Patient-derived xenografts undergo mouse-specific tumor evolution

Uri Ben-David[1], Gavin Ha[1,2], Yuen-Yi Tseng[1], Noah F Greenwald[1,3,4], Coyin Oh[1], Juliann Shih[1,2], James M McFarland[1], Bang Wong[1], Jesse S Boehm[1], Rameen Beroukhim[1–3,5,6,10] & Todd R Golub[1,7–10]

Patient-derived xenografts (PDXs) have become a prominent cancer model system, as they are presumed to faithfully represent the genomic features of primary tumors. Here we monitored the dynamics of copy number alterations (CNAs) in 1,110 PDX samples across 24 cancer types. We observed rapid accumulation of CNAs during PDX passaging, often due to selection of preexisting minor clones. CNA acquisition in PDXs was correlated with the tissue-specific levels of aneuploidy and genetic heterogeneity observed in primary tumors. However, the particular CNAs acquired during PDX passaging differed from those acquired during tumor evolution in patients. Several CNAs recurrently observed in primary tumors gradually disappeared in PDXs, indicating that events undergoing positive selection in humans can become dispensable during propagation in mice. Notably, the genomic stability of PDXs was associated with their response to chemotherapy and targeted drugs. These findings have major implications for PDX-based modeling of human cancer.

Cancer research relies on interrogating model systems that mirror the biology of human tumors. Cell lines cultured from human tumors have been the workhorse of cancer research, but the marked differences between the *in vitro* cell culture environment and the *in vivo* tumor environment raise concerns that these lines may not be fully representative of human tumors. Recently, there have been increasing efforts to use PDXs as models to study drug response[1–4]. These *in vivo* models are assumed to capture the cellular and molecular characteristics of human cancer better than simpler cell-line-based models[1,2].

As the value of PDX models depends on their faithful representation of primary tumors, it is important to assess whether PDXs retain their genomic and phenotypic characteristics throughout propagation. Thus far, the genomic stability of PDX models has primarily been evaluated indirectly, leading to the notion that PDXs are highly stable[3,5,6]. Consistent with this perception, PDX-based studies often involve the analysis of tumors from multiple passages[3]. However, hints that PDXs may be more genomically unstable than assumed have recently begun to emerge[7,8], emphasizing the need for a comprehensive analysis of PDX genomic evolution (**Supplementary Note**).

Here we systematically analyzed landscapes of aneuploidy and large CNAs in PDX models across multiple human cancers. We generated a comprehensive CNA catalog for 1,110 PDX samples from 24 cancer types and used these data to characterize CNA dynamics during PDX derivation and propagation, to study the origin of passaging-acquired CNAs, and to compare PDX genomic stability across cancer types. We also compared the CNA dynamics observed in PDXs to those of newly derived tumor cell lines and cell-line-derived xenografts (CLDXs). Finally, we compared the CNA landscapes of PDXs to those of human primary and advanced tumors. We found that, despite an overall similarity, the CNA landscapes of PDXs diverge substantially from those of their parental tumors during passaging. We discuss the potential implications of this divergence, including its effect on therapeutic response.

RESULTS

Generating a catalog of aneuploidy and CNAs in PDXs

To enable a comprehensive analysis of aneuploidy and CNAs in PDXs, we created an integrated CNA data set representing 1,110 PDXs. We first assembled data from DNA-based copy number measurements across multiple PDX passages, using published SNP arrays, comparative genomic hybridization (CGH) arrays and DNA sequencing data. Unfortunately, such DNA copy number data were only available for 177 PDX samples from five studies—too few to support a comprehensive analysis of CNA stability (**Supplementary Table 1**)[6,7,9–11]. In contrast, gene expression profiles were available for 933 PDX samples collected from 511 PDX models across 17 studies (**Supplementary Table 1**)[3,5,6,10–23]. To reconstruct chromosomal aneuploidy and large (>5-Mb) CNAs from these expression profiles, we used previously described computational inference algorithms that accurately identify CNAs on the basis of induced coordinated gene expression changes[24–26]. We validated the accuracy of this approach by analyzing PDXs for which both gene expression and SNP array data were available

[1]Cancer Program, Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA. [2]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. [3]Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. [4]Department of Neurosurgery, Brigham and Women's Hospital, Boston, Massachusetts, USA. [5]Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA. [6]Department of Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA. [7]Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. [8]Department of Pediatrics, Harvard Medical School, Boston, Massachusetts, USA. [9]Howard Hughes Medical Institute, Chevy Chase, Maryland, USA. [10]These authors jointly directed this work. Correspondence should be addressed to R.B. (rameen_beroukhim@dfci.harvard.edu) or T.R.G. (golub@broadinstitute.org).

