroenterology* **127**, 294–299 (2004).
19. Thomas, R. K. et al. Detection of oncogenic mutations in the EGFR gene in lung adenocarcinoma with differential sensitivity to EGFR tyrosine kinase inhibitors. *Cold Spring Harb. Symp. Quant. Biol.* **70**, 73–81 (2005).
20. Azam, M., Latek, R. R. & Daley, G. Q. Mechanisms of autoinhibition and STI-571/imatinib resistance revealed by mutagenesis of BCR-ABL. *Cell* **112**, 831–843 (2003).
21. Emery, C. M. et al. MEK1 mutations confer resistance to MEK and B-RAF inhibition. *Proc. Natl Acad. Sci. USA* **106**, 20411–20416 (2009).
22. Goetz, E. M., Ghandi, M., Treacy, D. J., Wagle, N. & Garraway, L. A. ERK mutations confer resistance to mitogen-activated protein kinase pathway inhibitors. *Cancer Res.* **74**, 7079–7089 (2014).
23. Lee, B. J. & Shah, N. P. Identification and characterization of activating ABL1 1b kinase mutations: impact on sensitivity to ATP-competitive and allosteric ABL1 inhibitors. *Leukemia* **31**, 1096–1107 (2017).
24. Marit, M. R. et al. Random mutagenesis reveals residues of JAK2 critical in evading inhibition by a tyrosine kinase inhibitor. *PLoS One* **7**, e43437 (2012).
25. Pahuja, K. B. et al. Actionable Activating Oncogenic ERBB2/HER2 Transmembrane and Juxtamembrane Domain Mutations. *Cancer Cell* **34**, 792–806.e5 (2018).
26. Tiedt, R. et al. A drug resistance screen using a selective MET inhibitor reveals a spectrum of mutations that partially overlap with activating mutations found in cancer patients. *Cancer Res.* **71**, 5255–5264 (2011).
27. Trowe, T. et al. EXEL-7647 inhibits mutant forms of ErbB2 associated with lapatinib resistance and neoplastic transformation. *Clin. Cancer Res.* **14**, 2465–2475 (2008).
28. Wagenaar, T. R. et al. Resistance to vemurafenib resulting from a novel mutation in the BRAFV600E kinase domain. *Pigment Cell Melanoma Res.* **27**, 124–133 (2014).
29. Williams, A. B. et al. Mutations of FLT3/ITD confer resistance to multiple tyrosine kinase inhibitors. *Leukemia* **27**, 48–55 (2013).
30. Zunder, E. R., Knight, Z. A., Houseman, B. T., Apsel, B. & Shokat, K. M. Discovery of drug-resistant and drug-sensitizing mutations in the oncogenic PI3K isoform p110α. *Cancer Cell* **14**, 180–192 (2008).
31. Ahler, E. et al. A combined approach reveals a regulatory mechanism coupling src's kinase activity, localization, and phosphotransferase-independent functions. *Mol. Cell* **74**, 393–408.e20 (2019).
32. Brenan, L. et al. Phenotypic characterization of a comprehensive set of MAPK1/ERK2 missense mutants. *Cell Rep.* **17**, 1171–1183 (2016).
33. Fowler, D. M. & Fields, S. Deep mutational scanning: a new style of protein science. *Nat. Methods* **11**, 801–807 (2014).
34. Giacomelli, A. O. et al. Mutational processes shape the landscape of TP53 mutations in human cancer. *Nat. Genet.* **50**, 1381–1387 (2018).
35. Gray, V. E., Hause, R. J. & Fowler, D. M. Analysis of large-scale mutagenesis data to assess the impact of single amino acid substitutions. *Genetics* **207**, 53–61 (2017).
36. Gray, V. E., Hause, R. J., Luebeck, J., Shendure, J. & Fowler, D. M. Quantitative missense variant effect prediction using large-scale mutagenesis data. *Cell Syst.* **6**, 116–124.e3 (2018).
37. Kitzman, J. O., Starita, L. M., Lo, R. S., Fields, S. & Shendure, J. Massively parallel single-amino-acid mutagenesis. *Nat. Methods* **12**, 203–206 (2015).
38. Majithia, A. R. et al. Prospective functional classification of all possible missense variants in PPARG. *Nat. Genet.* **48**, 1570–1575 (2016).
39. Matreyek, K. A. et al. Multiplex assessment of protein variant abundance by massively parallel sequencing. *Nat. Genet.* **50**, 874–882 (2018).
40. Mavor, D. et al. Determination of ubiquitin fitness landscapes under different chemical stresses in a classroom setting. *eLife* **5**, e15802 (2016).
41. Melnikov, A., Rogov, P., Wang, L., Gnirke, A. & Mikkelsen, T. S. Comprehensive mutational scanning of a kinase in vivo reveals substrate-dependent fitness landscapes. *Nucleic Acids Res.* **42**, e112 (2014).
42. Ray, A., Cowan-Jacob, S. W., Manley, P. W., Mestan, J. & Griffin, J. D. Identification of BCR-ABL point mutations conferring resistance to the Abl kinase inhibitor AMN107 (nilotinib) by a random mutagenesis study. *Blood* **109**, 5011–5015 (2007).

43. Kolde, R., Laur, S., Adler, P. & Vilo, J. Robust rank aggregation for gene list integration and meta-analysis. *Bioinformatics* **28**, 573–580 (2012).
44. Aronov, A. M. et al. Structure-guided design of potent and selective pyrimidylpyrrole inhibitors of extracellular signal-regulated kinase (ERK) using conformational control. *J. Med. Chem.* **52**, 6362–6368 (2009).
45. Martinez Molina, D. et al. Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. *Science* **341**, 84–87 (2013).
46. Yun, C. H. et al. The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proc. Natl Acad. Sci. USA* **105**, 2070–2075 (2008).
47. Chen, L., Fu, W., Zheng, L., Liu, Z. & Liang, G. Recent progress of small-molecule epidermal growth factor receptor (EGFR) inhibitors against C797S resistance in non-small-cell lung cancer. *J. Med. Chem.* **61**, 4290–4300 (2018).
48. Costa, D. B. et al. BIM mediates EGFR tyrosine kinase inhibitor-induced apoptosis in lung cancers with oncogenic EGFR mutations. *PLoS Med.* **4**, 1669–1679 (2007).
49. Costa, D. B., Schumer, S. T., Tenen, D. G. & Kobayashi, S. Differential responses to erlotinib in epidermal growth factor receptor (EGFR)-mutated lung cancers with acquired resistance to gefitinib carrying the L747S or T790M secondary mutations. *J. Clin. Oncol.* **26**, 1182–1184 (2008).
50. Ercan, D. et al. EGFR mutations and resistance to irreversible pyrimidine-based EGFR inhibitors. *Clin. Cancer Res.* **21**, 3913–3923 (2015).
51. Kohsaka, S. et al. A method of high-throughput functional evaluation of EGFR gene variants of unknown significance in cancer. *Sci. Transl. Med.* **9**, eaan6566 (2017).
52. Liu, Y. et al. Acquired EGFR L718V mutation mediates resistance to osimertinib in non-small cell lung cancer but retains sensitivity to afatinib. *Lung Cancer* **118**, 1–5 (2018).
53. Ou, S. I. et al. Emergence of novel and dominant acquired EGFR solvent-front mutations at Gly796 (G796S/R) together with C797S/R and L792F/H mutations in one EGFR (L858R/T790M) NSCLC patient who progressed on osimertinib. *Lung Cancer* **108**, 228–231 (2017).
54. Wang, Y. T., Ning, W. W., Li, J. & Huang, J. A. Exon 19 L747P mutation presented as a primary resistance to EGFR-TKI: a case report. *J. Thorac. Dis.* **8**, E542–E546 (2016).
55. Yamaguchi, F. et al. Acquired resistance L747S mutation in an epidermal growth factor receptor-tyrosine kinase inhibitor-naive patient: a report of three cases. *Oncol. Lett.* **7**, 357–360 (2014).
56. Goldberg, M. E. et al. Multiple configurations of EGFR exon 20 resistance mutations after first- and third-generation EGFR TKI treatment affect treatment options in NSCLC. *PLoS One* **13**, e0208097 (2018).
57. Yang, Z. et al. Investigating novel resistance mechanisms to third-generation EGFR tyrosine kinase inhibitor osimertinib in non-small cell lung cancer patients. *Clin. Cancer Res.* **24**, 3097–3107 (2018).
58. Bitencourt, R., Zalcberg, I. & Louro, I. D. Imatinib resistance: a review of alternative inhibitors in chronic myeloid leukemia. *Rev. Bras. Hematol. Hemoter.* **33**, 470–475 (2011).
59. Connor, L. M. O., Langabeer, S., McCann, S. R. & Conneally, E. Mutation mediated resistance to the tyrosine kinase inhibitors imatinib, dasatinib & nilotinib in Philadelphia positive leukaemia. *Blood* **112**, 4245–4245 (2008).
60. Muller, M. C. et al. Dasatinib treatment of chronic-phase chronic myeloid leukemia: analysis of responses according to preexisting BCR-ABL mutations. *Blood* **114**, 4944–4953 (2009).
61. Yoda, S. et al. Sequential ALK inhibitors can select for lorlatinib-resistant compound ALK mutations in ALK-positive lung cancer. *Cancer Discov.* **8**, 714–729 (2018).
62. Zhang, S. et al. The potent ALK inhibitor brigatinib (AP26113) overcomes mechanisms of resistance to first- and second-generation ALK inhibitors in preclinical models. *Clin. Cancer Res.* **22**, 5527–5538 (2016).
63. Boulbes, D. R. et al. HER family kinase domain mutations promote tumor progression and can predict response to treatment in human breast cancer. *Mol. Oncol.* **9**, 586–600 (2015).
64. Sun, Z. et al. Analysis of different HER-2 mutations in breast cancer progression and drug resistance. *J. Cell. Mol. Med.* **19**, 2691–2701 (2015).
65. Berger, A. H. et al. High-throughput phenotyping of lung cancer somatic mutations. *Cancer Cell* **30**, 214–228 (2016).
66. Ma, L. et al. CRISPR-Cas9-mediated saturated mutagenesis screen predicts clinical drug resistance with improved accuracy. *Proc. Natl Acad. Sci. USA* **114**, 11751–11756 (2017).
67. Jenkins, R. W. et al. Ex vivo profiling of PD-1 blockade using organotypic tumor spheroids. *Cancer Discov.* **8**, 196–215 (2018).
68. Zhu, Z. et al. Inhibition of KRAS-driven tumorigenicity by interruption of an autocrine cytokine circuit. *Cancer Discov.* **4**, 452–465 (2014).
69. Wu, J. Y. et al. Second-line treatments after first-line gefitinib therapy in advanced nonsmall cell lung cancer. *Int. J. Cancer* **126**, 247–255 (2010).
70. Wu, J. Y. et al. Effectiveness of tyrosine kinase inhibitors on "uncommon" epidermal growth factor receptor mutations of unknown clinical significance in non-small cell lung cancer. *Clin Cancer Res* **17**, 3812–3821 (2011).
71. Wu, S. G. et al. Frequent epidermal growth factor receptor gene mutations in malignant pleural effusion of lung adenocarcinoma. *Eur. Respir. J.* **32**, 924–930 (2008).
72. Kancha, R. K. et al. Differential sensitivity of ERBB2 kinase domain mutations towards lapatinib. *PLoS One* **6**, e26760 (2011).
73. Nagano, M. et al. High-throughput functional evaluation of variants of unknown significance in ERBB2. *Clin. Cancer Res.* **24**, 5112–5122 (2018).
74. Xu, X. et al. HER2 reactivation through acquisition of the HER2 L755S mutation as a mechanism of acquired resistance to HER2-targeted therapy in HER2$^+$ breast cancer. *Clin. Cancer Res.* **23**, 5123–5134 (2017).
75. Foster, S. A. et al. Activation mechanism of oncogenic deletion mutations in BRAF, EGFR, and HER2. *Cancer Cell* **29**, 477–493 (2016).
76. Gao, Y. et al. Allele-specific mechanisms of activation of MEK1 mutants determine their properties. *Cancer Discov.* **8**, 648–661 (2018).
77. He, M. et al. EGFR exon 19 insertions: a new family of sensitizing EGFR mutations in lung adenocarcinoma. *Clin. Cancer Res.* **18**, 1790–1797 (2012).
78. Kumar, R. D. & Bose, R. Analysis of somatic mutations across the kinome reveals loss-of-function mutations in multiple cancer types. *Sci. Rep.* **7**, 6418 (2017).
79. Schulze-Gahmen, U., De Bondt, H. L. & Kim, S. H. High-resolution crystal structures of human cyclin-dependent kinase 2 with and without ATP: bound waters and natural ligand as guides for inhibitor design. *J. Med. Chem.* **39**, 4540–4546 (1996).
80. Thevakumaran, N. et al. Crystal structure of a BRAF kinase domain monomer explains basis for allosteric regulation. *Nat. Struct. Mol. Biol.* **22**, 37–43 (2015).
81. Zhang, X., Gureasko, J., Shen, K., Cole, P. A. & Kuriyan, J. An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell* **125**, 1137–1149 (2006).
82. Wilson, F. H. et al. A functional landscape of resistance to ALK inhibition in lung cancer. *Cancer Cell* **27**, 397–408 (2015).
83. Yamada, T. et al. Paracrine receptor activation by microenvironment triggers bypass survival signals and ALK inhibitor resistance in EML4-ALK lung cancer cells. *Clin. Cancer Res.* **18**, 3592–3602 (2012).
84. Nayar, U. et al. Acquired HER2 mutations in ER$^+$ metastatic breast cancer confer resistance to estrogen receptor-directed therapies. *Nat. Genet.* **51**, 207–216 (2019).
85. Bose, R. et al. Activating HER2 mutations in HER2 gene amplification negative breast cancer. *Cancer Discov.* **3**, 224–237 (2013).
86. Leung, G. P. et al. Hyperactivation of MAPK signaling is deleterious to RAS/RAF-mutant melanoma. *Mol. Cancer Res.* **17**, 199–211 (2019).
87. Shah, N. H. & Kuriyan, J. Understanding molecular mechanisms in cell signaling through natural and artificial sequence variation. *Nat. Struct. Mol. Biol.* **26**, 25–34 (2019).

