

# High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response

Hui Gao<sup>1,7</sup>, Joshua M Korn<sup>1,7</sup>, Stéphane Ferretti<sup>2,7</sup>, John E Monahan<sup>3</sup>, Youzhen Wang<sup>4</sup>, Mallika Singh<sup>5,6</sup>, Chao Zhang<sup>4,6</sup>, Christian Schnell<sup>2</sup>, Guizhi Yang<sup>1</sup>, Yun Zhang<sup>4</sup>, O Alejandro Balbin<sup>1</sup>, Stéphanie Barbe<sup>2</sup>, Hongbo Cai<sup>1</sup>, Fergal Casey<sup>5</sup>, Susmita Chatterjee<sup>5</sup>, Derek Y Chiang<sup>1</sup>, Shannon Chuai<sup>4</sup>, Shawn M Cogan<sup>1</sup>, Scott D Collins<sup>1</sup>, Ernesta Dammassa<sup>2</sup>, Nicolas Ebel<sup>2</sup>, Millicent Embry<sup>5</sup>, John Green<sup>1</sup>, Audrey Kauffmann<sup>2</sup>, Colleen Kowal<sup>1</sup>, Rebecca J Leary<sup>1</sup>, Joseph Lehar<sup>3</sup>, Ying Liang<sup>4</sup>, Alice Loo<sup>1</sup>, Edward Lorenzana<sup>5</sup>, E Robert McDonald III<sup>1</sup>, Margaret E McLaughlin<sup>3</sup>, Jason Merkin<sup>1</sup>, Ronald Meyer<sup>3</sup>, Tara L Naylor<sup>3</sup>, Montesa Patawaran<sup>5</sup>, Anupama Reddy<sup>3,6</sup>, Claudia Röelli<sup>2</sup>, David A Ruddy<sup>3</sup>, Fernando Salangang<sup>5</sup>, Francesca Santacrose<sup>2</sup>, Angad P Singh<sup>1</sup>, Yan Tang<sup>5</sup>, Walter Tinetto<sup>2</sup>, Sonja Tobler<sup>2</sup>, Roberto Velazquez<sup>1</sup>, Kavitha Venkatesan<sup>1</sup>, Fabian Von Arx<sup>2</sup>, Hui Qin Wang<sup>3</sup>, Zongyao Wang<sup>4</sup>, Marion Wiesmann<sup>2</sup>, Daniel Wyss<sup>2</sup>, Fiona Xu<sup>4</sup>, Hans Bitter<sup>1</sup>, Peter Atadja<sup>4</sup>, Emma Lees<sup>5</sup>, Francesco Hofmann<sup>2</sup>, En Li<sup>4</sup>, Nicholas Keen<sup>1</sup>, Robert Cozens<sup>2</sup>, Michael Rugaard Jensen<sup>2</sup>, Nancy K Pryer<sup>5,6</sup>, Juliet A Williams<sup>1,8</sup> & William R Sellers<sup>1,8</sup>

**Profiling candidate therapeutics with limited cancer models during preclinical development hinders predictions of clinical efficacy and identifying factors that underlie heterogeneous patient responses for patient-selection strategies. We established ~1,000 patient-derived tumor xenograft models (PDXs) with a diverse set of driver mutations. With these PDXs, we performed *in vivo* compound screens using a 1 × 1 × 1 experimental design (PDX clinical trial or PCT) to assess the population responses to 62 treatments across six indications. We demonstrate both the reproducibility and the clinical translatability of this approach by identifying associations between a genotype and drug response, and established mechanisms of resistance. In addition, our results suggest that PCTs may represent a more accurate approach than cell line models for assessing the clinical potential of some therapeutic modalities. We therefore propose that this experimental paradigm could potentially improve preclinical evaluation of treatment modalities and enhance our ability to predict clinical trial responses.**

Roughly 85% of preclinical agents entering oncology clinical trials fail to demonstrate sufficient safety or efficacy to gain regulatory approval<sup>1–3</sup>. This high failure rate highlights a weak understanding of the complexity of human cancer, the continued limitations of the predictive value of existing preclinical models and the scale at which cancer models are interrogated in the preclinical setting<sup>4</sup>. Hence, there is a need for experimental systems that better replicate the diversity of human tumor biology in a preclinical setting.

Increasing amounts of evidence have suggested that PDXs faithfully recapitulate human tumor biology and predict patient drug response (reviewed by Tentler *et al.*<sup>5</sup>, Siolas and Hannon<sup>6</sup>, Rosfjord *et al.*<sup>7</sup> and Hidalgo *et al.*<sup>8</sup>) by directly comparing drug responses in patients and their corresponding xenografts. However, these studies have limited value in predicting potential clinical-trial response at the population level, owing to the use of a limited number of PDXs. Bertotti *et al.*<sup>9</sup> first reported testing the epidermal growth factor receptor (EGFR)

inhibitor cetuximab across a set of colorectal PDX models, and found concordance in the response of *EGFR*-amplified models and colorectal cancer (CRC) patients in the clinic to cetuximab. To extend such observations to a greater number of human cancers, we have generated an extensive collection of PDXs containing ~1,000 models, all characterized for their mutations, copy-number alterations and mRNA expression levels. We adopted the approach laid out by Migliardi *et al.*<sup>10</sup>, using these PDXs to perform a large-scale *in vivo* screen to model inter-patient response heterogeneity with a ‘one animal per model per treatment’ approach (1 × 1 × 1). By correlating genomic information with observed efficacy, we successfully validated genetic hypotheses and biomarkers derived from *in vitro* model systems, and identified novel therapies that cell line model systems have failed to capture. Furthermore, we obtained notably similar results when comparing the available clinical data with the response of the PDXs. Finally, deep-sequencing analysis of melanoma-resistant tumors from

<sup>1</sup>Oncology Disease Area, Novartis Institutes for Biomedical Research, Cambridge, Massachusetts, USA. <sup>2</sup>Oncology Disease Area, Novartis Institutes for Biomedical Research, Basel, Switzerland. <sup>3</sup>Department of Oncology Translational Medicine, Novartis Institutes for Biomedical Research, Cambridge, Massachusetts, USA.

<sup>4</sup>China Novartis Institutes for Biomedical Research, Shanghai, China. <sup>5</sup>Oncology Disease Area, Novartis Institutes for Biomedical Research, Emeryville, California, USA.

<sup>6</sup>Present addresses: Patrus Therapeutics, Inc., San Francisco, California, USA (M.S.); Janssen China R&D and Scientific Affairs, Shanghai, China (C.Z.); BioMarin Pharmaceutical, Inc., Novato, California, USA (N.K.P.); and Duke University, Durham, North Carolina, USA (A.R.). <sup>7</sup>These authors contributed equally to this work.

<sup>8</sup>These authors jointly directed this work. Correspondence should be addressed to J.A.W. ([juliet.williams@novartis.com](mailto:juliet.williams@novartis.com)).

Received 19 December 2014; accepted 26 August 2015; published online 19 October 2015; doi:10.1038/nm.3954

the PCT revealed mechanisms of resistance similar to those reported in patients. Together, these data demonstrate, retrospectively, the enhanced translatability of this *in vivo* experimental paradigm, and set the foundation for the use of this population-based approach for the potential prediction of human clinical trial responses.

## RESULTS

### The Novartis Institutes for BioMedical Research PDX encyclopedia (NIBR PDXE)

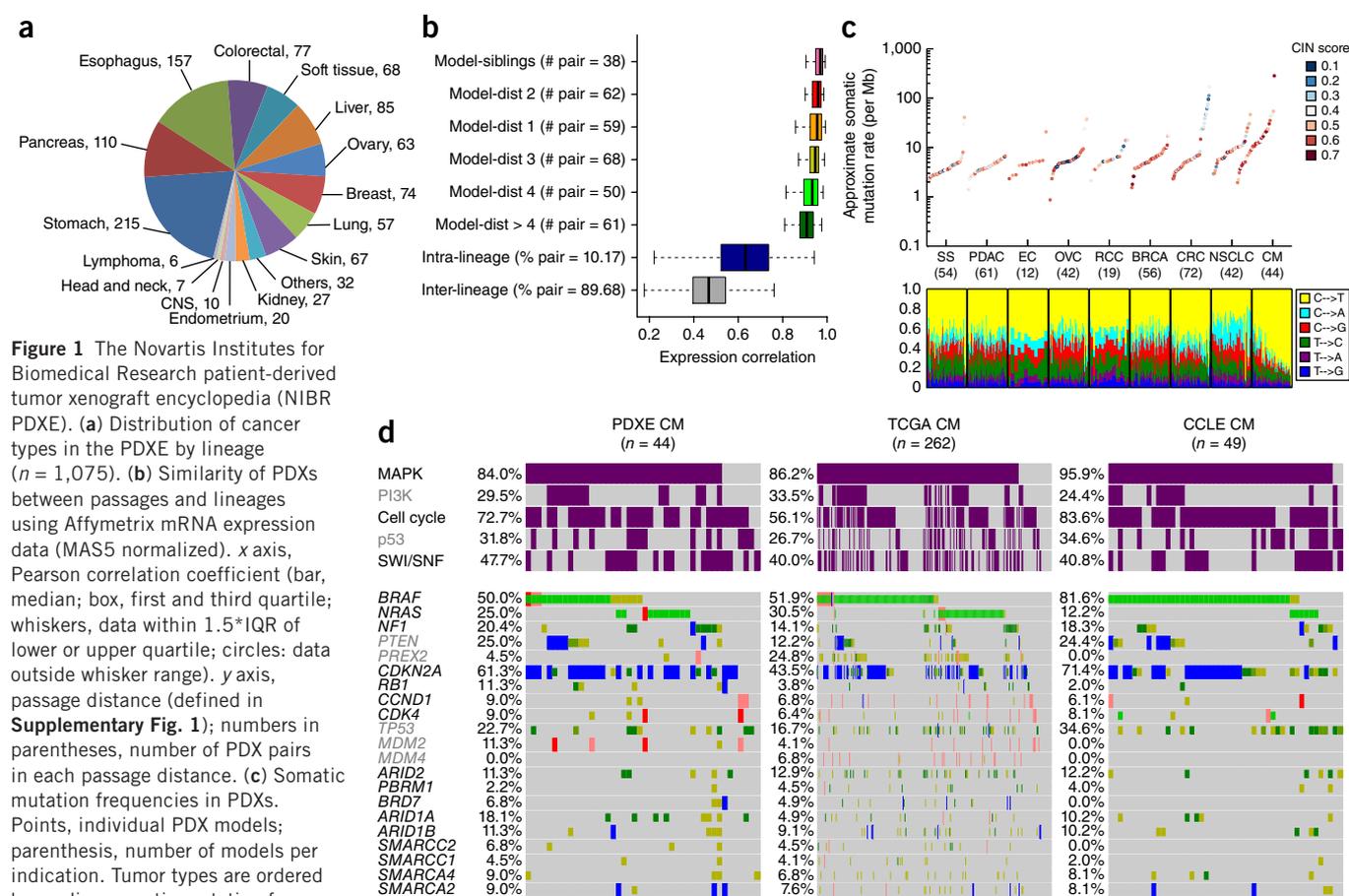
A total of 1,075 PDX models across the spectrum of common adult solid cancers were established (Fig. 1a). The identity of established PDXs was confirmed using single-nucleotide polymorphism (SNP) fingerprint analysis<sup>11</sup>, and the histology was compared to each patient's clinical diagnosis. Mutations, copy-number alterations and mRNA expression information for models used in the PCT study can be found with this article (Supplementary Table 1).