(**Supplementary Fig. 1**, **Supplementary Table 2** and **Supplementary Note**). Our final data set comprised CNA data for 1,110 PDX samples from 543 unique PDX models across 24 cancer types (**Fig. 1a** and **Supplementary Data 1** and **2**). For 342 of these PDX models, data were available from both the primary tumor and its derived PDX model(s), or from multiple PDX passages, thus enabling an analysis of tumor evolution (**Fig. 1a** and **Supplementary Table 1**).

We found the CNA landscapes of PDXs to be highly similar to those of their respective tumor types in The Cancer Genome Atlas (TCGA) (mean Pearson's r = 0.79; **Fig. 1b** and **Supplementary Fig. 2**), consistent with prior reports[6,8,10,11]. This confirms that, at the cohort level, PDX models are generally genomically representative of primary human tumors.

Tracking CNA dynamics during PDX derivation and propagation

We set out to follow CNA dynamics in individual PDX models, to assess their stability as well as their similarity to the tumors from which the models were derived. For each model, the earliest passage (in most cases, passage 0 (P0) or P1) was compared to later passages (range, P1–P16; median, P3) to determine the changes that occurred throughout passaging. A representative example of PDX model evolution is shown in **Figure 1c**.

We found that large (>5-Mb) CNAs arose rapidly in PDXs: 60% of the PDX models acquired at least one large chromosomal aberration within a single *in vivo* passage, and 88% acquired at least one large aberration within four passages (**Supplementary Fig. 3a**). The CNA landscape of PDX models thus gradually shifted away from those of the original primary tumors, with a median of 12.3% of the genome (range, 0–58.8%) affected by model-acquired CNAs within four passages (**Fig. 1d**). Of note, similar results were obtained using three different definitions of CNA prevalence—the proportion of the genome affected by CNAs (CNA fraction), the number of discrete events, and the proportion of altered genes (**Fig. 1d** and **Supplementary Fig. 3b,c**)—thereby highlighting the robustness of this finding.

There was no significant change in the overall number of CNAs throughout passaging (**Supplementary Fig. 3d**), indicating equal rates of emergence of new events and disappearance of existing ones. A median of 35.6% of the genome was affected by CNAs, consistent with prior estimates in primary tumors[27] (**Fig. 1e** and **Supplementary Fig. 3c**). The disappearance of CNAs during passaging was not due to changes in tumor sample purity (for instance, contamination with mouse tissue might dilute the CNA signal), as other primary events were readily detected at similar signal strength. Notably, approximately one of six large CNAs identified in PDX models at P4 was not observed in the respective primary tumor. A similar proportion of primary clonal CNAs could no longer be detected in PDXs by P4. This observed tumor evolution was not limited to large CNAs, but also affected mutations in cancer-related genes (**Supplementary Fig. 4** and **Supplementary Note**). We conclude that individual PDX models can quickly genomically diverge from their parental primary tumors.

Selection of preexisting subclones underlies CNA dynamics

Our observation that CNAs were often gained or lost during PDX passage might be explained by expansion of preexisting subclones, the acquisition and expansion of *de novo* events, or a combination of these mechanisms. Several lines of evidence suggest that clonal selection of preexisting subclones has a major role in shaping the CNA landscape of PDXs. First, CNAs accumulated with each passage, but their acquisition rate decreased over time (**Fig. 2a**). Second, apoptosis expression signatures gradually decreased with PDX passage number while

signatures of proliferation increased, in line with clonal selection of fitter clones (**Fig. 2b**). Third, the rates of model-acquired CNAs were similar in PDXs from primary tumors and from metastases (**Fig. 2c**), despite metastasis-derived PDXs being more aneuploid and exhibiting higher expression of genes associated with chromosomal instability (**Supplementary Fig. 5a,b**)[28]; this suggests that model-acquired CNAs predominantly result from clonal dynamics, rather than from genomic instability.