Methods

Saturation mutagenesis screens for testing CDK6 and CDK4 resistance to palbociclib. MITE was conducted on human CDK4 and CDK6 as described previously[32]. Saturation mutagenesis pools were made in pUC57 using codon-optimized WT sequences and subsequently cloned into a dox-inducible vector for downstream viral expression[32]. For CDK6 screening-scale spin-infections were performed in a 12-well format with ~5×10^7 cells per replicate to achieve a representation of at least 1,000 cells per variant following puromycin selection. Approximately 24 h after infection, all wells within a replicate were pooled and were split into T225 flasks. At 48 h after infection, cells were selected with puromycin for 4 d to remove uninfected cells. Six days after selection cells were seeded and allowed to adhere. Pretreatment, 'early time point' reference samples ($n=2$) were collected from Meljuso cells for sequencing to establish mutant abundance before drug treatment. Remaining replicates were treated with either a vehicle control (DMSO, $n=2$) or 2 μM palbociclib ($n=4$). Cells were passaged in drug or fresh media containing drugs every 3–4 d. Cells were harvested 20 d after treatment was initiated. Genomic DNA was isolated using Midi kits according to the manufacturer's protocol (Qiagen). CDK4 screens were undertaken in exactly the same way as CDK6.

Cell lines used in this study. Meljuso, A375, PC9 (PC14), H3122, T47D, 293T, SKMEL5 and WM266.4 cell lines were acquired via the Cancer Cell Line Encyclopedia (http://www.broadinstitute.org/ccle/home), confirmed to be mycoplasma negative and authenticated via Fluidigm single-nucleotide polymorphism-based fingerprinting.

Quantification of variant enrichment and depletion in high throughput screens. Non-WT CDK6 mutations identified through massively parallel sequencing were normalized to the total read counts per position using ORFCall script, accessed at https://github.com/tedsharpe/ORFCall. In these assays, sequencing cannot disambiguate WT reads that come from fully WT CDK6 from those emanating from the nonmutant portion of a mutant molecule. As such, following normalization wild-type reads are uninformative and were discarded. Replicate reproducibility was checked and for each mutation, LFC was calculated between the median of the drugged replicates and the median of the early time point replicates, (Median of ($\log_{2(\text{drugged A})}$, $\log_{2(\text{drugged B})}$, $\log_{2(\text{drugged C})}$, $\log_{2(\text{drugged D})}$)) – (Median of ($\log_{2(\text{early timepoint A})}$, $\log_{2(\text{early timepoint B})}$)). *Z*-scores were calculated for each mutation as LFC – (average LFC of all mutations)/(s.d. of the LFC across all mutations). CDK4 screens were analyzed in the same way as CDK6.

Alignment of kinase domains. We aligned the kinase domain protein sequences of human CDK6, CDK4, ERK2, ABL1, HER2 and BRAF obtained from canonical sequences available from Uniprot. The entire sequence was used for CDK6, CDK4 and ERK2, which are comprised almost entirely of the kinase domain. Initial 2D alignments using both Clustal Omega and MUSCLE were performed on all previously mentioned kinases using the entire protein sequence of ABL1, HER2 and BRAF. Crystal structures published in the Protein Data Bank (PDB) were obtained for CDK6 inhibited by palbociclib and compared using CHIMERA to relevant kinase domain crystal structures for all other kinases in the alignment using the University of California, San Francisco (UCSF) Chimera package[88]. By manually spot-checking each 2D alignment against 3D alignments for conserved residues known as the activating lysine, Gatekeeper residue, DFG motif and glycines of the p-loop we iteratively trimmed the sequence of ABL1, HER2 and BRAF to encompass only their kinase domains and realigned using both Clustal Omega and MUSCLE[89] tools to ensure that all structurally conserved residues lined up. Ultimately we used sequences encompassing residues 230–609 for ABL1, 706–1,063 for HER2 and 443–766 for BRAF together with the entire sequences for CDK6, CDK4 and ERK2 and a common numbering system was generated using MUSCLE alignments. We proceeded to compare resistance phenotypes for these positions across paralogous kinases at 419 positions.

Mapping of resistance screening data onto aligned kinase domains. For inclusion in this analysis, subsaturation mutagenesis datasets were required to have discovered alterations in the Gatekeeper residue as well as at least ten additional residues. For all screens, data were reduced down to a single metric of resistance representing the maximum phenotypic score observed at each residue of each kinase for which screening data were available. For ERK2, CDK6 and CDK4, we chose the substitution that produced the maximal resistance phenotype for each residue and used its *z*-score. Similarly, for subsaturation screening data, the maximum variant detection metric (BRAF) or colony/isolate count (HER2, BCR-ABL) was used. To identify 'resistance positions', we established phenotypic thresholds for each data set. For the ERK2, CDK4 and CDK6 data, mutants with a *z*-score > 1.95 were considered drug-resistant mutants, whereas for subsaturation mutagenesis screens, we used the thresholds reported in the original publications (BRAF) or considered all clones/isolates that survived a drug challenge (BCR-ABL, HER2). We then mapped data that passed these thresholds onto the amino acid alignment, revealing common drug-resistant positions. All proteins were human in origin.

Statistical calculations of drug-resistant hotspots. We used a robust-rank aggregation[43] to evaluate the statistical significance of drug-resistance hotspots, incorporating the recurrence of observation and strength of resistance phenotypes across kinases. For individual kinases that had been screened against dual inhibitors (ERK2 and BCR-ABL), we restricted our analysis to the maximum observed phenotypes across each pair of screens to avoid overweighting the analysis towards these proteins. While the quantification of the resistance phenotype included a diversity of metrics, including log(fold change), median enrichment and colony counts, we ultimately used these to rank-order the residues in each protein from the strongest to the weakest resistance phenotype, alleviating differences in absolute resistance metrics as a result of different detection methods. The significance of the observed recurrence was then computed as compared to a null model of uncorrelated ranking. The resulting *P* values of all paralogous positions were then corrected for multiple hypothesis testing using an FDR approach[90].