Minimum genetic drift of PDX models occurs after serial passages in mice<sup>12–19</sup>. We explored the effects of this minimum genetic drift on phenotypic stability of models by comparing Affymetrix transcriptome expression data obtained from 672 samples derived from multiple passages of 47 models of various lineages (Supplementary Fig. 1). Samples from the same model showed strong similarity,

with the highest correlation of expression values between xenograft tumors derived from the same parent (siblings) (Fig. 1b), suggesting minimum drift of a given PDX model between passages. Of note, there was a slightly decreased correlation with increased passage distance (Fig. 1b). To minimize the potential variability of drug response resulting from phenotypic drift, all PCT studies were conducted using models between passage (p)4 and p10.

### Somatic mutation frequency and genomic-landscape analysis

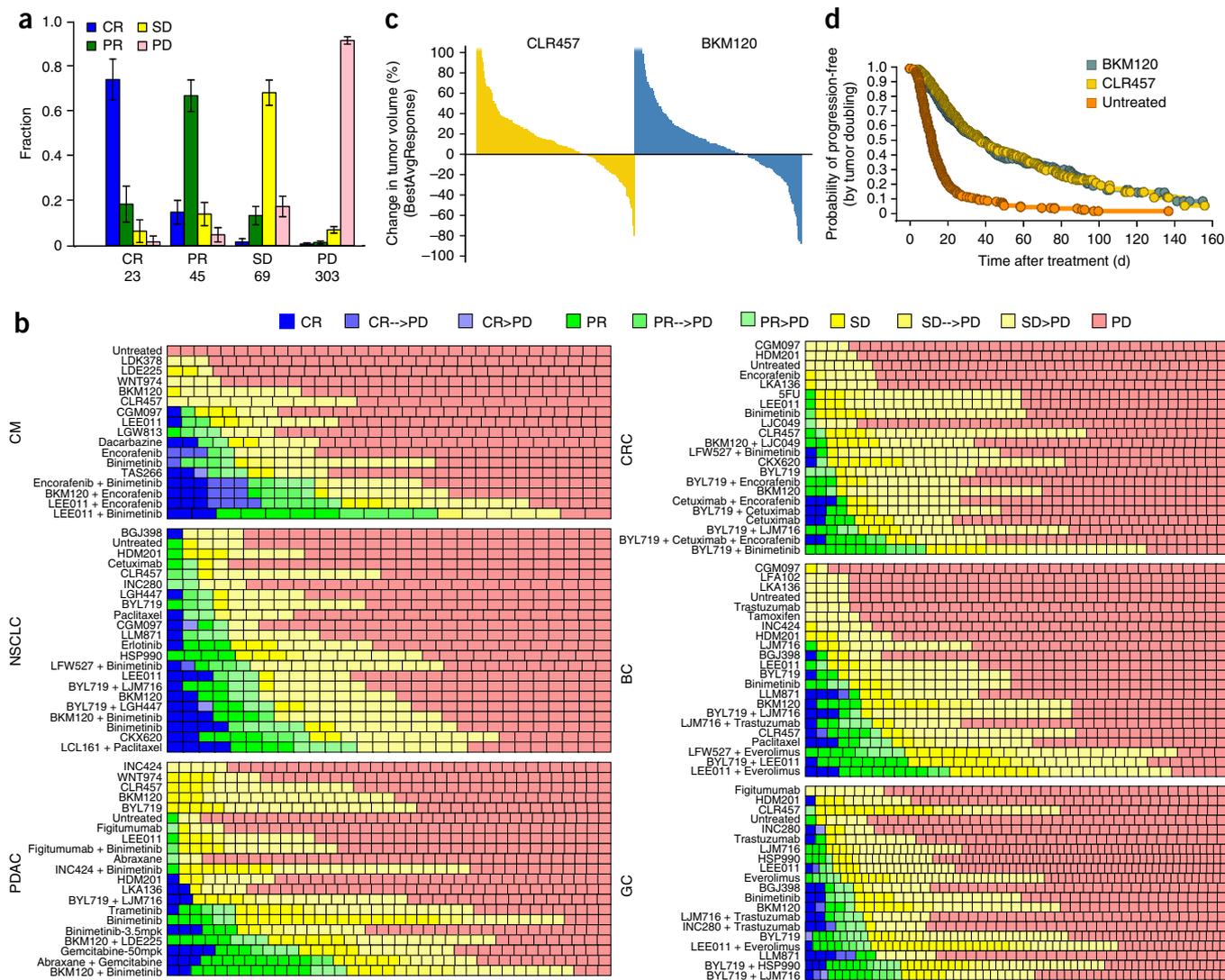
The mutation frequency, mutation-type distribution and chromosomal instability<sup>20</sup> of PDXs, cell lines and patient tumors in The Cancer Genome Atlas (TCGA) were characterized across indications. Despite a lack of matched normal samples for models from the NIBR PDXE and Cancer Cell Line Encyclopedia (CCLE), we observed a wide variation in mutation rates both between and within lineages from their respective cohorts, and an enrichment of particular mutational patterns in certain indications—notably C→A in lung and C→T in melanoma (Fig. 1c and Supplementary Fig. 2). The median and s.d. of the mutation rate per indication was highly correlated between samples from CCLE, PDXE and TCGA, and the correlation was highest between the PDXs and patient tumors ( $R = 0.94$  and  $R = 0.96$  for median and s.d., respectively; between cell lines and tumors,  $R = 0.51$



**Figure 1** The Novartis Institutes for Biomedical Research patient-derived tumor xenograft encyclopedia (NIBR PDXE). (a) Distribution of cancer types in the PDXE by lineage ( $n = 1,075$ ). (b) Similarity of PDXs between passages and lineages using Affymetrix mRNA expression data (MAS5 normalized). x axis, Pearson correlation coefficient (bar, median; box, first and third quartile; whiskers, data within  $1.5 \times \text{IQR}$  of lower or upper quartile; circles: data outside whisker range). y axis, passage distance (defined in Supplementary Fig. 1); numbers in parentheses, number of PDX pairs in each passage distance. (c) Somatic mutation frequencies in PDXs. Points, individual PDX models; parenthesis, number of models per indication; blue, homozygous deletions; salmon, amplification  $>5$  copies; red, amplification  $>8$  copies; light green, known COSMIC (Catalog of Somatic Mutations in Cancer) gain-of-function mutations; dark green, truncating mutations/frameshift or known COSMIC loss-of-function; mustard, novel mutation; purple, pathway altered in at least one gene; gene names colored black or gray to indicate inclusion in same-colored pathway listed above; percentages indicate percentage of samples altered for the given gene or pathway.

and  $R = 0.72$ ; between cell lines and PDXs,  $R = 0.53$  and  $R = 0.60$ , **Supplementary Fig. 3**). The median mutation rate showed a weaker correlation between cell lines and patient tumors, primarily because of a lack of hypermutators in melanoma cell line models. Specifically, 75/121 (62%) of TCGA samples have a mutation rate higher than 10 per megabase, compared to 23/44 (52%) of PDXE models and 13/49 (27%) of CCLE cell lines. All other lineages have a higher median mutation rate in cell lines than in primary tumors, possibly because of bias introduced by the amenability of tumors to growing on plastic or because post-somatic mutations accrue in cell lines during passaging. PDXE samples have a much smaller increased mutation rate (probably because of residual unfiltered germline variants) and s.d., which is consistent with less divergence from the patient tumors.

The mutational and copy-number landscape of *in vivo* and *in vitro* models was compared at both the gene and pathway levels to clinical samples for five indications (**Fig. 1d**; **Supplementary Figs. 4a–d** and **5a–e**). The frequency of genetic alterations across the three data sets was remarkably consistent; however, we see greater correspondence between PDX models and clinical samples than between the CCLE models and clinical samples. Several mutations are consistently not captured in the PDXs and cell lines (for example, *MDM4* (encoding MDM4, p53 regulator) amplification in breast cancer (BRCA) patients and *PHGDH* (encoding phosphoglycerate dehydrogenase) amplification in pancreatic ductal adenocarcinoma (PDAC) patients). It is not yet clear whether the absence of these genes in PDXs and cell lines is due to a potential selection bias toward *in vitro* and *in vivo* establishment



**Figure 2** Systematic approach for *in vivo* compound profiling using PDXs (PCT), and its reproducibility. **(a)** Feasibility assessment of  $1 \times 1 \times 1$  PCT approach by Pearson correlation analysis.  $x$  axis, number of majority response from each response category;  $y$  axis, fraction of individual animal response relative to the majority (average  $\pm$  s.e.m.). A total of 2,138 single-animal response data were collected and categorized from 440 unique treatment models (Online Methods). CR: complete response; PR: partial response; SD: stable disease; PD: progressive disease. **(b)** Summary of compound sensitivity in the PCTs. The BestAvgResponse was used to make response calls (Online Methods), and each square represents a PDX. A total of 62 treatment groups were tested in 277 PDXs across six indications (BRCA (breast cancer,  $n = 43$ ), CM (cutaneous melanoma,  $n = 33$ ), CRC (colorectal carcinoma,  $n = 59$ ), GC (gastric cancer,  $n = 64$ ), NSCLC (non-small cell lung carcinoma,  $n = 36$ ) and PDAC (pancreatic ductal adenocarcinoma,  $n = 42$ )). Arrow (CR $\rightarrow$ PD, PR $\rightarrow$ PD, SD $\rightarrow$ PD, and CR $\rightarrow$ PD, PR $\rightarrow$ PD, SD $\rightarrow$ PD) indicates disease progression; > indicates progression seen after 64 d; > indicates progression in <64 d. **(c)** Waterfall plot of responses to the PI3K inhibitors CLR457 ( $n = 205$ ) and BKM120 ( $n = 213$ ) across all indications; each bar represents an individual PDX. **(d)** Kaplan-Meier progression-free survival curve of PDXs treated with CLR457 ( $n = 205$ ) and BKM120 ( $n = 213$ ) across all indications.

of cells with these mutations, or whether it is simply a result of the relatively smaller sample size in the PDXE and CCLE pools.