If our hypothesis that acquired CNAs are the result of positive biological selection of existing minor subclones is correct, one would expect that the same minor clones would be enriched in multiple independent grafts of the same tumor (transplanted into 'sibling' P0 mice). Five such PDX pairs (representing breast, lung, pancreas and skin cancer PDXs) were available for analysis (**Fig. 2d**). The similarity in model-acquired CNAs between sibling PDXs was indeed significantly higher than the similarity between lineage-controlled 'non-sibling' PDXs ($P < 1 \times 10^{-5}$, stratified bootstrap test; **Fig. 2e**). This finding suggests that clonal evolution occurs through directional selection of preexisting subclones, consistent with observations in breast and hematopoietic cancers[7,29].

To further test this idea, we analyzed loss of heterozygosity (LOH) events. Because LOH is an irreversible event at the cellular level, an LOH 'reversion' at the population level can only be explained by expansion of cells that did not undergo LOH in the first place. We queried previously published whole-genome sequencing data from 15 breast cancer pairs of primary tumors and PDXs[7] and identified 5 cases of LOH reversion (**Fig. 2f** and **Supplementary Fig. 6**). These analyses thus confirm that rare preexisting subclones, which are not readily detected in the primary tumor, can expand and become the dominant clone in PDXs.

We conclude that CNA dynamics are strongest during engraftment and the first few *in vivo* passages, continue at a reduced rate throughout model propagation, and result primarily from selection of preexisting subclones.

The degree of genomic instability in PDXs mirrors that of primary tumors

As human cancer types differ considerably in their CNA prevalence and acquisition rate (also referred to as the degree of genomic instability, or DGI), we next compared CNA dynamics in PDXs across cancer types. We found that the rate of model-acquired CNAs varied significantly ($P = 0.001$, Wilcoxon rank-sum test, when comparing the most stable to the least stable tumor type), with brain tumor PDXs being the most stable and gastric tumor PDXs being the least stable (with a median of 0% and 4.2% of the genome affected by CNAs per passage, respectively) (**Fig. 3a**).

We therefore asked whether this spectrum of PDX aneuploidy reflects the aneuploidy levels of human cancer types. We measured aneuploidy in TCGA data according to two metrics. First, we used the percentage of samples with whole-genome duplication (WGD)[27]. Across seven tissues for which data were available from both TCGA and PDX data sets, the CNA acquisition rate in PDXs correlated strongly with WGD prevalence in TCGA samples (Spearman's ρ = 093, $P = 0.003$; **Fig. 3b** and **Supplementary Fig. 7a**). Second, we found that the median number of PDX-acquired arm-level CNAs and the median number of arm-level CNAs acquired during tumor development in TCGA samples were correlated across ten different cancer types (Spearman's $\rho = 0.76$, $P = 0.010$; **Fig. 3c**). We thus conclude that DGI variation among PDX tumor types parallels that of primary tumors.