Vectors used. For saturation mutagenesis screens, human CDK6 and CDK4 were synthesized in pUC57 using optimized codons. These vectors were subsequently used for MITE and used for saturation mutagenesis screens for palbociclib resistance. For low-throughput validation, full-length WT and mutant CDK6 for K443M, D104S, F98E, F28R, G25S, T58F and V45M were synthesized in pdonor223 without codon optimization and without stop codons (GenScript) and then stop codons added using QuikChange Lightning mutagenesis. The following CDK6 mutants in pdonor223 were generated from the above WT CDK6 using QuikChange Lightning (Agilent) mutagenesis: I19W, F98E, H100F and K43M/V45M. K43M was added additionally to the above mutants in pdonor223 to create K43M/H100F, K43M/F98E and H100F. CDK4 WT pdonor223 plasmids were synthesized by GenScript. CDK4 mutants I12A, F93E, H95F and V37M were generated from WT CDK4 using QuikChange Lightning mutagenesis. All CDK4 and CDK6 mutants were confirmed using Sanger sequencing then recombined into expression vectors using Gateway LR Clonase II (Invitrogen) for constitutive expression in Meljuso cells. EML4-ALK WT and L1152R in the pLVX317 vector were kindly gifted from Fredrick Wilson, Yale University. WT ERK2 was synthesized in pdonor223 (GenScript) as indicated[32]. ERK2 (human) mutations were generated using QuikChange Lightning mutagenesis, and all ERK2 constructs were transferred by LR-mediated recombination into the lentiviral vector (AddGene) pLVX307 for constitutive expression in A375 cells. Constructs containing human HER2 WT and L755S mutations were obtained from Nick Wagle's lab at the Dana-Farber Cancer Institute[84]. Human *CSNK2A1* gene cDNAs were obtained from Harvard's PlasmID (Boston) and mutagenesis of CSNK2A1 was performed using InFusion-based mutagenesis (Takara Bio). Sanger sequencing was done to verify mutants. Gateway Cloning was used to transfer cDNA into plX304 vector. For human TBK1, plx980-TBK1 and subsequent mutants were obtained from D.A.B.'s lab at the Dana-Farber Cancer Institute. For human EGFR, WT, T790M, L858R and T790M/L858R mutants were obtained in pLX317 from the Broad Institute collection. All other EGFR mutations were made by incorporating WT protein into pdonor223 and using QuikChange Lightning mutagenesis and then transferred by LR-mediated recombination into the lentiviral vector (AddGene) pLVX307 for constitutive expression in PC9 or H3122 cells. Human BRAF WT, BRAF V600E, MEK1 WT and MEK1DD were obtained in pdonor223 and pdonor221 from the Broad Institute collection and stop sites were added using QuikChange Lightning mutagenesis. Other BRAF and MEK1 mutations were generated using QuikChange Lightning mutagenesis using these templates. All BRAF and MEK1 constructs were transferred by LR-mediated recombination into the lentiviral vector (AddGene) pLVX307 for constitutive expression in A375 cells. For MEK1 V5pulldown experiments, the V5-tagged versions (without a stop codon between the gene end and the V5-tag encoded by pLVX307) were used.

Lentiviral infection of cell lines. Unless otherwise indicated, lentivirus was produced by transiently transfecting 293T cells with 1 μg of lentivirus expression vector, 900 ng of psPAX2 and 100 ng of VSV-G packaging system and 6 μl of Fugene transfection reagent (Promega). After 72 h, viral supernatant was harvested, aliquoted and placed at −80 °C for at least 24 h. All cells were plated in six-well plates with 2 ml of media per well then infected 24 h later with 200–500 μl virus and an appropriate amount of polybrene were then added to each well. Plates were spun at 2,250 r.p.m. on a tabletop centrifuge for 1 h. After centrifugation, the media was removed and replaced with complete growth media. After 24–48 h media the appropriate amount of puromycin was added and cells were selected until an uninfected control sample had been completely killed off. The cell number plated 24 h before infection, the amount of polybrene and the amount of puromycin used were as follows: for all CDK4 and CDK6 expressing cells, Meljuso cells were infected with 300,000–400,000 cells plated in each well of a six-well dish with 2 μg ml^{-1} of polybrene and selected with 1 μg ml^{-1} of puromycin. For EGFR expressing cells, PC9 or H3122 cells were plated at approximately 300,000–400,000 cells per well with 8 μg ml^{-1} polybrene and selected with 1 μg ml^{-1} puromycin. BRAF, MEK1 and ERK2 expressing viruses were infected with approximately 300,000–400,000 A375 cells per well, 7.5 μg ml^{-1} polybrene, 1 μg ml^{-1} puromycin. For cells with CRISPR/Cas9-mediated knockout of RB1, an sgRNA targeting RB1 was expressed from a lentiviral vector and viral particles were produced in 293T cells using the Δ8.9 and VSV-G packaging plasmids transfected with XtremeGENE9 (Sigma Aldrich). Infection of Meljuso cell lines was done in growth

media supplemented with 5 μg ml^{-1} polybrene and selected with 2 μg ml^{-1} of blasticidin for 72 h.

Transient transfection of MEK1 mutations in HEK293T cells. HEK293T cells were transiently transfected with 10 μg of C-terminally V5-tagged human MEK1 lentivirus vector, 750 μl of optimem and 50 μl of Fugene (Promega) preincubated for 45 min and then added to 10 ml of media on roughly 70% confluent 10 cm plates and harvested 24–48 h later.

Population doublings for CDK6 and CDK4 mutants. WT and mutant CDK6 expression Meljuso cells were seeded into 2 ml of RPMI media supplemented with 10% FBS and 1% pen-strep in six-well plates at a density of 300,000 cells per well in the absence or presence of 2 μM palbociclib (Selleck Chemicals). Every 3–4 d, media was removed, cells were washed with 1 ml of PBS, trypsinized in 200 μl of Trypsin-EDTA, then resuspended with 800 ml of media. The solution was washed with another 1 ml of media and then the number of cells in 500 μl of the 2 ml of solution was counted on a ViCell and recorded, and 300,000 cells were then re-seeded into fresh six-well plates. The volume was raised to 2 ml of media in the absence or presence of 2 μM palbociclib (Selleck Chemicals). For some of the sensitive cell lines, slightly less than 300,000 cells were plated for the last two time points of the 14-d experiment, but this was taken into account when calculating the cumulative population doubling. CDK4 population doublings were performed in the same manner as CDK6 except that only 1 μM of palbociclib was used.

Cellular thermal shift assays. CETSA was performed as previously reported for CDK6 (ref. [45]) with the following modifications. For each mutant tested, 22,000,000 cells were spun in a 50-ml conical tube for 3 min at 300*g*, and supernatant was then carefully aspirated and washed by pipetting in and swirling 10 ml of PBS. The PBS was aspirated, and the cells were resuspended in 2.2 ml of cold buffer (25 mM Tris pH 7.5, 2 mM DTT, 10 mM $MgCl_2$, 5 mM beta-glycerophosphate, 1× complete protease inhibitors (Roche) with EDTA) and flash-frozen in liquid nitrogen. Cells were thawed and immediately refrozen in liquid nitrogen. This was repeated once more. From this step forward, with the exception of the temperature gradient, lysate remained on ice or at 4 °C. The lysed cells were spun at 20,000*g* on a tabletop centrifuge for at least 30 min at 4 °C. The cleared lysate was removed and saved, leaving behind a small amount of liquid so as not to disturb the pellet, and added to a conical tube.

The sample was split into two equivalent volumes on ice (500 μl to 1 ml), and to one was added 20 mM palbociclib to a final concentration of 100 μM (such as 5 μl to 1 ml) and to the other tube was added the equivalent volume of water. The lysate solutions were then mixed with gentle pipetting, avoiding bubbles, and incubated on ice for at least 30 min. Following incubation, each sample was split into 50-μl aliquots into PCR tubes on ice. PCR tubes were then placed in a gradient PCR machine and subjected to 3 min at respective temperatures (40, 43, 46, 49, 52, 55, 58, 61 or 64 °C) followed by 3 min at room temperature, and then returned to ice. Samples were then transferred to Eppendorf tubes, and spun on a tabletop centrifuge at 20,000*g* for at least 30 min at 4 °C. A total of 30 μl of each sample was removed carefully not to disturb the pellet and 6 μl of 6× gel loading dye was then added to each sample. Western blots were run on the samples to evaluate soluble CDK6 as indicated in the western blot method. Bands for soluble CDK6 were quantitated using ImageStudio software, and the intensity of each band was normalized to the 40 °C band and duplicate experiments were graphed on Graphpad to produce melting curves. To calculate melting temperatures, all replicate points were fitted together in Graphpad to a Boltzman sigmoidal equation[91] and then the calculated V_{50}, being the temperature at which the curve is halfway between the top and the bottom of the curve, represents the melting temperature reported here. The associated s.e.m. of the curve fit represents the error in the melting temperature. To calculate changes in melting temperature, melting temperature in the absence of drug were subtracted from melting temperatures in the presence of drug and the errors were combined using standard propagation of error calculations as follows: Error in the change in melting temperature = the square root of ((error in melting temperature with palbociclib)^2 + (error in the absence of palbociclib)^2).