### Feasibility assessment of the PCT (1 × 1 × 1) experimental design

The PCT response criteria were defined by modifying Response Evaluation Criteria In Solid Tumors (RECIST) criteria<sup>21</sup> (mRECIST, see Online Methods). We evaluated the reproducibility of 2,138 single-animal response data from 440 unique treatment models taken from historical experiments. On the basis of the majority response of single animals in each treatment model, the models were categorized into four response categories: 23 complete responses (mCR), 45 partial responses (mPR), 69 stable diseases (mSD) and 303 progressive diseases (mPD). For each ‘majority response’ category, at least 66% of the individual responses matched the majority response, and fewer than 10% of individual responses were off by more than one mRECIST category (Fig. 2a). Furthermore, when we combined the response categories (mCR, mPR and mSD) into a single ‘responder’ category, the response calls made on a single mouse were consistent with the majority response 95% of the time, which strongly supports the rationale of using one animal to reflect the true response. These data justify the 1 × 1 × 1 experimental approach, facilitating efficient robust assessment of drug response across a large number of models to determine population-based response rates.

### The PCT approach for *in vivo* compound profiling

Similarly to what a cancer patient experiences in the clinic, each mouse received one treatment in a given trial. In addition, each tumor was enrolled in multiple treatment trials, enabling the assessment of inter-patient response heterogeneity across multiple disease lineages. A total of 38 unique therapeutic entities were screened as either a single treatment (36 unique entities) or in combination (26 unique entities), and model responses were categorized (Fig. 2b and Supplementary Table 1).

The aggregate trial reproducibility of the PCT approach was assessed by comparing the responses of compounds that modulate the same molecular target. Two structurally distinct pan-PI3K inhibitors, BKM120 and CLR457, were tested in >200 PDXs across all six test indications. Strikingly, at the population level, both compounds

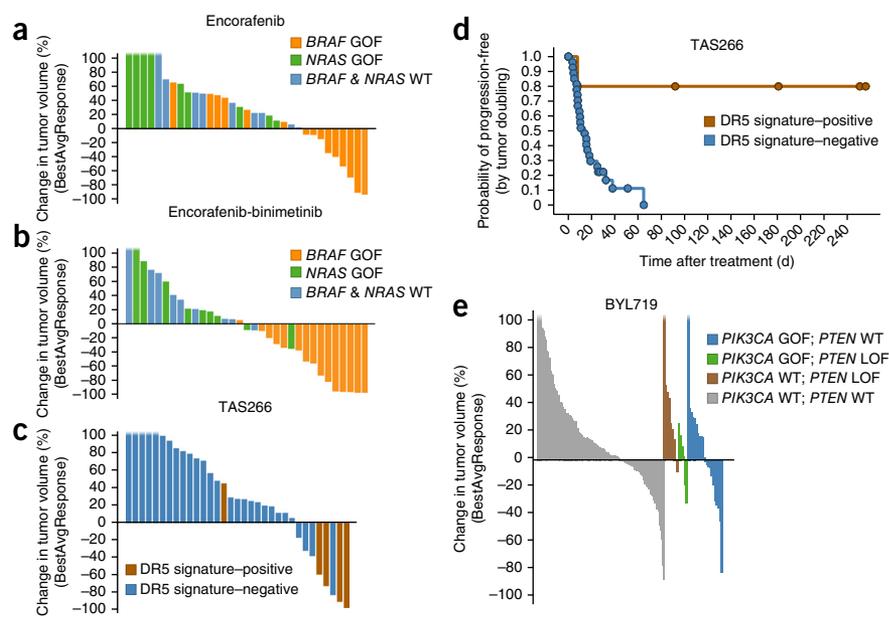
had a highly similar response rate (mCR + mPR + mSD): 48% for BKM120 and 54% for CLR457 (Fig. 2c). At the individual patient level, response rates were also significant (Supplementary Fig. 6;  $R = 0.56$ ). We also found increased progression-free survival (PFS, defined here for PDX as time until tumor volume reaches 200% of baseline) compared to the PFS of untreated controls (BKM120, hazard ratio 0.28, 95% interval 0.23–0.35,  $P < 1 \times 10^{-16}$ ; CLR457, hazard ratio 0.28, 95% interval 0.23–0.35,  $P < 1 \times 10^{-16}$ ) (Fig. 2d).

In addition, two combinatorial treatments—LEE011 (an inhibitor of cell cycle regulatory kinases CDK4/6)-everolimus (mTOR inhibitor); and LCL161 (IAP inhibitor)-paclitaxel—had comparable response rates and median survivals in repeated trials when assessed at the population level (although the responses of individual PDXs were not always consistent) (Supplementary Fig. 6b–f). These data strongly support the selectivity of the biological effects that are observed with compounds modulating the same genetic target, and thus reinforce the validity and reproducibility of the 1 × 1 × 1 PCT approach.

### PCT predicts patient response to targeted therapies

The translatability of the 1 × 1 × 1 PCT concept was assessed by comparing the responses of targeted therapies observed in PDXs to clinical outcomes in cancer patients. Encorafenib is a highly potent and selective RAF inhibitor, and 69% of melanoma cell lines with a mutation in *BRAF* (encoding BRAF, B-raf proto-oncogene, serine/threonine kinase)- responded to encorafenib treatment *in vitro* (cutoff 0.2 μM, Supplementary Fig. 7a). When administered at clinically relevant exposure levels, as predicted, 67% of PDXs mutated in *BRAF* at the position encoding amino acid V600 responded to encorafenib, with 42% demonstrating mCR or mPR (Fig. 3a). These data are consistent with the reported phase 2 clinical trial response rates of the BRAF inhibitors vemurafenib<sup>22</sup> and dabrafenib<sup>23</sup>. Furthermore, none of the *NRAS* (encoding the neuroblastoma RAS viral (v-ras) oncogene homolog protein, NRAS)-mutated PDXs responded to encorafenib treatment, consistent with reports that melanoma patients with mutated *NRAS* are relatively insensitive to type 1 and type 1.5 BRAF inhibitors compared to patients with wild-type *NRAS*<sup>24,25</sup>.

About 30% of *BRAF*-mutated melanomas are intrinsically insensitive to BRAF inhibitors<sup>22,23</sup>. Improved responses in this population can be achieved by inhibiting both BRAF and its downstream target MEK kinase<sup>26</sup>. Administering a combination of encorafenib



**Figure 3** PCT predicts targeted therapy response and validates predictive gene signature. (a,b) Waterfall plot of response to the BRAF inhibitor encorafenib ( $n = 33$ ) (a) and encorafenib in combination with the MEK inhibitor binimetinib ( $n = 33$ ) (b) among melanoma PDXs. GOF, gain of function; WT, wild type. (c) Waterfall plot of response to DR5 agonist TAS266 among DR5 signature-positive and DR5 signature-negative melanoma PDXs ( $n = 33$ ). (d) Kaplan-Meier PFS curve with TAS266 treatment in melanoma, stratified by DR5 predictive signature ( $n = 33$ ). (e) Waterfall plot of response to PI3K $\alpha$  inhibitor BYL719 among PI3K $\alpha$ - and PTEN- mutated PDXs across five indications ( $n = 205$ ). LOF, loss of function; blue, PIK3CA GOF/PTEN WT ( $n = 21$ ); green, PIK3CA GOF/PTEN LOF ( $n = 6$ ); brown, PIK3CA WT/PTEN LOF ( $n = 9$ ); gray, PIK3CA WT/PTEN WT ( $n = 169$ ).

**Table 1** Candidate resistance mechanisms in melanoma PDXs with acquired resistance to MAPK pathway inhibitors

PDX model	Treatment	Dosing	Response	Duration of response (d)	MAPK pathway candidate resistance mechanism
X-3676	Encorafenib; binimetinib	20 mg/kg q.d. 10 mg/kg b.i.d.	mSD	39	<i>BRAF</i> amplification
X-3483	Encorafenib	20 mg/kg q.d.	mCR	143	<i>BRAF</i> amplification
X-3676	Encorafenib; BKM120	20 mg/kg q.d. 35 mg/kg q.d.	mSD	35	<i>BRAF</i> amplification
X-2992	Encorafenib; BKM120	20 mg/kg q.d. 35 mg/kg q.d.	mPR	125	<i>MAP2K1</i> <sup>E203K</sup>
X-2613	Binimetinib; LEE011	10 mg/kg b.i.d. 250 mg/kg q.d.	mPR	241	<i>MAP2K2</i> <sup>Q218P</sup>

q.d., daily; b.i.d., twice daily.

and binimetinib (a small-molecule MEK inhibitor) synergistically inhibited the growth of *BRAF*-mutated cell lines *in vitro* (Supplementary Fig. 7b,c). Binimetinib single-agent treatment yielded moderate anti-tumor activity in the melanoma PCT (Supplementary Fig. 8); by contrast, the combined treatment increased the response rate to 100% among *BRAF*-mutated PDXs, with 72% displaying mCR or mPR (Fig. 3b). This is concordant with the results seen in a combination phase 1/2 trial in *BRAF*(V600)-mutant melanoma patients, in which a combination of dabrafenib and trametinib resulted in a 94% response rate, with 76% PR and CR<sup>26,27</sup>. The consistencies between preclinical and clinical data support the utility of the PCT approach for reproducing clinical patient responses to targeted therapies.