As we found that expansion of preexisting subclones had a major role in shaping the CNA landscape of PDXs, we next examined whether the tissue-specific rate of CNA dynamics correlates with the
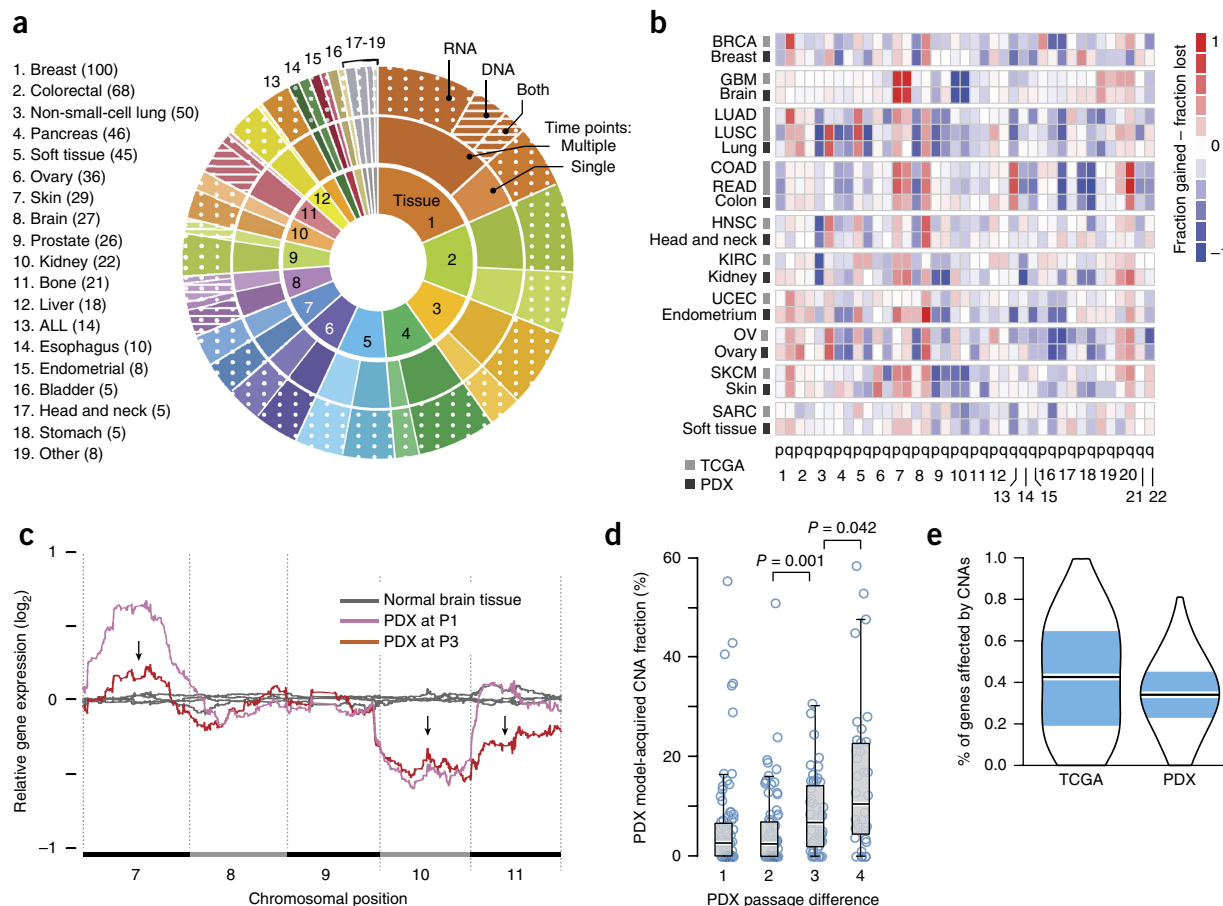
a

1. Breast (100)
2. Colorectal (68)
3. Non-small-cell lung (50)
4. Pancreas (46)
5. Soft tissue (45)
6. Ovary (36)
7. Skin (29)
8. Brain (27)
9. Prostate (26)
10. Kidney (22)
11. Bone (21)
12. Liver (18)
13. ALL (14)
14. Esophagus (10)
15. Endometrial (8)
16. Bladder (5)
17. Head and neck (5)
18. Stomach (5)
19. Other (8)