Differential scanning fluorescence assays. WT CDK6, I19W CDK6 and V45M CDK6 were subcloned into plx307 with an N-terminal FLAG tag. H100F CDK6 and F98E CDK6 were made by QuikChange Lightning site-directed mutagenesis of FLAG-tagged WT CDK6. For each construct, 5–6 10-cm plates of ~70% confluent 293T cells were each transiently transfected with 50 μl of FuGene reagent (Promega) and 20 μg of plasmid DNA; 24 h later the media was changed and 24 h after that cells were trypsinized, washed with PBS and spun at 3 min at 300*g*. The fluid was aspirated and the cell pellet was flash-frozen in liquid nitrogen. Cell pellets were thawed in 3 ml of 1× cell lysis buffer (Cell Signaling Technology) and left on ice for 30 min to lyse. The lysate was cleared by spinning on a tabletop centrifuge at 21,130*g* for 1 h to remove the debris. Cleared lysate for each form of CDK6 was combined. For each mutant three Eppendorf tubes of 50 μl of anti-FLAG M2 affinity gel suspension (Sigma) were prepared by spinning for 3 min at 1,000*g*, removing the solution, washing with 500 μl each of 1× cell lysis buffer (Cell Signaling Technology). Cleared lysate for each mutant was split into three tubes and rotated at 4° for 2–2.5 h. Each tube was spun at 1,000*g* for 3 min, lysate was removed and the beads were washed 2× with 750 μl of cell lysis buffer (Cell Signaling Technology) followed by 1× with 750 μl of 200 mM NaCl, 50 mM Tris 7.5, and 1x complete protease inhibitors (Roche). Each tube of FLAG-CDK6 was eluted first in 100 μl of 283 μg ml^{-1} 3× FLAG peptide (Sigma) in 250 mM NaCl, 50 mM Tris 7.5 and 1× complete protease inhibitors by rocking at 4° for 30–40 min, spinning at 1,000*g* for 3 min and removing the fluid from the gel. A second elution was also performed in 100 μl of the same elution buffer but rocking for 15 min. The six eluates for each mutant were combined and spun through a Pierce spin column (Thermo Scientific) to remove residual gel; buffer was exchanged to remove excess 3× FLAG peptide and concentrated using repeated concentration, dilution and concentration in a spin-X UF 500 -μl 10k MWCO PES spin column (Corning) using a buffer consisting of 200 mM NaCl, 50 ml Tris 7.5 and 1× complete protease inhibitor. The protein was concentrated to 3.4–3.6 mg ml^{-1} and 1 M DTT was added to a final concentration of 10 mM. The purified protein constructs were buffer exchanged in the assay buffer 25 mM Tris pH 7.5, 10 mM $MgCl_2$, 2 mM DTT and β-glycerophosphatase inhibitor using a 0.5-ml Zeba Spin desalting column (7 kDa MWCO) following the manufacturer's protocol. The sample preparation for the DSF assay included incubating the different protein constructs (final concentration) 13.5 μM with either water alone (control) or with 40 μM palbociclib for 30 min before the addition of Sypro orange dye (final concentration 10×). The final reaction volume of 10 μl is pipetted into a 384 multi-well clear plate (Light Cycler-480), sealed using a foil and spun down at 1,000*g* for 2 min. The fluorescence emission intensity of the Sypro orange dye was monitored in a Light Cycler-480 instrument by running the protein melt program. The program was run in a continuous acquisition mode from 25 °C to 95 °C with a ramp rate of 0.06 °C s^{-1} (with 10 acquisitions per °C). The data (first derivative of the fluorescence signal as a function of temperature) were analyzed using the Light Cycler-480 software for measuring the melting temperature values with and without the compound. A separate control with buffer alone and buffer + compound was run using the same plate to subtract out the potential interactions of the buffer components and/or the compound with the dye. Error was propagated as in the CETSA experiments.

Drug sensitivity evaluation of EGFR, ERK2, EML-ALK, BRAF and MEK1 mutants. EGFR mutant-expressing PC9 or H3122 cells and EML4-ALK H3122 mutant-expressing cells were seeded into 384-well, white-walled, clear bottom plates at a density of 800 or 1,000 cells per well, respectively. Doxycycline-inducible ERK2, MEK1 and BRAF-mutant-expressing A375 cells were seeded into 384-well, white-walled, clear bottom plates at a density of 500 cells per well. Twenty-four hours after seeding, the initial viability (day 0) of a subset of plated cells was established at the same time as compound addition using the CellTiterGlo viability assay (Promega). Doxycycline (from a 0.3% Tween stock solution) was added to the remaining cells where applicable and drugs (from DMSO solutions) were added using an HP D300 Digital Dispenser. After the addition of drugs, cells were incubated for 96 h and cell viability was measured using a CellTiterGlo viability assay (Promega). Normalized growth rate inhibition was calculated using day 0 and endpoint DMSO data.

Drug sensitivity or protein activity evaluation of HER2 mutations. T47D HER2 wild-type and mutant-expressing cells were plated in RPMI with 10% charcoal-stripped serum (CSS) at 1,000 cells per well of 96-well Viewplates (Perkin-Elmer). After 2 d of serum starvation, cells were switched over to complete media or treated with estradiol (E2, Sigma-Aldrich) as appropriate, and treated with a range of doses of the corresponding drug (either lapatinib (Selleck Chemicals) or GDC-0810 (Medkoo)). Cells were re-treated after 3 d. One week after the start of treatment, viability was determined using a CellTiterGlo viability assay (Promega).

Validation of drug resistance for TBK1 mutations. Cytokine release assays were initiated by plating 3×10^5 HEK293T cells onto a six-well plate, where they were pretreated with 5 μM MRT or 5 μM MMB for 12 h and transfected using X-tremeGENE HP DNA transfection reagent (Roche catalog no. 06366236001) with plx980-TBK1 or plx980-TBK1 L15W in the presence of MRT or MMB. Conditioned media were collected 24 h after transfection. CXCL10 (R&D Systems catalog no. DIP100) and CCL5 (R&D Systems catalog no. DRN00B) ELISAs were performed according to manufacturer's instructions. Immunoblotting was performed as previously described[68].

Validation of drug resistance for Casein Kinase II mutations. Lentivirus production and infection of the *CSNK2A1* gene cDNA into HEK293T cells were performed as described above. Cells were plated in 96-well plates 16–24 h before treatment. Drug was added to the indicated concentration, with controls treated with equal volumes of DMSO. Measurement of cell number was estimated by CellTiterGlo (Promega) 48 h after drug treatment. SKMEL5 (ATCC) transduced with indicated viruses were treated with CX4945 (Selleck) or vehicle control for 24 h before analysis by western blotting.

Western blots. Adherent cells were washed with ice-cold PBS and lysed with 1% NP-40 buffer (150 mM NaCl, 50 mM Tris pH 7.5, 2 mM EDTA pH 8, 25 mM NaF and 1% NP-40) containing Complete Mini EDTA-free protease inhibitors (Roche)

and Phosphatase Inhibitor Cocktails I and II (CalBioChem). Lysate was cleared by spinning at 15,000 r.p.m. on a tabletop centrifuge for at least 10 min, and soluble lysate protein content was quantified (BCA assay) normalized, then reduced and denatured (95 °C), and resolved by SDS gel electrophoresis on 4–12% Bis–Tris gels (Invitrogen). Resolved protein was transferred to nitrocellulose membranes using the iBlot system (Invitrogen), blocked in LiCOR blocking buffer, probed with either mouse or rabbit primary antibodies, followed by LiCOR secondary IRDye 680RD goat anti-mouse (no. 926-68070) and IRDye 800CW goat anti-rabbit (no. 926-32211) antibodies, and then imaged on a Licor Odyssey CLx Infrared Imaging system. For TBK1 western blots, cells were lysed in RIPA buffer containing 1× protease inhibitors (Roche) and phosphatase inhibitors (50 mM NaF and 100 mM Na_3VO_4). CSNK2A1 western blotting was done as previously described[92].

Primary antibodies used. Phospho-Rb (Ser780) (D59B7) rabbit mAb (Cell Signaling Technology, 8180S), Rb (4H1) mouse mAb (Cell Signaling Technology, no. 9309S), Cdk6 (C-21) rabbit polyclonal IgG (Santa Cruz Biotechnology, sc-177), Cdk4 (DCS-35) mouse monoclonal IgG_1 (Santa Cruz Biotechnology, sc-23896), beta-actin (8H10D10) mouse mAb (Cell Signaling Technology, no. 3700S), monoclonal anti-vinculin antibody (Sigma-Aldrich, V9131), GAPDH (14C10) rabbit mAb (Cell Signaling Technology, no. 2118S), ALK (31F12) mouse mAb (Cell Signaling Technology, no. 3791S), phospho-ALK (Tyr1278/1282/1283) antibody (Cell Signaling Technology, no. 3983S), Akt (pan) (40D4) mouse mAb (Cell Signaling Technology, no. 2920), phospho-Akt (Ser473) (D9E) XP rabbit mAb (Cell Signaling Technology, no. 4060L), p44/42 MAPK (Erk1/2) (L34F12) mouse mAb (Cell Signaling Technology, no. 4696S), phospho-p44/42 MAPK (Erk1/2) (Thr202/Tyr204) antibody (Cell Signaling Technology, no. 9101S), Cofilin (D3F9) XP rabbit mAb (Cell Signaling Technology, no. 5175S), Raf-B antibody (F-7) mouse monoclonal IgG (Santa Cruz Biotechnology, sc-5284), MEK1/2 rabbit mAB (Cell Signaling Technology, no. 9122L), phospho-MEK1/2 (Ser217/221) (41G9) rabbit mAb (Cell Signaling Technology, no. 9154), RSK1 (D6D5) rabbit mAb (Cell Signaling Technology, no. 8408), phospho-p90RSK (Ser380) antibody (Cell Signaling Technology, no. 9341), EGF Receptor (1F4) mouse mAb (Cell Signaling Technology, no. 2239), phospho-EGF Receptor (Tyr1068) (D7A5) XP rabbit mAb (Cell Signaling Technology, no, 3777), GAPDH (D4C6R) mouse mAb (Cell Signaling Technology, no. 97166),TBK1/NAK antibody (Cell Signaling Technology, no. 3013), phospho-IRF3 (Ser396) (4D4G) rabbit mAb (Cell Signaling Technology, no. 4947), IRF3 (D6I4C) XP rabbit mAb (Cell Signaling Technology, no. 11904), phospho-Akt1 (Ser129) (D4P7F) rabbit mAb (Cell Signaling Technology, no. 13461), HER2/erbb2 (D8F12) XP rabbit (Cell Signaling Technology, no. 4290S), Akt1 (C73H10) rabbit mAb (Cell Signaling Technology, no. 2938), and c-myc rabbit antibody (Cell Signaling Technology, no. 9402) were used in this study.

Mapping of phenotypic data onto structures. Mapping and graphics of phenotypic data onto crystal structures were performed with the UCSF Chimera package[88]. Chimera is developed by the Resource for Biocomputing, Visualization and Informatics at the UCSF (supported by NIGMS P41-GM103311). Phenotypes were mapped using the 'define and render by attribute' functions of the program.

Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All previously unpublished saturation mutagenesis data generated or analyzed during this study are included in this published article (and its Extended Data information files). All screening data described in this project can be found in the Supplementary Information or have been deposited in the National Center for Biotechnology Information Short Read Archive (BioProject accession number PRJNA559517). Source data for Fig. 3a,c,e, Fig. 4, Fig. 5a–d,g–j, Fig. 6a,b,d–f, Fig. 7b,d–f, Extended Data Fig. 4a, Extended Data Fig. 6d,e,g and Extended Data Fig. 7c,e are available online. Aggregated mutational scanning data, including previously published datasets, are available from the corresponding author on reasonable request.

Code availability

All code used to analyze mutant screening data (ORFCall) can be accessed at https://github.com/tedsharpe/ORFCall.