### PCT successfully reveals the mechanisms of resistance

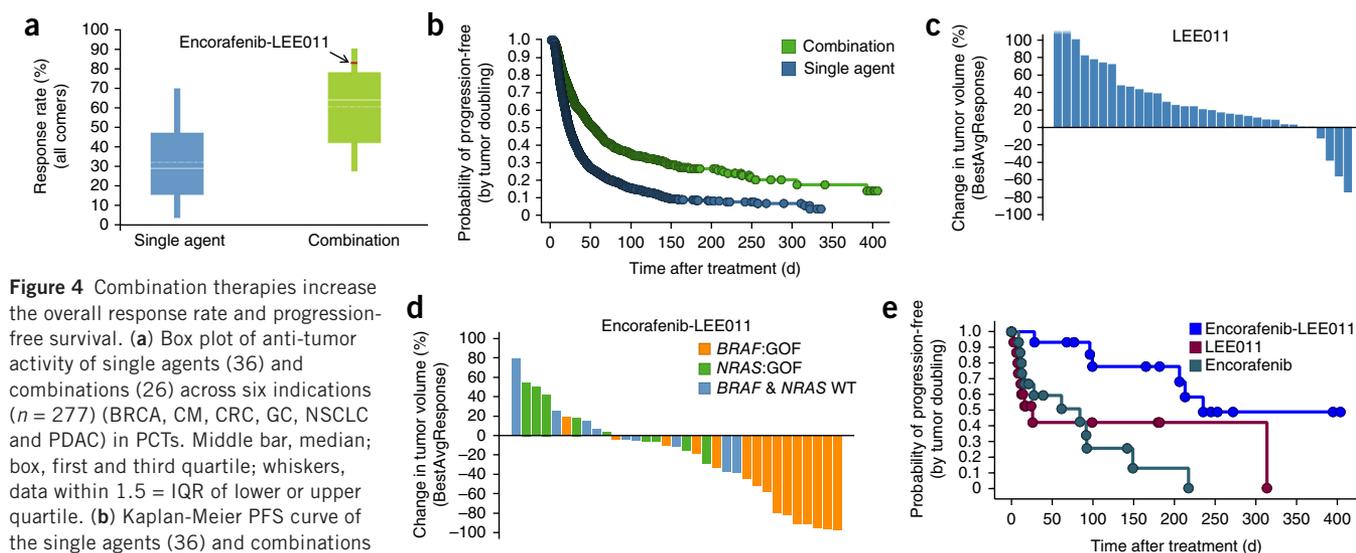
To enable the robust study of escape mechanisms, PDXs showing primary responses were all treated until resistance developed, after which they were subjected to DNA- and RNA-sequencing analysis. We focused our initial resistance analysis on the MAPK inhibitors. Three PDXs treated with either encorafenib as a single agent, or in combination with BKM120 or binimetinib, developed resistance through *BRAF* amplification (Table 1 and Supplementary Fig. 9a–c),

These data strongly suggest that PDXs and PCTs are well-suited tools for discovering mechanisms of resistance.

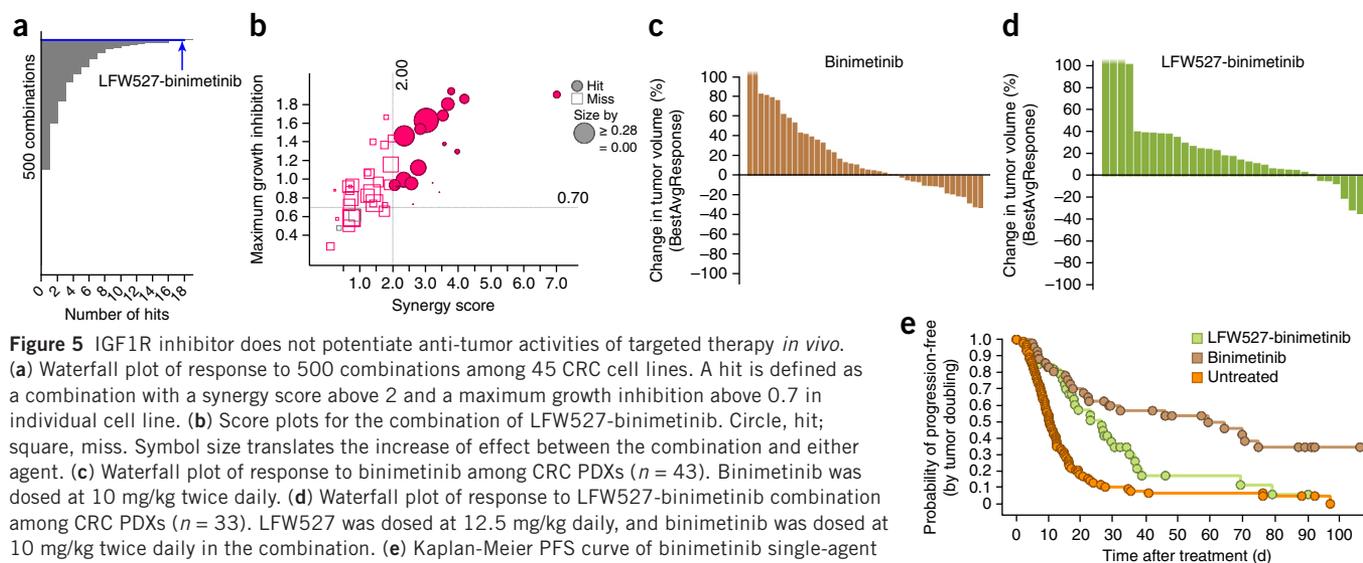
### PCT functionally supports predictive biomarker hypotheses

We next asked whether a PCT could be used to provide initial validation of genetic biomarker hypotheses emerging from *in vitro* cell line studies. TAS266 is a nanobody developed to activate the DR5 (death receptor 5, encoded by *TNFRSF10B*)-mediated apoptosis pathway in cancer cells<sup>32</sup>. A signature predictive of treatment response, developed from an *in vitro* cell line screen, was defined as when DR5 expression > median DR5 expression and caspase-8 expression > median caspase-8 expression. We assessed response to TAS266 in the melanoma PCT, and 18% of PDXs displayed a range of sensitivity (mPR and mCR) (Fig. 3c). We retrospectively applied the predictive signature to this data set, and found that 80% (4/5) of signature-positive models responded to the treatment. Moreover, the signature-positive xenografts had significantly better PFS (hazard ratio 0.09, 95% interval 0.01–0.71,  $P < 1 \times 10^{-16}$ ) than did the signature-negative cohorts (Fig. 3d).

A PIK3CA (encoding phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit  $\alpha$ ) mutation has been reported as the



**Figure 4** Combination therapies increase the overall response rate and progression-free survival. (a) Box plot of anti-tumor activity of single agents (36) and combinations (26) across six indications ( $n = 277$ ) (BRCA, CM, CRC, GC, NSCLC and PDAC) in PCTs. Middle bar, median; box, first and third quartile; whiskers, data within 1.5 = IQR of lower or upper quartile. (b) Kaplan-Meier PFS curve of the single agents (36) and combinations (26) across six indications in the PCTs ( $n = 277$ ). The y axis is the percentage of animals on study as calculated by tumor doubling (when a tumor volume has doubled compared to its baseline tumor volume), after which a model is considered to have progressed on treatment. (c) Waterfall plot of response to LEE011 among melanoma PDXs ( $n = 33$ ). (d) Waterfall plot of response to the encorafenib-LEE011 combination among melanoma PDXs ( $n = 33$ ). (e) Kaplan-Meier PFS curve of encorafenib and LEE011 single agents and encorafenib-LEE011 combination among *BRAF*-mutated melanoma PDXs ( $n = 15$ ).



**Figure 5** IGF1R inhibitor does not potentiate anti-tumor activities of targeted therapy *in vivo*. (a) Waterfall plot of response to 500 combinations among 45 CRC cell lines. A hit is defined as a combination with a synergy score above 2 and a maximum growth inhibition above 0.7 in individual cell line. (b) Score plots for the combination of LFW527-binimetinib. Circle, hit; square, miss. Symbol size translates the increase of effect between the combination and either agent. (c) Waterfall plot of response to binimetinib among CRC PDXs ( $n = 43$ ). Binimetinib was dosed at 10 mg/kg twice daily. (d) Waterfall plot of response to LFW527-binimetinib combination among CRC PDXs ( $n = 33$ ). LFW527 was dosed at 12.5 mg/kg daily, and binimetinib was dosed at 10 mg/kg twice daily in the combination. (e) Kaplan-Meier PFS curve of binimetinib single-agent ( $n = 43$ ) and LFW527-binimetinib combination ( $n = 33$ ) among CRC PDXs.

foremost positive predictor of sensitivity to BYL719, a selective PI3K- $\alpha$  inhibitor, in breast cancers<sup>33</sup>. The response rates between PDXs harboring *PIK3CA* and/or *PTEN* (encoding phosphatase and tensin homolog) mutations and PDXs with wild-type *PIK3CA* and *PTEN* were similar (50% versus 54%); however, we observed differential responses among *PIK3CA*-mutated PDXs with different *PTEN* statuses (Fig. 3e). More than 67% of PDXs with *PIK3CA* mutations and wild-type *PTEN* responded, whereas only 11% of the wild-type *PIK3CA* and *PTEN*-mutated PDXs did, adding further evidence that people with a *PIK3CA* mutation are more likely to benefit from BYL719 treatment than are people with *PTEN* mutations<sup>34</sup>. These data suggest that PCTs, combined with comprehensive genomic characterization, can provide an effective *in vivo* validation step for emerging predictive gene signatures.