Figure 1 The landscape of aneuploidy and copy number alterations in PDXs. (**a**) Distribution of cancer types in our PDX data set ($n = 543$ unique models). In the inner circle, models are divided by lineage: each cancer type is denoted by color and number. In the middle circle, models are divided by the number of time points analyzed: multiple time points are denoted by a darker color and enable PDX evolution to be followed throughout *in vivo* propagation. In the outer circle, models are divided by the biological material from which CNAs were inferred: DNA (stripes), RNA (dots) or both (stripes and dots). (**b**) Heat map comparing the landscapes of arm-level CNAs in lineage-matched PDXs and primary TCGA tumors, showing an overall high degree of concordance (mean Pearson's $r = 0.79$). The color of each chromosome arm corresponds to the difference between the fraction with gains and the fraction with losses of that arm. (**c**) A representative example of PDX model evolution. Shown are gene expression moving-average plots for normal brain tissue (gray), a GBM PDX model at P1 (pink) and a GBM PDX model at P3 (red), demonstrating the disappearance of trisomy 7, the retention of monosomy 10 and the emergence of monosomy 11 within two *in vivo* passages. (**d**) Gradual evolution of CNA landscapes throughout PDX passaging. Box plots present the model-acquired CNA fraction as a function of the number of passages between measurements. Bar, median; box, 25th to 75th percentile (interquartile range, IQR); whiskers, data within 1.5 times the IQR; circles, all data points. *P* values indicate significance from a Wilcoxon rank-sum test. (**e**) Violin plots present the proportion of genes affected by CNAs in TCGA and PDX tumor samples (all tissue types combined), showing an overall similarity between the two data sets. Bar, median; colored rectangle, 25th to 75th percentile; the width of each violin indicates frequency at the corresponding CNA fraction level.

degree of heterogeneity that characterizes each cancer type. Indeed, the CNA acquisition rate in PDXs correlated well with the median number of clones of the respective primary tumor type[30] across six cancer types that could be matched (Spearman's $\rho = 0.82$, $P = 0.044$; **Supplementary Fig. 7b**). Interestingly, melanoma had the highest degree of intratumoral heterogeneity, but only a moderate level of DGI in PDXs, and was therefore the only cancer type that substantially deviated from the observed correlation; the correlation became even stronger when melanoma was removed from the analysis (Spearman's $\rho = 1$, $P < 2.2 \times 10^{-16}$; **Fig. 3d** and **Supplementary Note**).

The combined results of these analyses suggest that PDX models have characteristic tissue-specific levels of CNA dynamics that correspond both to the DGI and to the degree of heterogeneity of the respective primary tumor type. As genetic heterogeneity is closely associated with aneuploidy levels and DGI in primary tumors[30–32], either of these factors—or both together—might explain the observed correlations.

CNA recurrence analysis identifies distinct selection pressures in PDXs and in primary tumors

A key question is whether the clonal dynamics observed in PDXs mimic those observed in human patients. To address this question, we asked whether recurrent arm-level genetic events that are observed in human tumors remain under selection pressure when the tumors are transplanted into mice; loss of these signature events would signal key differences in selection pressures between human and mouse hosts.