References

88. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
89. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
90. Klipper-Aurbach, Y. et al. Mathematical formulae for the prediction of the residual beta cell function during the first two years of disease in children and adolescents with insulin-dependent diabetes mellitus. *Med. Hypotheses* **45**, 486–490 (1995).
91. Jafari, R. et al. The cellular thermal shift assay for evaluating drug target interactions in cells. *Nat. Protoc.* **9**, 2100–2122 (2014).
92. Haq, R. et al. Oncogenic BRAF regulates oxidative metabolism via PGC1α and MITF. *Cancer Cell* **23**, 302–315 (2013).

Acknowledgements

We thank A. Burgin, C. Painter, B. Tomson and M. Rees for discussions and critical reading of the manuscript.

Author contributions

N.S.P. and C.M.J. designed the study and wrote the paper. D.H., M.D.C., L.B., O.C., S.K., U.N., A.W., S.P., Y.L., J.C., M.S., C.Z., T.K.H., P.R. and P.P. contributed data for this project. X.Y, C.Z. and N.S.P. analyzed the data, T.S.M., D.A.B., R.H., F.P., D.E.R and C.M.J. supervised the work.

Competing interests

C.M.J. is a full-time employee of Novartis Institutes of Biomedical Research, Inc. T.M. is a full-time employee of 10x Genomics.

Additional information

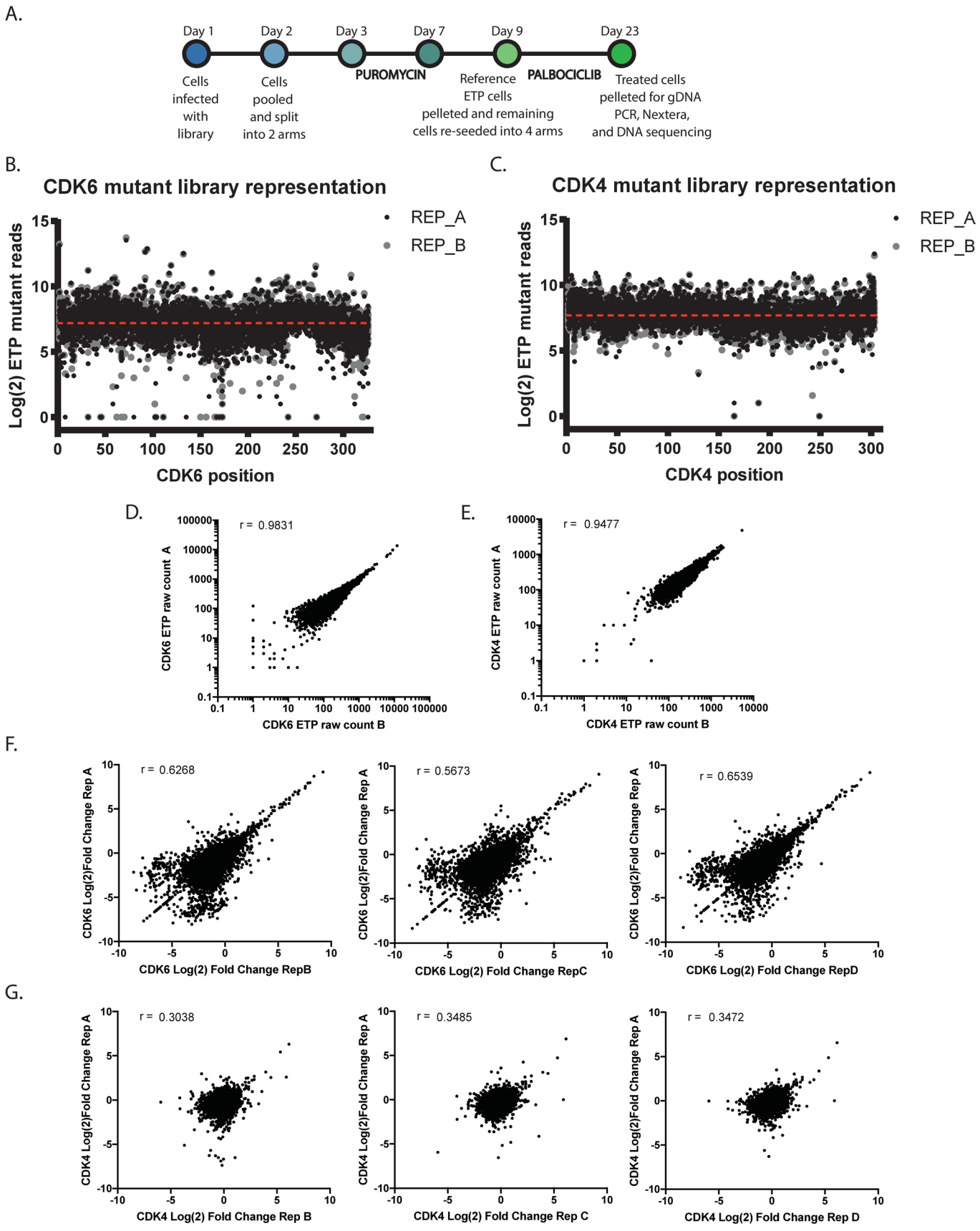
Extended data is available for this paper at https://doi.org/10.1038/s41594-019-0358-z.

Supplementary information is available for this paper at https://doi.org/10.1038/s41594-019-0358-z.

Correspondence and requests for materials should be addressed to C.M.J.

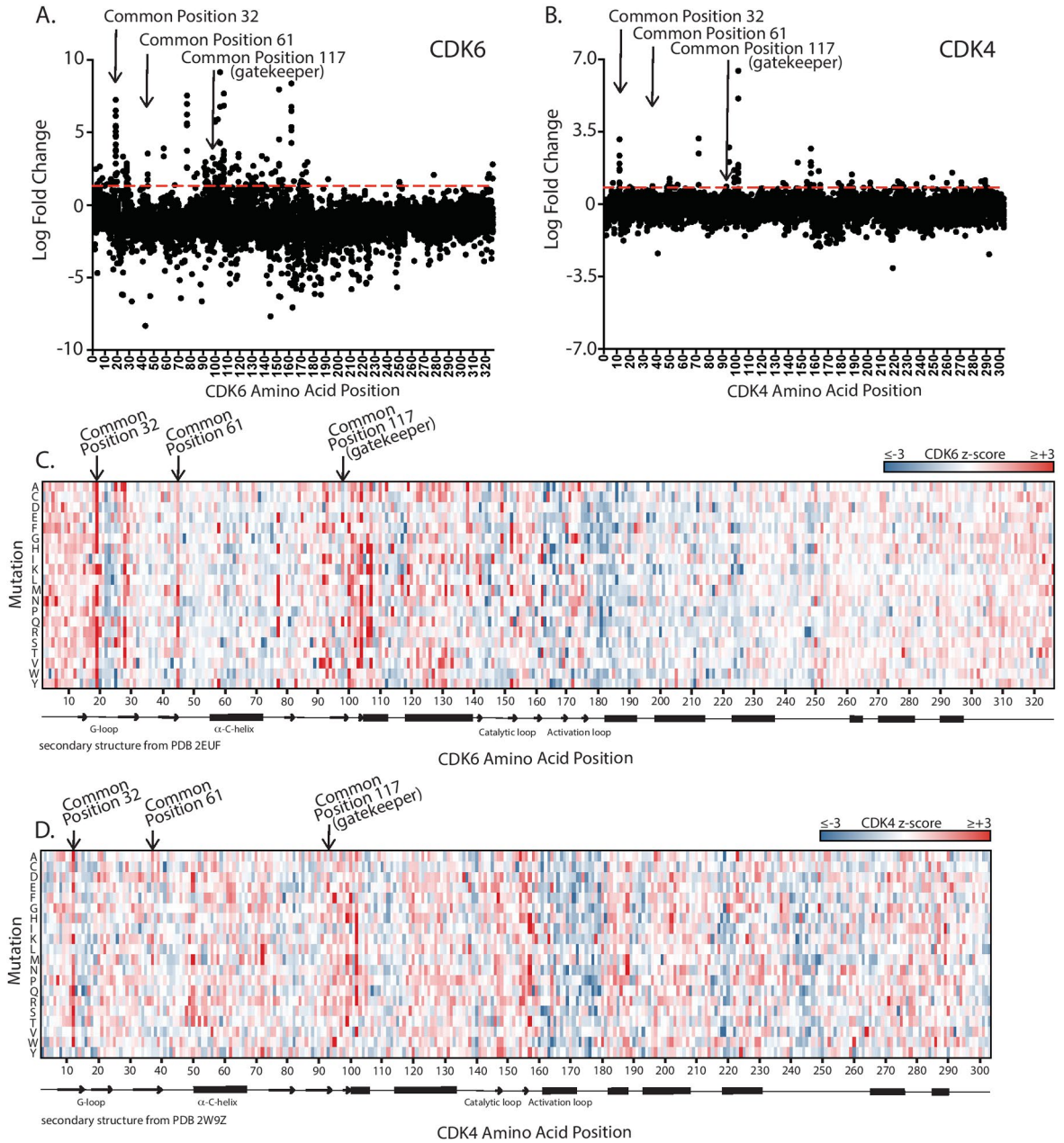
Peer review information Anke Sparmann was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.