#### Multiple CDK4-inhibitor combinations dramatically prolong PFS

More than a third of our PCTs explored combination treatments. The median response rate was 30% across 36 single agents, compared to 64% for 26 combinations tested among the unselected PDXE population (Fig. 4a). Furthermore, combinations produce a significantly longer median PFS period than do single agents (22 d for single agent versus 55 d for combination; hazard ratio 0.52, 95% interval 0.48–0.57,  $P < 1 \times 10^{-16}$ ) because they delay the emergence of resistance (Fig. 4b). Among the tested combinations, encorafenib combined with CDK4/6 inhibitor LEE011 showed dramatically enhanced activity compared to the single agents (Figs. 3a and 4c), with a 100% response rate, including 87% mPR and mCR (Fig. 4d) in melanoma PDXs. Pharmacokinetics (PK) studies demonstrated no drug-drug interaction (DDI)<sup>4</sup> (Supplementary Fig. 10). Furthermore, this combination significantly improved PFS and delayed the development of resistance compared to encorafenib alone (hazard ratio 0.17; 95% interval 0.06–0.52,  $P = 1.8 \times 10^{-3}$  (Fig. 4e). It is noteworthy that 40% of the PDXs treated with the encorafenib-LEE011 combination never developed resistance under continuous treatment for up to 200 d (Fig. 4e).

The ability of LEE011 to boost efficacy was further exemplified by its combinatorial effect with BYL719 in breast cancer (Supplementary Fig. 11a–d), suggesting that combining CDK4/6 inhibitors with other targeted agents is a highly effective approach for maximally inhibiting

the oncogenic pathway and preventing or delaying the development of drug resistance. Notably, BYL719 did not show increased combinatorial potential when combined with LEE011 *in vitro*, demonstrating a potential disconnect between *in vitro* cell-based assays and *in vivo* PDX-based PCT screens (Supplementary Fig. 11e,f).

#### IGF1R inhibitors do not potentiate the activity of targeted therapies

Another example in which a disconnect between *in vitro* cell-based assays and *in vivo* PCTs has been observed is in the ability of IGF1R (insulin-like growth factor 1 receptor) inhibitors to potentiate the anti-tumor activities of multiple targeted therapies. LFW527, a low-molecular-weight (LMW) IGF1R inhibitor, has been identified in a large-scale combination screen as having strong synergy in 18 out of 45 CRC cell lines when combined with binimetinib (Fig. 5a,b). This combination was screened in the CRC trial across 35 PDXs. At clinically relevant exposure levels, binimetinib demonstrated mild anti-tumor activity with a response rate of 55% (1 mPR and 22 mSD) (Fig. 5c). Contrary to what was observed *in vitro*, the combination of LFW527 with binimetinib did not improve the response rate (Fig. 5d), or improve PFS when compared to binimetinib single-agent treatment (hazard ratio 1.43, 95% interval 0.83–2.47,  $P = 0.19$ ) (Fig. 5e). PK studies with both IGF1R inhibitors showed no DDI issues encountered (data not shown), suggesting that the lack of synergy *in vivo* is unlikely to be due to insufficient target inhibition in the combination groups. Similarly, LFW527 potentiated the *in vitro* activity of binimetinib in the non-small cell lung carcinoma (NSCLC) cells (Supplementary Fig. 12a,b) and PDAC cells (Supplementary Fig. 13a,b), but failed to demonstrate any synergy in the PCT studies (Supplementary Fig. 12c–e and Supplementary Fig. 13c–e), regardless of whether an LMW inhibitor or a fully humanized antibody was used. Furthermore, such a disconnect between the *in vitro* and *in vivo* studies was not limited to the binimetinib with IGF1R inhibitor treatment combination. When combined with the mTOR inhibitor everolimus, LFW527 again demonstrated strong synergy *in vitro*, but it neither enhanced response rate nor improved PFS in breast cancer trials *in vivo* (data not shown). These data together provide examples of a disconnect between *in vitro* screens conducted in cell lines and *in vivo* PCTs conducted in PDXs.

## DISCUSSION

We have demonstrated the effectiveness of the  $1 \times 1 \times 1$  PCT experimental paradigm for examining population-based *in vivo* compound screens. It enables insight into inter-patient response heterogeneity in an efficient manner, and helps to identify responsive subpopulations, thus enabling the discovery of predictive biomarkers. In addition, it can be used to identify clinically relevant mechanisms of resistance. We propose this experimental paradigm for preclinical drug evaluation to enhance the predictability for phase 1/2 clinical trials.

As with all preclinical models, there are limitations with PDXs, including the lack of an intact immune system, differential influences of mouse stroma versus human stroma and, in addition, the underrepresentation of very specific genotypes and specific lineage subtypes. Nonetheless, conservation of major genetic alterations found in patient tumors in the corresponding PDXs underlies the extensive utility of these models in the preclinical studies. Eirew *et al.*<sup>14</sup> have demonstrated clonal dynamics during *in vivo* engraftment in breast cancer patient xenografts, although once established, the xenografts are genetically relatively stable. This could have a profound impact when we consider using personalized PDXs to guide patient therapy, as in the Avatar trial setting<sup>8</sup>. The data presented in the current study, however, are concerned with the population response. As PDXs closely resemble the genomic landscape of human cancers at the population level, there is a strong rationale for performing preclinical drug screens to investigate the population-based inter-patient response heterogeneity.

In addition, we have demonstrated that PDXs may well have an advantage over long-established cell lines. Our genomic analysis revealed that various signaling pathways are under- or overrepresented in cell lines across lineages. For example, we found underrepresentation of alterations in the PI3K pathway in NSCLC and overrepresentation of the transforming growth factor (TGF)- $\beta$  pathway in PDAC and receptor tyrosine kinase (RTK) alterations in BRCA. In contrast, these pathways are accurately represented in the PDX collection at a similar mutation frequency as that reported in patient tumors. The discrepancies between *in vitro* and *in vivo* may well reflect the clonal bias inherent in immortalized cells propagated on plastic. Another well-understood example is that the dependency of cancer cells on developmental pathways is lost upon culturing extensively *in vitro* and is not re-acquired once returned to an *in vivo* environment<sup>35</sup>. Furthermore, when considering the fact that only a limited number of cell lines are able to grow *in vivo* as xenograft models, PDXs more broadly represent inter-patient diversity.

The liabilities with cell-based models are borne out in discrepancies with respect to drug response. An *in vitro* combination screen in melanoma failed to identify the combinatorial effect of the CDK4/6 inhibitor with other targeted therapies, whereas the PCT subsequently did reveal this effect (as exemplified by the combination of LEE011 and encorafenib). Our studies additionally show differential combinatorial effects with IGF1R inhibitors *in vitro* and *in vivo*. Indeed, there are several literature reports showing that combinatorial inhibition of IGF1R and MEK1/2 (refs. 36–38) or of PI3K/mTOR<sup>39</sup> are efficacious not only in cell proliferation assays *in vitro* but also, importantly, in cell line-derived xenografts *in vivo*. These positive results have led to a number of clinical trials in CRC and NSCLC, as well as in several other indications, yet the data from the clinical trials are fairly disappointing<sup>40–43</sup>, concordant with our results in PDX models.

Analysis of the melanoma-resistant tumors under continuous drug treatment by deep-sequencing analyses has revealed similar mechanisms of resistance to both single-agent (encorafenib) and combination treatments (encorafenib-binimetinib and encorafenib-BKM120)

reported in the clinic, supporting the idea that the PCT approach can recapitulate the mechanisms of drug resistance found in patients. Additionally, we have identified a novel mutation (*MAP2K2*<sup>Q218P</sup>) that could potentially confer resistance to the binimetinib-LEE011 combination currently being evaluated in the clinic to treat *NRAS*-mutated melanomas. Das Thakur *et al.*<sup>44</sup> reported *BRAF* amplification as a mechanism conferring resistance to *BRAF* inhibitors; despite the use of multiple animals in each treatment group, this was the only mechanism identified. In contrast, the population-based PCT studies uncovered diverse mechanisms of resistance to *BRAF* inhibition, demonstrating a very efficient approach to studying drug resistance in an *in vivo* setting, and consequently, providing opportunities to develop strategies that are likely to mitigate resistance mechanisms at an early stage of drug development.

To our knowledge, the herein described PDX clinical trial concept is the first to evaluate the reproducibility and translatability of the  $1 \times 1 \times 1$  concept to drug response using an extensive, well-characterized PDX collection. This approach represents a new experimental paradigm through which to address the tumor biology of cancer patients, and interrogate targeted therapies in *in vivo* models that are more relevant to the clinic than are traditional oncology models, which could potentially improve the ability of preclinical oncology studies to predict patient response in the clinic.

## METHODS

Methods and any associated references are available in the [online version of the paper](#).

*Note: Any Supplementary Information and Source Data files are available in the online version of the paper.*

## ACKNOWLEDGMENTS

We thank B. Gruenenfelder, M. Lechevalier and D. Thomis for project management; R. Mosher and M. Murakami for their advice on the pathology of PDXs; L. Barys, P. Fordjour, M. Gallagher, B. Gorbacheva, N. Houde, E. Kurth, J.A. Kwon, Y. Oei, K. O'Malley, D. Rakiec and C. Taurus for their technical support; D. Fox for IT support; S.-M. Maira, C. Fritsch and M. Yao for their helpful discussion; M. Stump, L. Kifule and P. Zhu for support with *in vitro* proliferation screens; J. Ledell for Chalice software development; and J. Steiger for data interpretation and project management. We received tumor specimens from the US National Disease Research Interchange, the US National Cancer Institute, the Maine Medical Center, the Tufts Medical Center, the Mt Group Inc. and GenenDesign, and we are grateful to the people who consented to donate their tissues to support this work.