We identified 61 recurrent arm-level CNAs across TCGA tumor types and followed them in PDXs. Surprisingly, events that were recurrent in the TCGA data set (reflecting positive selection in humans) tended to disappear throughout PDX passaging. Specifically, among lineage-matched PDXs, we observed 116 model-acquired events that were in the opposite direction (gain instead or loss, or vice versa) to the recurrent TCGA CNAs, and only 79 model-acquired events were in the same direction ($P = 0.01$, McNemar's test **Fig. 4a**).
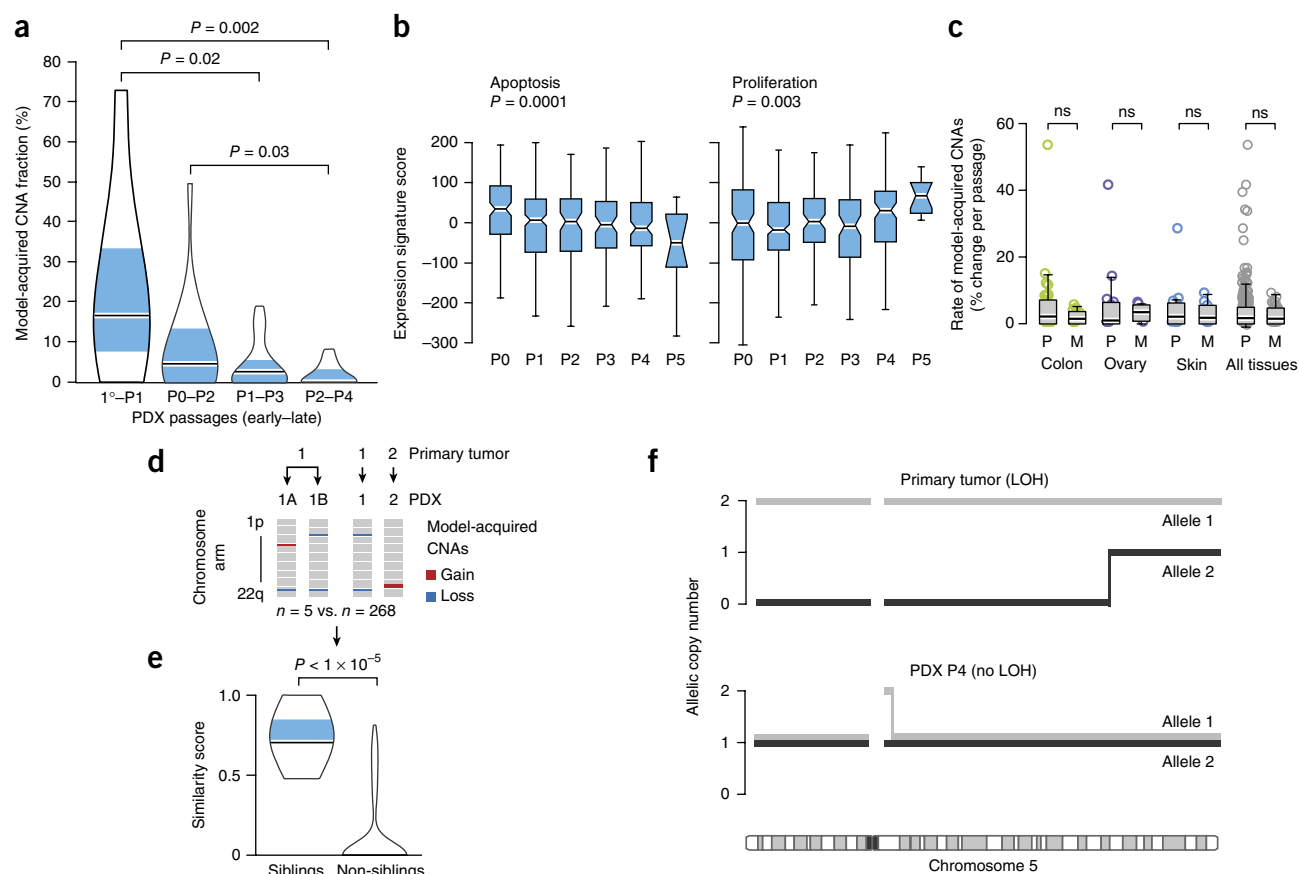
Figure 2 Selection of preexisting subclones underlies CNA dynamics. (**a**) The rate of model-acquired CNAs decreases with PDX passaging. Violin plots present the fraction of CNAs acquired within two *in vivo* passages as a function of passage number. *P* values indicate significance from a Wilcoxon rank-sum test. 1°, primary tumor. (**b**) Apoptosis decreases and proliferation increases with PDX passaging. Box plots present apoptosis (left) and proliferation (right) gene expression signature scores as a function of passage number. Bar, median; box, 25th to 75th percentile (interquantile range, IQR); whiskers, data within 1.5 times the IQR. *P* values indicate significance from a Kruskal–Wallis test. (**c**) Similar CNA acquisition rates in PDXs from primary tumors and from metastases. Box plots present the rate of model-acquired CNAs as a function of tumor source (P, primary; M, metastasis), across three available tissue types. Box-plot features are as in **b**. ns, not significant (Wilcoxon rank-sum test). (**d**) Schematic showing the calculation of pairwise similarity scores for PDX models coming from the same primary tumor but propagated independently in mice (sibling PDXs; *n* = 5) and for PDX models coming from distinct primary tumors (non-sibling PDXs; *n* = 268). (**e**) Sibling PDXs tend to acquire more similar aberrations than lineage-matched non-sibling PDXs. Violin plots present the similarity scores of sibling and non-sibling PDXs. *P* value indicates significance from a lineage-controlled permutation test. (**f**) Alleles that seem to have been lost in primary tumors can 'reappear' in PDXs, demonstrating expansion of rare preexisting subclones throughout PDX propagation. Plots present the copy number of both chromosome 5 alleles in a primary tumor and its derived PDX. LOH is identified in the primary tumor along most of chromosome 5, but both alleles are detected in a 1:1 ratio in the PDX derived from that primary tumor.