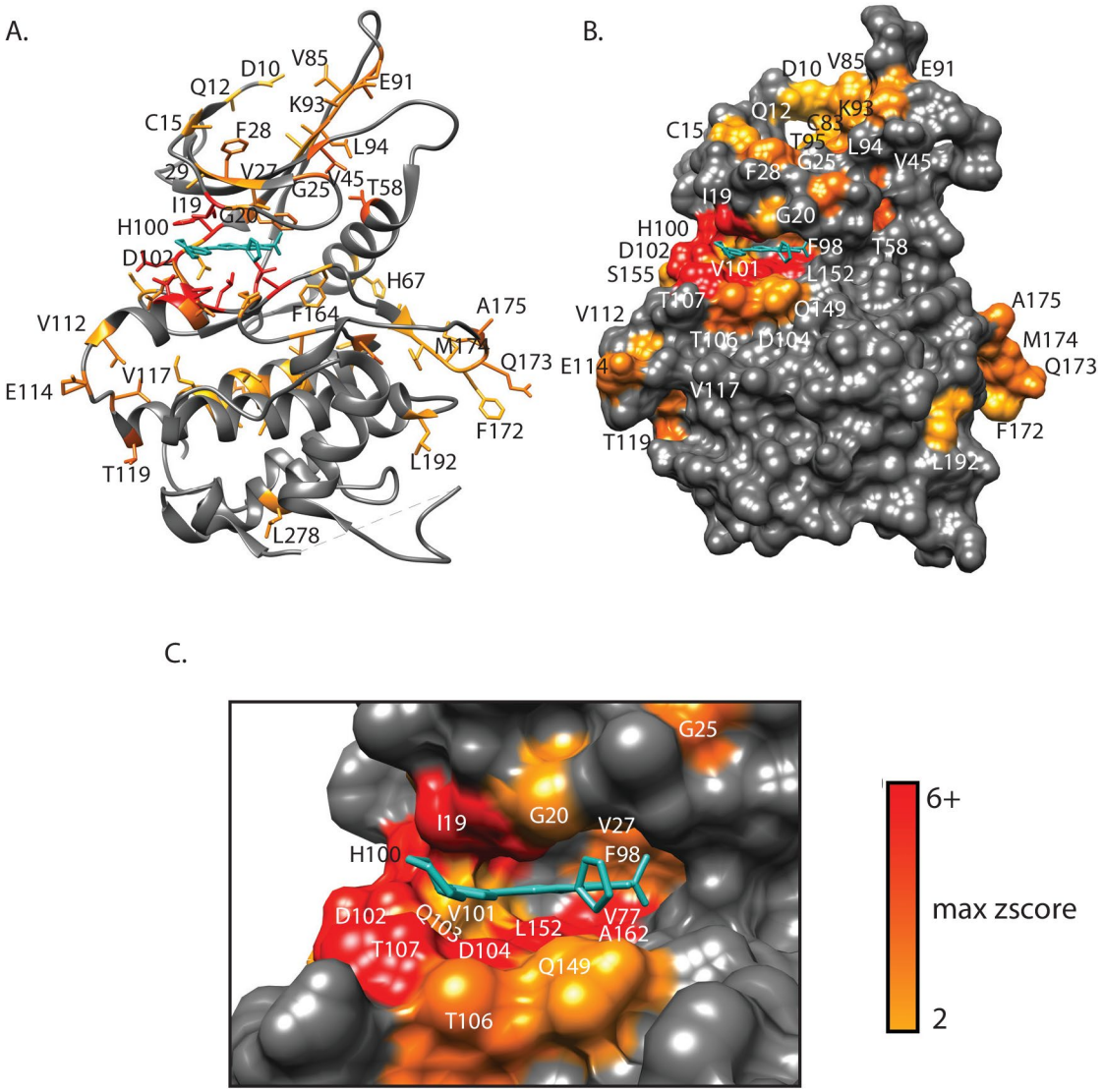


Extended Data Fig. 1 | See next page for caption.

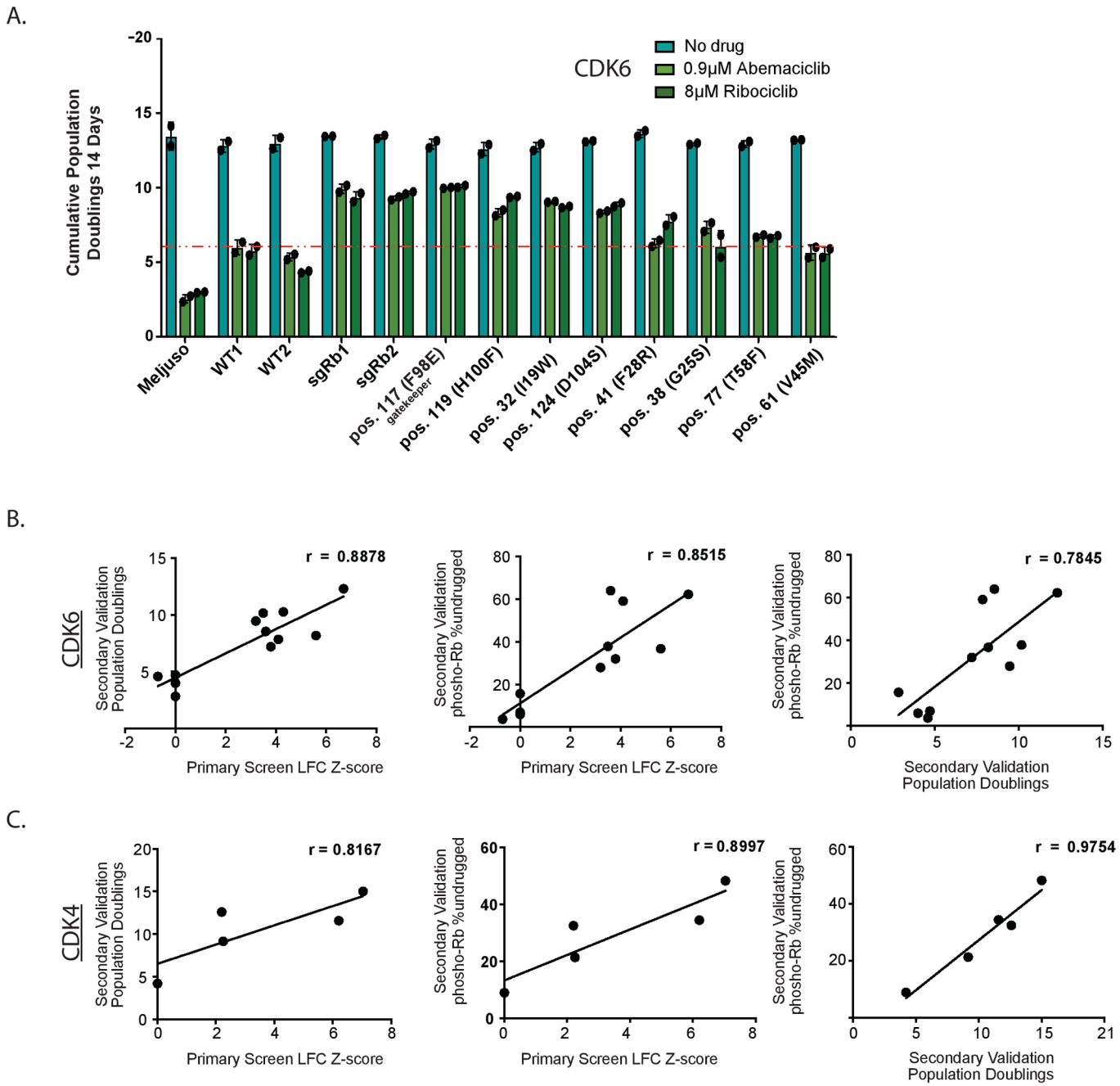
Extended Data Fig. 1 | Saturation mutagenesis of CDK4 and CDK6. a Summary schematic of CDK6 and CDK4 saturating mutagenesis screens. **b** CDK6 and **c** CDK4 early time point mutant representation illustrated across the length of each protein. Each dot represents one of two replicate measurements (**a and b**). **d** and **e** Early time point reference correlations for CDK6 and CDK4 mutants. Each dot represents one of two replicate measurements. The pearson correlation score is reported on each graph. **f** Log(2)-fold-changes for each of four CDK6 replicate measurements was individually calculated with respect to the average of two replicate measurements of early time point mutant representation and shown as black dots. Replicated correlations of Log(2)-Fold-Changes Replicate A versus all three other replicates are shown with their corresponding Pearson correlations. **g** Log(2)-fold-changes for each of four CDK4 replicate measurements was individually calculated with respect to the average of two replicate measurements of early time point mutant representation and shown as black dots. Replicated correlations of Log(2)-Fold-Changes Replicate A versus all three other replicates are shown with their corresponding Pearson correlations.



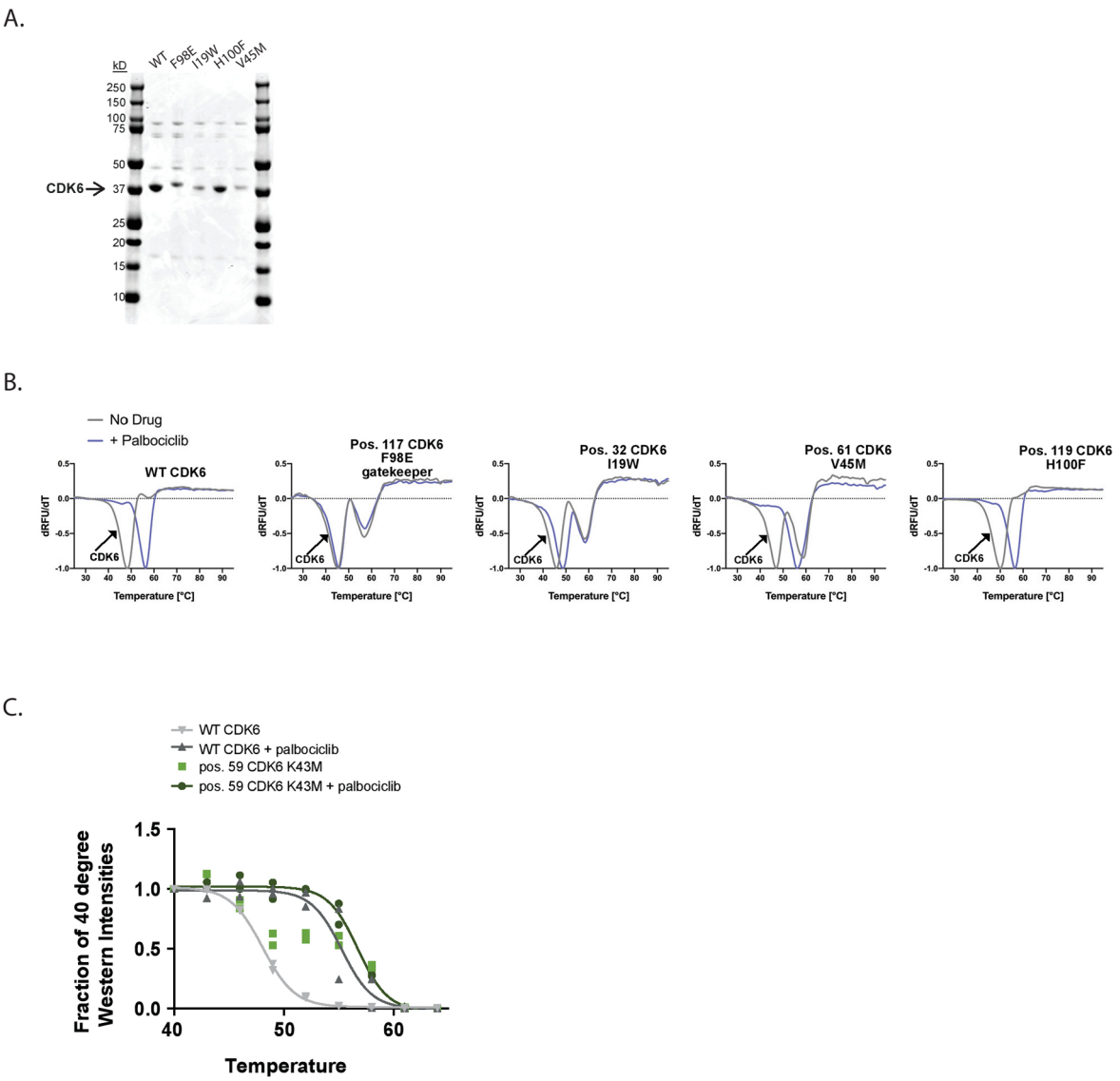
Extended Data Fig. 2 | Analysis of CDK4/6 saturation mutagenesis data. a and **b** Scatter plots of the average Log(2)-Fold-Changes (LFCs, y-axis, n=4 replicate measurements) and native amino acid position (x-axis) for the complete mutagenesis data for CDK6 (**a**) and CDK4 (**b**) palbociclib resistance screens. Red lines indicate z-scores of 1.95, our significance threshold for all saturation data (see Methods). Normalized data can be found in Supplementary Tables 1-2. **c** and **d** Heatmap visualization of the data presented in (**a**) and (**b**) for CDK6 (**c**) and CDK4 (**d**). Colors correspond to the z-score of the LFCs. Darker blue indicates a z-score of -3 or lower and darker red indicates a z-score of +3 or higher (more red = more palbociclib resistant).