## AUTHOR CONTRIBUTIONS

H.G., S.F., Y.W., M.S., C.Z., C.S., G.Y., S.B., H.C., S. Chatterjee, S.M.C., S.D.C., N.E., M.E., C.K., E. Lorenzana, M.P., C.R., F. Salangsang, F. Santacroce, Y.T., W.T., S.T., R.V., F.V.A., Z.W., D.W. and F.X. performed the PCT trials; H.G., G.Y., Y.Z., S.M.C., J.G., C.K., A.L., R.V., Z.W. and F.X. performed PDX model development; E.D., Y.L., M.E.M. and R.M. performed histopathologic analysis; J.M.K. and E.R.M. led the genomic landscape analysis; J.M.K., F.C., S. Chuai, A.K., J.M., J.L., A.R. and K.V. performed computational biology and bioinformatics analysis; O.A.B., D.Y.C., J.R.L. and A.P.S. performed pan-cancer panel analysis for melanoma resistance; J.E.M., J.G., T.L.N. and D.A.R. performed or directed nuclear acid extraction, quality control and genomic data generation; H.Q.W. performed the PK analysis of the encorafenib and LEE011 combination; M.W. led the *in vitro* combination screens; H.G., J.M.K., J.M., A.R., O.A.B., D.Y.C. prepared figures and tables for the main text and supplementary information; H.G., J.M.K., J.E.M., S.F., M.S., C.S., O.A.B., A.P.S., D.Y.C., M.W., H.B., J.A.W. and W.R.S. wrote and edited the main text and supplementary information; P.A., R.C., M.R.J., N.K.P., J.A.W., E. Li, E. Lees, F.H., N.K. and W.R.S. contributed to project oversight and advisory roles; J.A.W. and W.R.S. provided overall project leadership.

## COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the [online version of the paper](#).

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

1. Arrowsmith, J. & Miller, P. Trial watch: phase II and phase III attrition rates 2011–2012. *Nat. Rev. Drug Discov.* **12**, 569 (2013).
2. Arrowsmith, J. Trial watch: Phase II failures: 2008–2010. *Nat. Rev. Drug Discov.* **10**, 328–329 (2011).
3. DiMasi, J.A., Reichert, J.M., Feldman, L. & Malins, A. Clinical approval success rates for investigational cancer drugs. *Clin. Pharmacol. Ther.* **94**, 329–335 (2013).
4. Paul, S.M. *et al.* How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat. Rev. Drug Discov.* **9**, 203–214 (2010).
5. Tentler, J.J. *et al.* Patient-derived tumour xenografts as models for oncology drug development. *Nat. Rev. Clin. Oncol.* **9**, 338–350 (2012).
6. Siolas, D. & Hannon, G.J. Patient-derived tumor xenografts: transforming clinical samples into mouse models. *Cancer Res.* **73**, 5315–5319 (2013).
7. Rosfjord, E., Lucas, J., Li, G. & Gerber, H.P. Advances in patient-derived tumor xenografts: from target identification to predicting clinical response rates in oncology. *Biochem. Pharmacol.* **91**, 135–143 (2014).
8. Hidalgo, M. *et al.* Patient-derived xenograft models: an emerging platform for translational cancer research. *Cancer Discov.* **4**, 998–1013 (2014).
9. Bertotti, A. *et al.* A molecularly annotated platform of patient-derived xenografts (“xenopatiens”) identifies HER2 as an effective therapeutic target in cetuximab-resistant colorectal cancer. *Cancer Discov.* **1**, 508–523 (2011).
10. Migliardi, G. *et al.* Inhibition of MEK and PI3K/mTOR suppresses tumor growth but does not cause tumor regression in patient-derived xenografts of RAS-mutant colorectal carcinomas. *Clin. Cancer Res.* **18**, 2515–2525 (2012).
11. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).
12. DeRose, Y.S. *et al.* Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes. *Nat. Med.* **17**, 1514–1520 (2011).
13. Ding, L. *et al.* Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* **464**, 999–1005 (2010).
14. Eirew, P. *et al.* Dynamics of genomic clones in breast cancer patient xenografts at single-cell resolution. *Nature* **518**, 422–426 (2015).
15. Hennessey, P.T. *et al.* Promoter methylation in head and neck squamous cell carcinoma cell lines is significantly different than methylation in primary tumors and xenografts. *PLoS ONE* **6**, e20584 (2011).
16. Julien, S. *et al.* Characterization of a large panel of patient-derived tumor xenografts representing the clinical heterogeneity of human colorectal cancer. *Clin. Cancer Res.* **18**, 5314–5328 (2012).
17. Mattie, M. *et al.* Molecular characterization of patient-derived human pancreatic tumor xenograft models for preclinical and translational development of cancer therapeutics. *Neoplasia* **15**, 1138–1150 (2013).
18. Einarsdottir, B.O. *et al.* Melanoma patient-derived xenografts accurately model the disease and develop fast enough to guide treatment decisions. *Oncotarget* **5**, 9609–9618 (2014).
19. de Plater, L. *et al.* Establishment and characterisation of a new breast cancer xenograft obtained from a woman carrying a germline *BRCA2* mutation. *Br. J. Cancer* **103**, 1192–1200 (2010).
20. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* **6**, 80–92 (2012).
21. Therasse, P. *et al.* New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J. Natl. Cancer Inst.* **92**, 205–216 (2000).
22. Sosman, J.A. *et al.* Survival in BRAF V600-mutant advanced melanoma treated with vemurafenib. *N. Engl. J. Med.* **366**, 707–714 (2012).
23. Ascierto, P.A. *et al.* Phase II trial (BREAK-2) of the BRAF inhibitor dabrafenib (GSK2118436) in patients with metastatic melanoma. *J. Clin. Oncol.* **31**, 3205–3211 (2013).
24. Kaplan, F.M., Shao, Y., Mayberry, M.M. & Aplin, A.E. Hyperactivation of MEK-ERK1/2 signaling and resistance to apoptosis induced by the oncogenic B-RAF inhibitor, PLX4720, in mutant N-RAS melanoma cells. *Oncogene* **30**, 366–371 (2011).
25. Halaban, R. *et al.* PLX4032, a selective BRAF(V600E) kinase inhibitor, activates the ERK pathway and enhances cell migration and proliferation of BRAF melanoma cells. *Pigment Cell Melanoma Res.* **23**, 190–200 (2010).
26. Flaherty, K.T. *et al.* Combined BRAF and MEK inhibition in melanoma with BRAF V600 mutations. *N. Engl. J. Med.* **367**, 1694–1703 (2012).
27. Robert, C. *et al.* Improved overall survival in melanoma with combined dabrafenib and trametinib. *N. Engl. J. Med.* **372**, 30–39 (2015).
28. Shi, H. *et al.* Acquired resistance and clonal evolution in melanoma during BRAF inhibitor therapy. *Cancer Discov.* **4**, 80–93 (2014).
29. Shi, H. *et al.* Melanoma whole-exome sequencing identifies (V600E)B-RAF amplification-mediated acquired B-RAF inhibitor resistance. *Nat. Commun.* **3**, 724 (2012).
30. Wagle, N. *et al.* MAP kinase pathway alterations in BRAF-mutant melanoma patients with acquired resistance to combined RAF/MEK inhibition. *Cancer Discov.* **4**, 61–68 (2014).
31. Rizos, H. *et al.* BRAF inhibitor resistance mechanisms in metastatic melanoma: spectrum and clinical impact. *Clin. Cancer Res.* **20**, 1965–1977 (2014).
32. Papadopoulos, K.P. *et al.* Unexpected hepatotoxicity in a phase I study of TAS266, a novel tetravalent agonistic nanobody targeting the DR5 receptor. *Cancer Chemother. Pharmacol.* **75**, 887–895 (2015).
33. Fritsch, C. *et al.* Characterization of the novel and specific PI3K $\alpha$  inhibitor NVP-BYL719 and development of the patient stratification strategy for clinical trials. *Mol. Cancer Ther.* **13**, 1117–1129 (2014).
34. Juric, D. *et al.* Convergent loss of PTEN leads to clinical resistance to a PI(3)K $\alpha$  inhibitor. *Nature* **518**, 240–244 (2015).
35. Sasai, K. *et al.* Shh pathway activity is down-regulated in cultured medulloblastoma cells: implications for preclinical studies. *Cancer Res.* **66**, 4215–4222 (2006).
36. Flanigan, S.A. *et al.* Overcoming IGF1R/IR resistance through inhibition of MEK signaling in colorectal cancer models. *Clin. Cancer Res.* **19**, 6219–6229 (2013).
37. Molina-Arcas, M., Hancock, D.C., Sheridan, C., Kumar, M.S. & Downward, J. Coordinate direct input of both KRAS and IGF1 receptor to activation of PI3 kinase in KRAS-mutant lung cancer. *Cancer Discov.* **3**, 548–563 (2013).
38. Ebi, H. *et al.* Receptor tyrosine kinases exert dominant control over PI3K signaling in human KRAS mutant colorectal cancers. *J. Clin. Invest.* **121**, 4311–4321 (2011).
39. Friedbichler, K. *et al.* Pharmacodynamic and antineoplastic activity of BI 836845, a fully human IGF ligand-neutralizing antibody, and mechanistic rationale for combination with rapamycin. *Mol. Cancer Ther.* **13**, 399–409 (2014).
40. Moran, T. *et al.* Activity of dalotuzumab, a selective anti-IGF1R antibody, in combination with erlotinib in unselected patients with Non-small-cell lung cancer: a phase I/II randomized trial. *Exp. Hematol. Oncol.* **3**, 26 (2014).
41. Scagliotti, G.V. *et al.* Randomized, phase III trial of figitumumab in combination with erlotinib versus erlotinib alone in patients with nonadenocarcinoma nonsmall-cell lung cancer. *Ann. Oncol.* **26**, 497–504 (2015).
42. Brana, I. *et al.* A parallel-arm phase I trial of the humanised anti-IGF-1R antibody dalotuzumab in combination with the AKT inhibitor MK-2206, the mTOR inhibitor ridaforolimus, or the NOTCH inhibitor MK-0752, in patients with advanced solid tumours. *Br. J. Cancer* **111**, 1932–1944 (2014).
43. Di Cosimo, S. *et al.* Combination of the mTOR inhibitor ridaforolimus and the anti-IGF1R monoclonal antibody dalotuzumab: preclinical characterization and phase I clinical trial. *Clin. Cancer Res.* **21**, 49–59 (2015).
44. Das Thakur, M. *et al.* Modelling vemurafenib resistance in melanoma reveals a strategy to forestall drug resistance. *Nature* **494**, 251–255 (2013).