We identified 12 recurrent events in TCGA samples that were preferentially lost throughout PDX passaging across five cancer types (glioblastoma multiforme (GBM), breast, lung, colon and pancreatic cancers; **Fig. 4b** and **Supplementary Fig. 8**). Events that tend to disappear throughout PDX propagation should be less prevalent in high-passage than in low-passage PDXs. Indeed, 9 of the 12 events that tended to disappear in PDXs, including the hallmark gains of chromosomes 1q and 8q in breast cancer and chromosome 7 in GBM and the hallmark losses of chromosome 10 in GBM and chromosome 4q in non-small-cell lung cancer, were less common in high-passage PDXs (**Fig. 4c** and **Supplementary Fig. 9**).

Taken together, these data demonstrate that PDXs can lose recurrent chromosomal aberrations that are believed to have causal roles in the development of human tumors. This suggests that the selection pressures that led to the acquisition and retention of these hallmark CNAs in patients may no longer exist in the mouse model environment.

Distinct CNA dynamics during tumor progression in PDXs and in human patients

To further assess whether the clonal dynamics observed in PDXs are indeed fundamentally different from those occurring during tumor evolution in patients, we next analyzed CNA dynamics during the progression of primary tumors into advanced disease (metastases and recurrences). We predicted that, if recurrent CNAs tend to disappear in PDXs owing to mouse-specific selection pressures, this trend should not be found during tumor evolution in humans.

To address this question, we compiled CNA data from 306 tumor samples of matched primary tumors (*n* = 132) and their derived metastases or recurrences (*n* = 174) across five cancer types represented in our PDX data set (colon, lung, endometrial, brain, and head and neck)[33–40] (**Supplementary Table 3**). By comparing each metastasis or recurrence to its matched primary tumor, we found that tumor progression in patients was associated with a dramatic shift in CNA landscape, with a median of 18.2% of the genome (range, 0–95.4%) affected by progression-acquired
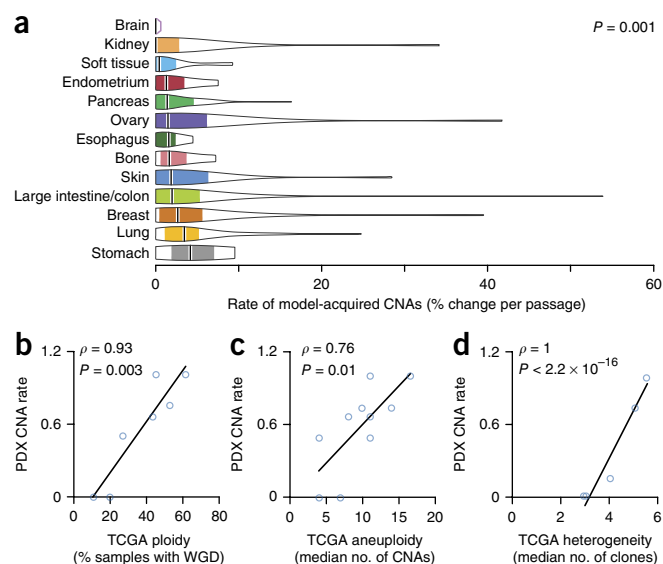
Figure 3 The genomic instability of PDXs mirrors that of primary tumors. (**a**) The DGI of PDXs is cancer type specific. Violin plots present the rate of CNA acquisition throughout PDX propagation in 13 cancer types, for which data were available from at least three PDX models. The *P* value indicates significance from a Wilcoxon rank-sum test. (**b**) The DGI of PDXs and that of primary tumors correlate extremely well. In PDXs, tissue DGI was defined as the median number of CNAs per passage. In TCGA tumors, tissue DGI was defined as the fraction of samples with WGD. (**c**) This correlation holds when the tissue DGI is defined, both for PDXs and TCGA tumors, by the median number of arm-level CNAs. (**d**) The DGI of PDXs also correlates extremely well with the intratumoral heterogeneity of primary tumors (excluding skin tissue). The DGI of PDXs was defined as the median number of arm-level CNAs per passage. The heterogeneity of primary tumors was defined as the median number of clones per tumor. Spearman's ρ values and *P* values indicate the strength and significance of the correlations, respectively.