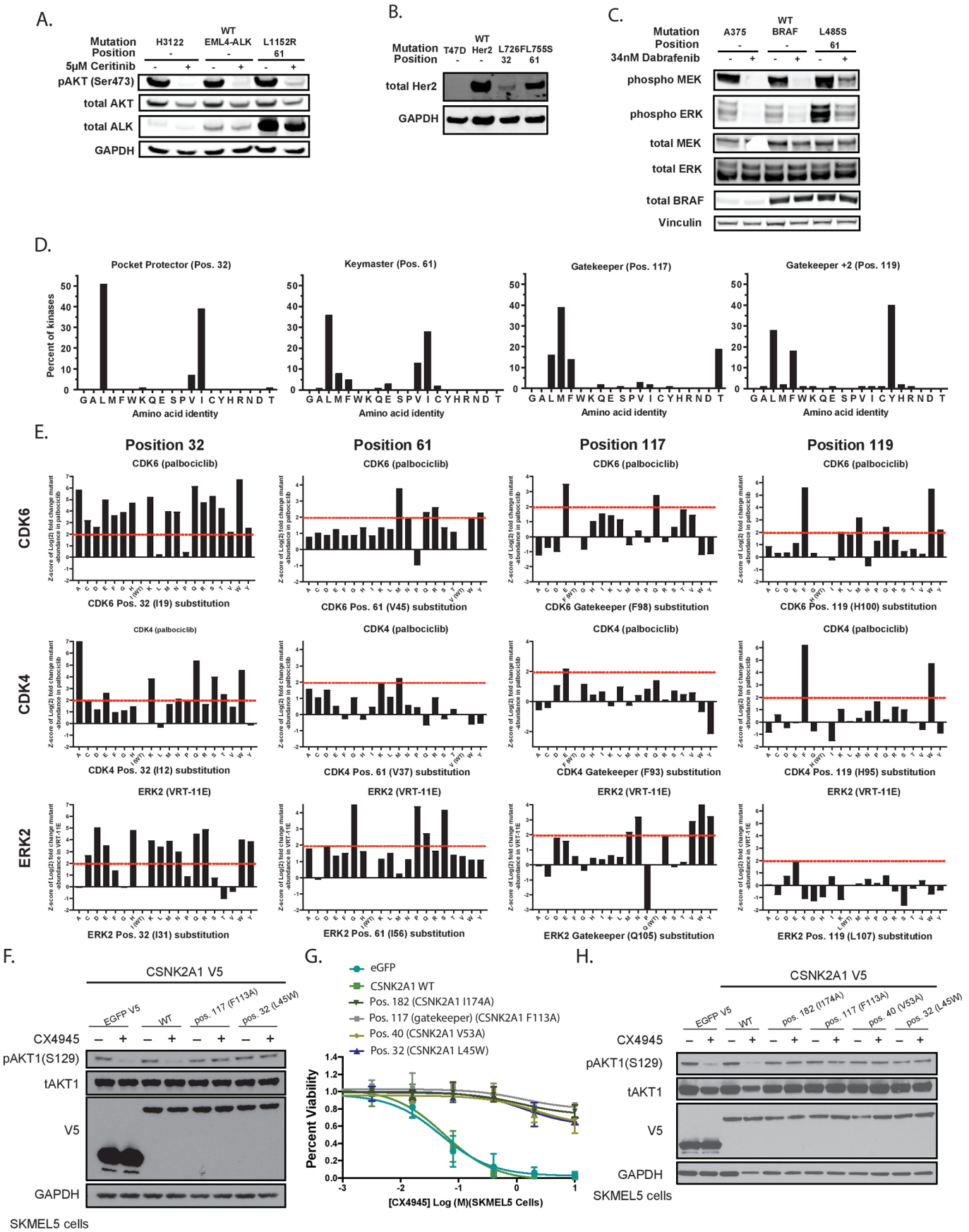
Extended Data Fig. 3 | Spatial projection of CDK4/6 saturation mutagenesis phenotypic data onto the 3D structure of CDK6. All data for the CDK6 palbociclib resistance screen was reduced to a single metric representing the maximum z-score observed at each position (average of n=4 independent replicates for each of n=19 substitutions). For positions that showed maximum z-scores over 1.95, the z-score was mapped onto the crystal structure of CDK6 bound to palbociclib (PDB 2EUF), either as a ribbon (**a**) or space-filling diagram (**b**). **c** View of the palbociclib binding pocket in CDK6 (PDB 2EUF). For **a**-**c**, light orange represents a z-score greater than 1.95 and red represents a z-score of 6 or higher.



Extended Data Fig. 4 | Validation of CDK4/6 saturation mutagenesis data. a Common drug resistant positions found in the CDK6 screen for palbociclib resistance are tested here for resistance to two other CDK4/CDK6 inhibitors: abemaciclib and ribociclib. Average cumulative population doublings after 14 days of Meljuso cells exogenously expressing wild-type (WT) or mutant CDK6 in the absence of drug are shown in turquoise. Treatment with 0.9 μM abemaciclib is shown in light green, and 8 μM ribociclib is shown in dark green, (n=2 independent experiments), error bars represent standard deviation. The correlation of resistance phenotypes in the CDK6 (**b**) or CDK4 (**c**) primary screens compared to population doublings in validation experiments, to phosphorylated Rb (compared to no drug) validation experiments, as well as the correlation of validation experiments to each other. Pearson correlations for each set are reported on the graph. Each dot represents a single measurement of the representative western blot shown in main figures.

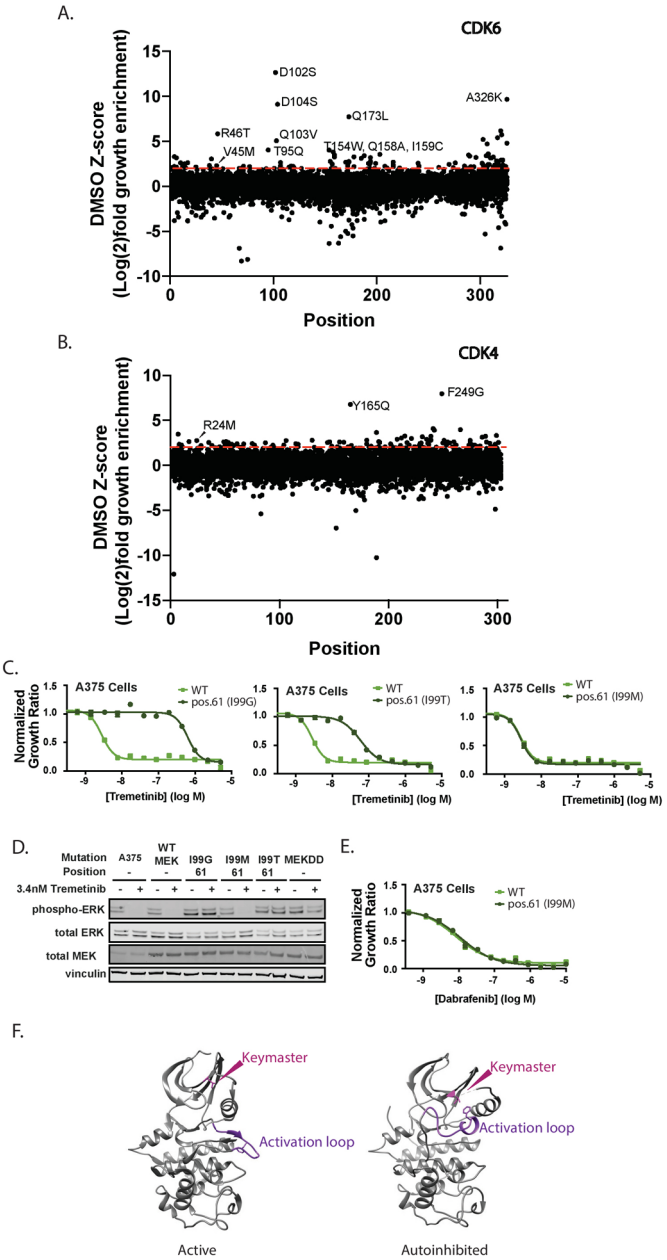


Extended Data Fig. 5 | Assessment of drug binding to CDK6 mutants using CETSA and DSF. a Coomasie stained acrylamide gel of purified FLAG-CDK6 mutants used for DSF assays. **b** DSF assay for investigating the interaction of the drug palbociclib with wild-type and mutant CDK6. The normalized first derivative melt profiles of purified wild-type CDK6, F98E, I19W, V45M, and H100F mutants in the absence (gray trace) or presence (blue trace) of the drug palbociclib. The black arrow indicates the melting profile of CDK6 alone in the absence of the drug. The temperature at which the derivative is the most negative (minimum) is used to calculate the melting temperature (T_m, DT_m), average of n = 5 replicate measurements. **c** CETSA assays of kinase dead CDK6 (K43M) in the presence or absence of palbociclib, average of n=2 replicate measurements.



Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Confirmation of generalizable kinase mutants beyond CDK4/6. a WT and mutant EML4-ALK drug sensitivity were measured by Western Blotting of downstream phosphorylation of AKT, representative image of n=2 individual experiments. **b** Levels of WT and mutant HER2 exogenously expressed in T47D cells were measured by western blotting, representative image of n=2 individual experiments. **c** Activity of BRAF position 61 (L485S) was measured by western blotting for downstream phosphorylation of MEK and ERK, representative image of n=2 individual experiments. **d** Bar graphs showing the percentage of all kinases harboring each amino acid at the wild-type residue of position 32, 61 117 and 119. All data is from the kinome alignment described in Kumar, R. D. & Bose, R. (2017). **e** Bar graphs illustrating the average resistance phenotypes (z-score) caused by all possible amino acid substitutions at the Pocket Protector, Keymaster, Gatekeeper and Gatekeeper +2 residues of CDK4 (n=4 replicate measurements), CDK6 (n=4 replicate measurements) and ERK2 (n=6 replicate measurements), based on CDK4/6 primary screening data and data from Brenan et al. (2016), n= **f**) Activity of CSNK2A1 mutants as assessed by western blot of downstream signaling, including phosphorylated AKT (pAKT1), representative image of n=2 individual experiments. **g** Proliferation assay assessing the sensitivity of SKMEL5 cells transduced with WT or mutant CSNK2A1 and treated with CX4945, plotted as the mean of 2 replicate measurements. **h** Activity of CSNK2A1 mutants as assessed by western blot of downstream signaling, including phosphorylated AKT (pAKT1) corresponding to samples in (**g**), representative image of n=2 individual experiments.



Extended Data Fig. 7 | Phenotypes of mutant CDK4/6 in standard growth conditions. Scatter plots of the z-scores of the Log(2)-Fold-Changes and native amino acid position for DMSO (vehicle only, n=2 replicate measurements) calculated relative to the early time point reference for CDK6 (**a**) or CDK4 (**b**), n=2 replicate measurements. **c** Proliferation assay assessing the sensitivity of A375 cells transduced with WT or mutant MEK1 treated with the MEK-inhibitor Trametinib (n=4 replicate measurements for each mutant at each concentration). **d** Activity of WT or mutant MEK1 as assessed via western blot analysis of downstream signaling, representative blot of n=2 individual experiments. **e** Dabrafenib complementation assays to measure activity of A375 cells transduced with WT or I99M mutant MEK1 (n=4 replicate measurements for each mutant at each concentration). **f** Illustration of active and auto-inhibited forms of kinases as illustrated by CDK2 structures (PDB: 2JGZ (active) and PDB: 1HCK (auto-inhibited). Activation loops are colored purple while keymaster positions are colored magenta.

nature research

Corresponding author(s): Cory Johannessen

Last updated by author(s): Nov 15, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Capture of western blot images, ImageStudio (v 2.1.10).
Data analysis	The following commercially and publicly available software was used for data analysis or collections: Mutational calling for mutagenesis screens, ORFCall (v1.0); protein alignments, Clustal Omega (v 1.2.4), MUSCLE (v 3.8) and the UCSF Chimera package (v 1.13.1); plotting and curve fitting of drug response curves and CETSA, PRISM Graphpad (v 7.0c); capture of western blot images, ImageStudio (v 2.1.10); generation of logo plots, Seq2Logo (v 2.0); statistical analysis of aggregated phenotypic data, R (v 3.5.2, 2018-12-20), R Packages used, RobustRankAggreg (v 1.1), stats (base, v 3.5.2).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All previously unpublished saturation mutagenesis data generated or analyzed during this study are included in this published article (and its Extended Data information files). All screening data described in this project can be found in Supplementary Information or has been deposited in the NCBI Short Read Archive (BioProject accession number PRJNA559517). Aggregated mutational scanning data, including previously published data sets, are available from the corresponding author upon reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Replicate numbers (n=3) in primary screens were determined from down sampling prior validated datasets to ensure the we exceeded the minimum number of samples needed (n=2).
Data exclusions	No data was excluded.
Replication	We individually validated mutants identified in the primary screening data, as well as validated predictions made on other proteins not included in the primary data/analysis.
Randomization	Cells expressing mutant cDNA libraries were randomized into early time point, DMSO treatment or drug treatment arms.
Blinding	Blinding was not relevant to this study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☐	☐ Palaeontology
☐	☐ Animals and other organisms
☐	☐ Human research participants
☐	☐ Clinical data

Methods

n/a	Involved in the study
☐	☐ ChIP-seq
☐	☐ Flow cytometry
☐	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Phospho-Rb (Ser780) (D59B7) Rabbit mAb (Cell Signaling Technology, #8180S), Rb (4H1) Mouse mAb (Cell Signaling Technology, #9309S), Cdk6 (C-21) rabbit polyclonal IgG (Santa Cruz Biotechnology, sc-177), Cdk4 (DCS-35) mouse monoclonal IgG1 (Santa Cruz Biotechnology, sc-23896), beta-Actin (8H10D10) Mouse mAb (Cell Signaling Technology, #3700S), Monoclonal Anti-vinculin antibody (Sigma-Aldrich, V9131), GAPDH (14C10) Rabbit mAb (Cell Signaling Technology, #2118S), ALK (31F12) Mouse mAb (Cell Signaling Technology, #3791S), Phospho-ALK (Tyr1278/1282/1283) Antibody (Cell Signaling Technology, #3983S), Akt (pan) (40D4) Mouse mAb (Cell Signaling Technology, #2920), Phospho-Akt (Ser473) (D9E) XP® Rabbit mAb (Cell Signaling Technology, #4060L), p44/42 MAPK (Erk1/2) (L34F12) Mouse mAb (Cell Signaling Technology, #4696S), Phospho-p44/42 MAPK (Erk1/2) (Thr202/Tyr204) Antibody (Cell Signaling Technology, #9101S), Cofilin (D3F9) XP® Rabbit mAb (Cell Signaling Technology, #5175S), Raf-B Antibody (F-7) Mouse Monoclonal IgG (Santa Cruz Biotechnology, sc-5284), MEK1/2 Rabbit mAB (Cell Signaling Technology, #9122L), Phospho-MEK1/2 (Ser217/221) (41G9) Rabbit mAb, (Cell Signaling Technology, #9154), RSK1 (D6D5) Rabbit mAb (Cell Signaling Technology, #8408), Phospho-p90RSK (Ser380) Antibody (Cell Signaling Technology, #9341), EGF Receptor (1F4) Mouse mAb (Cell Signaling Technology, #2239), Phospho-EGF Receptor (Tyr1068) (D7A5) XP® Rabbit mAb (Cell Signaling Technology, #3777), GAPDH (D4C6R) Mouse mAb (Cell Signaling Technology, #97166),TBK1/NAK Antibody (Cell Signaling Technology, #3013), Phospho-IRF-3 (Ser396) (4D4G) Rabbit mAb (Cell Signaling Technology, #4947), IRF-3 (D6I4C) XP® Rabbit mAb (Cell Signaling Technology, #11904), Phospho-Akt1 (Ser129) (D4P7F) Rabbit mAb (Cell Signaling Technology, #13461), HER2/erbb2 (D8F12) XP rabbit (Cell Signaling Technology, #4290S) , Akt1 (C73H10) Rabbit mAb (Cell Signaling Technology, #2938), and c-myc Rabbit Antibody (Cell Signaling Technology, #9402) were used in this study.

Validation

Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	All cell lines were acquired via the Broad Institute Cancer Cell Line Encyclopedia, via ATCC
Authentication	Every cell line was authenticated within 6 months of this submission via SNP-based profiling and alignment to reference profiles.
Mycoplasma contamination	All cell lines used in this study were assessed for mycoplasma contamination at the time of authentication.
Commonly misidentified lines (See ICLAC register)	PC9 cells are identifiable as both PC9 and PC14.

Palaeontology

Specimen provenance	*Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).*
Specimen deposition	*Indicate where the specimens have been deposited to permit free access by other researchers.*
Dating methods	*If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.*

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	*For laboratory animals, report species, strain, sex and age OR state that the study did not involve laboratory animals.*
Wild animals	*Provide details on animals observed in or captured in the field; report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.*
Field-collected samples	*For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.*
Ethics oversight	*Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, gender, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*
Recruitment	*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*
Ethics oversight	*Identify the organization(s) that approved the study protocol.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	*Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.*
Study protocol	*Note where the full trial protocol can be accessed OR if not available, explain why.*
Data collection	*Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.*

Outcomes — *Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.*

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication.
For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.

Files in database submission
Provide a list of all files available in the database submission.

Genome browser session
(e.g. UCSC)
Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.

Methodology

Replicates
Describe the experimental replicates, specifying number, type and replicate agreement.

Sequencing depth
Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.

Antibodies
Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.

Peak calling parameters
Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.

Data quality
Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.

Software
Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation
Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.

Instrument
Identify the instrument used for data collection, specifying make and model number.

Software
Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.

Cell population abundance
Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.

Gating strategy
Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type	*Indicate task or resting state; event-related or block design.*
Design specifications	*Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.*
Behavioral performance measures	*State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).*

Acquisition

Imaging type(s)	*Specify: functional, structural, diffusion, perfusion.*
Field strength	*Specify in Tesla*
Sequence & imaging parameters	*Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.*
Area of acquisition	*State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.*

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software	*Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*
Normalization	*If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.*
Normalization template	*Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.*
Noise and artifact removal	*Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).*
Volume censoring	*Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.*

Statistical modeling & inference

Model type and settings	*Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).*
Effect(s) tested	*Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.*

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference (See Eklund et al. 2016)	*Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.*
Correction	*Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).*

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity	*Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).*
Graph analysis	*Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).*

Multivariate modeling and predictive analysis

Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.