## ONLINE METHODS

**Patient tumor specimen collection and annotation.** Patient tumor specimens (untreated or treated) were obtained from non-profit and commercial vendors in the US and China (PDXs were established and exclusively used in China for tumors originating from patients in China). All patients provided informed consent for the tumor samples procured by Novartis, Inc. from: the National Disease Research Interchange, Philadelphia, Pennsylvania, USA; National Cancer Institute, Rockville, Maryland, USA; Maine Medical Center, Portland, Maine, USA; Tufts Medical Center, Boston, Massachusetts, USA; The Mt Group Inc., Van Nuys, California, USA; and GenenDesign, Shanghai, China. Clinical and pathologic data were entered and maintained in Novartis databases.

**Generation of PDX models.** Mice were maintained and handled in accordance with the Novartis Institutes for BioMedical Research (NIBR) Animal Care and Use Committee protocols and regulations. Patient tumor material was collected in culture medium and kept on wet ice for engraftment within 24 h after resection. Upon arrival, necrotic and supporting tissues were carefully removed using a surgical blade. Three ~30–50-mg tissue fragments were flash frozen and stored at –80 °C for genomic profiling, and ~50–100-mg tissue fragments were fixed in 10% neutral-buffered formalin and paraffin embedded for histopathologic analysis. Approximately 20–30-mg tissue fragments were implanted subcutaneously into the flank region of athymic nude (nu/nu) female mice using a trocar. Successfully engrafted tumor models were then passaged and banked after three passages in mice. The identity of the established PDXs was confirmed by SNP48 analysis before and after PCT studies.

**PCT and drug treatment.** Established and genomically characterized PDXs of BRCA, CM, CRC, gastric cancer (GC), NSCLC and PDAC were used for the PCT study. Tumors were subcutaneously implanted into nude mice at passages between p4 and p10. Tumors (~200 mm<sup>3</sup>) were enrolled on a rolling base and treated with single agents or combinations without pre-selection on the basis of their genetic or proteomic, or growth latency/doubling time. A total of 38 unique therapeutic entities were screened, encompassing five cytotoxic chemotherapeutics and 33 targeted agents (including small molecule, antibody and antibody-drug conjugate), used as a single treatment (36) or in combination (26). Approximately 20 treatments were tested in each of six indications (BRCA, CM, CRC, GC, NSCLC, PDAC) with 29–45 models for each indication. The maximum tolerated dose (MTD) was used for the agents that have not entered the clinics yet, whereas the clinically relevant dose (CRD) was used for agents that are currently used or evaluated in the clinic. The CRD was converted from a human dose by matching the human blood exposure to mouse blood exposure. The standard treatment schedule was 21 d. Fast-growing tumors were dosed until the tumors reached ~1,500 mm<sup>3</sup>, and treatment was continued until the tumor volume doubled at least twice for slow-growing tumors. The tumor size was evaluated twice weekly by caliper measurements, and the approximate volume of the mass was calculated using the formula  $(l \times w \times w) \times (\pi/6)$ , where  $l$  is the major tumor axis and  $w$  is the minor tumor axis. In the absence of progression or an adverse event, treatment was continued for at least 90 d.

**Response calls.** The response was determined by comparing tumor volume change at time  $t$  to its baseline: % tumor volume change =  $\Delta Vol_t = 100\% \times ((V_t - V_{initial}) / V_{initial})$ . The BestResponse was the minimum value of  $\Delta Vol_t$  for  $t \geq 10$  d. For each time  $t$ , the average of  $\Delta Vol_t$  from  $t = 0$  to  $t$  was also calculated. We defined the BestAvgResponse as the minimum value of this average for  $t \geq 10$  d. This metric captures a combination of speed, strength and durability of response into a single value. The criteria for response (mRECIST) were adapted from RECIST criteria<sup>21</sup> and defined as follows (applied in this order): mCR, BestResponse < –95% and BestAvgResponse < –40%; mPR, BestResponse < –50% and BestAvgResponse < –20%; mSD, BestResponse < 35% and BestAvgResponse < 30%; mPD, not otherwise categorized. Mice that were sacrificed because of an adverse event before they had completed 14 d on trial were removed from the data set.

**DNA and RNA extraction.** RNA and DNA from ~30 mg of flash-frozen tissue were purified with the QIAcube (Qiagen cat. no. 9001292) automated sample preparation

platform using the Qiagen ALLPrep DNA/RNA Mini Kit (cat. no.80204). The RNA concentration and integrity was evaluated with the Agilent 2100 Bioanalyzer (cat. no. G2940CA) using the Agilent RNA 6000 Nano Kit (cat. no. 5067-1511) and protocol. DNA quantification was assessed with the NanoDrop 8000.

**PDX identity validation by SNP genotyping.** The identities of tissue samples derived from the PDX models were established by SNP genotyping DNA derived from the original patient tumor, as well as tissue from each PDX model at different passages. The SNP genotyping panel uses a set of 48 intergenic SNPs originally developed by the Broad Institute and previously described<sup>11</sup>. The assays were performed at Asuragen, Inc. using a SNaPshot multiplex assay format (Life Technologies, Inc.) with the analysis performed using a Novartis-developed application.

**Copy-number analysis (SNP array).** Copy-number analysis was derived from the profiling of total DNA on the Affymetrix genome-wide human SNP Array 6.0 chip using instrumentation and protocols from Affymetrix. The profiling was performed at either the Novartis PHARMA core facility in Basel, Switzerland, or by CiToxLAB in France. The DNA samples were profiled in batches of 96 chips containing 15 normal (HapMap) DNA samples for QC and normalization as described previously<sup>11</sup>. The raw CEL files were returned, and the QC steps were performed using Affymetrix Genotyping Console. The MAPD (median of the absolute values of all pairwise differences) parameter was the critical QC metric. Chips with a MAPD value >0.35 were excluded from analysis. The SNP genotyping QC metrics (contrast QC and call rate) were not relied on because the human DNA was derived from tumor cells. Varying amounts of mouse DNA did not substantially affect sample QC, and samples without human DNA failed before reaching the MAPD calculation. The CEL files from chips passing QC were used for copy-number analysis. Both a copy-number segmentation file and a gene-level copy-number value were generated using Partek Genomic Suite 6.6 genomic segmentation algorithm (Partek, Inc.). CIN scores were calculated as the s.d. of the mean copy-number across chromosome arms.

**Expression arrays.** Total RNA with an acceptable integrity (Agilent RNA integrity number (RIN) > 7.5) was used for gene expression profiling. The profiling was performed at either the Novartis PHARMA core facility in Basel, Switzerland, or by CiToxLAB in France using instrumentation and protocols from Affymetrix. The Human Genome 133 Plus 2.0 gene chip (Affymetrix, Inc.) was used with the Affymetrix 3' IVT One-Cycle target labeling protocol. The raw CEL files were returned, and the QC steps were performed using an Affymetrix Genotyping Console. Chip quality was assessed with selected quality-control parameters (background, % present calls, scaling factor and the 3'/5' ratios of  $\beta$ -actin (*ACTB*) and glyceraldehyde dehydrogenase (*GAPDH*) reference genes) generated using the MAS 5.0 algorithm (MicroArray Suite 5) developed by Affymetrix. All samples had a 'scaling factor' between 0.71 and 1.15 and the 'background average' value ranged from 44.5 to 57.2. Finally, the 3'/5' ratios for the *GAPDH* gene were between 0.88 and 0.93.

**RNA-seq analysis and mutation calls.** Total RNA was used as input to the Illumina mRNA-Seq 8 Sample Prep Kit (cat. no. RS-100-0801) or TruSeq RNA Sample Prep Kit v2-Sets A/B (48Rxn) (cat. no. FC-122-1001 and FC-122-1002), depending on the date of the RNA-Seq library generation. RNA input ranged from 0.25  $\mu$ g to 2  $\mu$ g with RIN (RNA Integrity Number) scores from 5.1 to 10.0. A size fragment range between 200 and 300 bp was selected on a 2% gel using Bio-Rad low-range ultra-agarose (cat. no. 161-3107) for the mRNA-Seq 8 sample Prep Kit manually. The gel-free protocol was employed for the TruSeq RNA Sample Prep Kit according to the manufacturer's specifications, and was performed on the Beckman Coulter Biomek FXp robotics platform. The standard RNA fragmentation profile was used, as recommended by Illumina (94 °C for 5 min for the mRNA-Seq 8 sample Prep Kit, and 94 °C for 8 min for the TruSeq RNA Sample Prep Kit). The PCR amplified RNA-Seq library products were then quantified using the Agilent DNA 1000 kit (cat. no. 5067-1504) on the Agilent 2100 Bioanalyzer or the Standard Sensitivity NGS Fragment Analysis Kit (cat. no. DNF-479) on the Advance Analytical Fragment Analyzer. The samples were diluted to 10 nM in Qiagen Elution Buffer (Qiagen material number 1014609), denatured and loaded at a range of 2.25–10 pM for clustering. The PE

Cluster Generation Kit v4 (PE-203-4001) was used with the Illumina Genome Analyzer IIX. The TruSeq PE Cluster Kit v2 – cBot – HS (cat. no. PE-300-2001) and the TruSeq PE Cluster Kit v3 – cBot – HS kit (cat. no. PE-401-3001) were used with the Illumina HiSeq2000/2500.

The resultant flow cells were loaded onto the Illumina GAIIx using the TruSeq Sequencing by Synthesis (SBS) Kit v5 – GA (36-cycle) reagents (cat. no. FC-104-5001). The HiSeq2000/2500 was loaded with the TruSeq SBS Kit – HS (50 cycles) reagents (cat. no. FC-401-1002) or the TruSeq SBS Kit v3 – HS (200 cycles) reagents (cat. no. FC-401-3001).

The RNA-seq libraries were sequenced at a range of 75–100-bp paired-end reads with a 7-bp index using the standard Illumina primers. The sequence intensity files were generated by an instrument using the Illumina Real-Time Analysis software. The intensity files were demultiplexed, and FASTQ files were created using the CASAVA software suite (version dependent upon the date of analysis and the current CASAVA package available; the latest version used was 1.8.2).

The FASTQ files were then processed as in Korpál *et al.*<sup>45</sup>, modified to align simultaneously to the mm10 and GRCh37 genomes and transcriptomes to allow for both human (tumor) and mouse (stroma) alignment. Mutations were called only on GRCh37 (ref. 45). To calculate mutation rates, non-COSMIC mutations that appeared in >50 samples or in dbSNP v138 were removed as probable germline variants, false positives or alignment artifacts.

**Somatic mutation calls.** For mutation calls from TCGA samples, we downloaded somatic single-nucleotide variants (SNVs) identified from whole exome sequencing and copy-number alterations identified by Affymetrix SNP6.0 from the cBio portal<sup>46</sup> and mutation rates across a variety of cancer types from Lawrence *et al.*<sup>47</sup>. For the PDXE and CCLE samples, probable somatic SNVs were identified using transcriptome sequencing and the CCLE<sup>11</sup>, and copy-number alterations were determined using Affymetrix SNP6.0 arrays.

**Pan-cancer panel analysis and SNV and copy-number calls.** Targeted sequencing (SureSelect Human All Exon V4, Agilent) was performed on the DNA of treated and control samples using a HiSeq-2500. The sequencing data was processed as follows: First, sequence reads were aligned with the reference human genome (build hg19) and the Mouse genome (build mm10) using the Burrows-Wheeler Aligner (BWA)<sup>48</sup> to create a BAM file. Next, reads perfectly mapping to the mouse genome were removed, and the initial BAM files were cleaned with Picard to mark PCR duplicates (<http://broadinstitute.github.io/picard/>). Then, the Genome Analysis ToolKit was used for local realignment and base quality score recalibration<sup>49,50</sup>. SNVs were identified with the GATK UnifiedGenotyper in multi-sample mode and copy-number variants called using ExomeCNV. SNVs were retained only if at least five supporting reads were identified. SNVs were annotated with dbSNP v.138, COSMIC v.68, and SnpEff and copy-number variants were called using ExomeCNV<sup>51</sup>.

GATK was run in multi-sample mode to increase the sensitivity on detecting SNVs at low allele frequency. Single nucleotide variants were annotated using dbSNP (v.138) and COSMIC (v.68). Functional annotation was obtained using the SnpEff algorithm<sup>20</sup>. The following criteria were employed to annotate high quality called SNVs (quality filter = PASS and total depth of coverage over the position  $\geq 100$  reads) observed in the treated samples as *de novo* or pre-existent SNVs in the control sample: first, *de novo* SNVs should be supported by more than five reads with the alternate allele in the treated sample, and by fewer than five reads in the control sample. Second, pre-existing SNVs should be supported by more than five reads with the alternative allele in both the treated and the control samples.

Copy-number calling in the PDX data was performed by obtaining the ratio of coverage between the resistant sample and the respective vehicle sample. An amplified region is expected to have a higher proportion of reads mapping to it, relative to its vehicle, whereas a copy-number loss would have fewer reads relative to its vehicle. We used ExomeCNV to call copy-numbers on the PDX data. ExomeCNV takes depth of coverage files generated by GATK as input and outputs the estimated copy-numbers for all regions of the sample. It calculates the log coverage ratio between the resistant sample and the vehicle, and calls CNVs for each exon individually. We assumed the contamination rate to be zero, and used other default parameters. We attempted to normalize the G+C profile for every library so as to obtain a uniformly distributed coverage for GC.

The normalization was performed by obtaining a running median of coverage values across localized regions. Copy-numbers for individual exons were called using ExomeCNV on the normalized coverage profiles.

Individual exonic CNV values were combined into per gene CNV values by using the median CNV value of all the exons across a gene. Arm level, as well as focal copy-number changes, were called this way. Copy-numbers were not called for non-autosomal chromosomes, owing to a lack of knowledge of the gender of the samples. We verified the CNV values for two of the resistant-vehicle pairs against SNP 6.0 data to very good accuracy levels.

**Pharmacological characterization in cell lines.** Cell culture and pharmacological treatment were done as reported previously<sup>11</sup>. Briefly, cell lines were dispensed into 1,536-well plates (optimized for tissue culture) with a final volume of 5  $\mu$ l and a concentration of 250 cells per well. 12–24 h after plating, 20 nl of each compound dilution series were transferred to the 1,536-well plates (containing the tumor cells) using slotted pins (V&P Scientific). This yielded final drug concentration ranges of 8  $\mu$ M–2.5 nM (8-point dose response assays) by 3.16-fold dilutions and a final DMSO concentration of just less than 0.4%. The cell-compound mixtures were incubated for 72–84 h; afterwards, cell numbers were determined by measuring the amount of ATP per well using Cell Titer Glo (Promega). Luminescence per well was measured using a ViewLux plate reader (Perkin Elmer). Dose-response curves were generated, and the half-maximal inhibitory concentration (IC<sub>50</sub>) was calculated.

**Cell line combination synergy testing.** The identities of the cell lines used in the *in vitro* screens were confirmed using SNP genotyping DNA compared to SNP genotyping DNA derived from the parental cells. The SNP genotyping panel uses a set of 48 intergenic SNPs originally developed by the Broad Institute and previously described<sup>11</sup>. The cells were maintained at 37 °C in 5% CO<sub>2</sub>.

Chemical inhibitors were dissolved into DMSO at a 10 mM stock concentration, and diluted to yield the desired concentrations for each compound used in the screens. Each pairwise combination was tested at multiple concentrations using a matrix in which each drug was added to the cells at five concentrations, with alternate matrix points and six additional matrix points omitted to allow for more efficient screening. 12–24 h after plating, the compound was acoustically transferred in droplets from the source plate to the assay plates containing cells, until the desired concentration was achieved. The cell-compound mixtures were incubated for 72–84 h; afterwards, cell numbers were determined by measuring the amount of ATP per well using Cell Titer Glo (Promega).

Luminescence for each well was measured using a ViewLux plate reader (PerkinElmer). Inhibitory responses were calculated for each treated well relative to the median luminescence level from untreated wells across the same assay plate. Cell proliferation was calculated for every dose combination, as described above. The growth effect of each combination was measured by the inhibition of the cellular response of the treatment relative to the untreated level (vehicle alone). For untreated vehicle and treated levels  $V$  and  $T$ , we calculated a fractional inhibition  $Z = 1 - T/V$ . The inhibition ranges from 0% at the untreated level to 100% when  $T = 0$ . Using the Chalice software, provided by Horizon Discovery Inc<sup>52</sup>, the response of the combination was compared to its single agents, relative to the widely used Loewe model that utilizes drug-with-itself dose-additivity as a reference. The strength of combination effects was calculated using the Synergy score  $S = k \ln f_X \ln f_Y \sum \max(0, Z_{\text{data}}) \max(0, Z_{\text{data}} - Z_{\text{Loewe}})$ , summed over all the measured doses in the matrix<sup>52</sup>. Here  $f_X$  and  $f_Y$  are the threefold dilution factors between adjacent dosing points, and  $k$  is a scaling factor to account for untested matrix data points.  $Z_{\text{data}}$  is the measured response value at a dosing point, and  $Z_{\text{Loewe}}$  is the expected Loewe dose-additive model response at any combined concentration  $(X, Y)$ , calculated from sigmoidal fits to the single-agent response curves in the dose matrix<sup>52</sup>. In order to generate a measure for growth inhibition beyond stasis, day 0 measurements were taken and deducted from the end measurements. The following definitions were introduced to describe the extent of growth inhibition: no inhibition = 0, stasis = 1 and total cytotoxicity = 2. To arrive at these values for the fractional growth inhibition,  $Z(\text{GI}) = 1 - (Y/100)$  was used.  $Y$  was calculated using  $Y = (50 \times (X1 - \text{mean}(X1, \text{AC}) \text{mean}(X2, \text{NC}_T0) + \text{mean}(X2, \text{AC}))/(\text{mean}(X1, \text{NC}) - \text{mean}(X1, \text{AC}) \text{mean}(X2, \text{NC}_T0) + \text{mean}(X2, \text{AC})) \times (X1/X1 + \text{sign}(X1 - X1/X1 \times (\text{mean}(X1, \text{AC}) + \text{mean}(X2, \text{NC}_T0) - \text{mean}(X2, \text{AC})))) + (50 \times (X1 - \text{mean}(X1, \text{AC}) - \text{mean}(X2, \text{NC}_T0) +$

$\text{mean}(X2,AC)/(\text{mean}(X2,NC\_T0) - \text{mean}(X2,AC)) \times (X1/X1 + \text{sign}(X1/X1 \times (\text{mean}(X1,AC) + \text{mean}(X2,NC\_T0) - \text{mean}(X2,AC)) - X1))$  using the following definitions:  $X1$  = individual data point (for example, treated well);  $\text{mean}(X1,AC)$  = AVG of media BG wells;  $\text{mean}(X1,NC)$  = AVG of wells with the vehicle-treated control;  $\text{mean}(X2,NC\_T0)$  = AVG of wells with cells on day 0;  $\text{mean}(X2,AC)$  = AVG of media BG wells on day 0.

**Pharmacokinetics analysis of encorafenib and LEE011 combination.** Female nude mice, aged 7–8 weeks, were used for the PK study. LEE011 at 150 mg/kg, encorafenib at 20 mg/kg or the two in combination was given orally for one dose. After treatment, three mice from each treatment group were euthanized at 1, 4, 8, 16 and 24 h. Blood was collected by cardiac puncture. Plasma samples were obtained by centrifugation, and analyzed by Exploratory Bioanalytical, inVentiv Health Clinical Lab, Princeton, NJ.

45. Korpai, M. *et al.* An F876L mutation in androgen receptor confers genetic and phenotypic resistance to MDV3100 (enzalutamide). *Cancer Discov.* **3**, 1030–1043 (2013).
46. Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2**, 401–404 (2012).
47. Lawrence, M.S. *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218 (2013).
48. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
49. DePristo, M.A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).
50. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
51. Sathirapongsasuti, J.F. *et al.* Exome sequencing-based copy-number variation and loss of heterozygosity detection: ExomeCNV. *Bioinformatics* **27**, 2648–2654 (2011).
52. Lehár, J. *et al.* Synergistic drug combinations tend to improve therapeutically relevant selectivity. *Nat. Biotechnol.* **27**, 659–666 (2009).