CNAs (**Fig. 4d**). This change was greater, on average, than the change observed in PDXs ($P < 1 \times 10^{-5}$, stratified bootstrap test; **Fig. 1d**), likely reflecting the much longer time periods (often years) between paired resections and the strong treatment-associated selection pressures.

However, in contrast to the disappearance of recurrent CNAs during PDX passaging, the opposite was observed during tumor evolution in patients: recurrent CNAs more often emerged than disappeared during tumor progression. We observed 158 progression-acquired events that were in the same direction as the recurrent TCGA CNAs, and only 101 model-acquired events were in the opposite direction ($P = 0.0005$, McNemar's test; **Fig. 4a**). The relative proportion of recurrent CNAs that emerged, in comparison to the proportion that disappeared, during tumor evolution was significantly different between PDXs and advanced human disease ($P = 0.0001$ Chi-squared test; **Fig. 4a**). Therefore, these data further demonstrate that distinct selection pressures shape the CNA landscapes of tumors during their evolution in humans and in mouse hosts.

Genomic instability of PDXs is comparable to that of cell lines and CLDXs

PDXs are generally considered to reflect primary human tumors more faithfully than cell lines, owing to the artificial cell culture environment[1,41]. However, the subcutaneous environment in immunodeficient mice also differs considerably from the natural human host. To address the assumption that PDXs better preserve the fidelity to
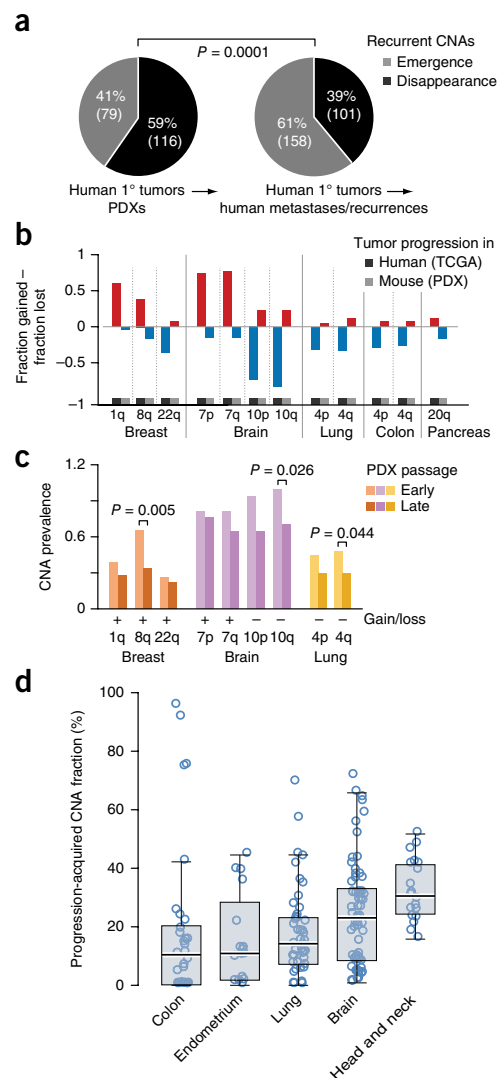

Figure 4 Tumor evolution of PDXs diverges from that of primary tumors. (**a**) Left, recurrent arm-level TCGA CNAs tend to disappear throughout PDX passaging. The pie chart presents the number of model-acquired events that were in the opposite direction to the recurrent TCGA CNAs versus the number of events in the same direction. Right, recurrent arm-level TCGA CNAs tend to emerge throughout tumor progression in patients. The pie chart presents the number of progression-acquired events that were in the opposite direction to the recurrent TCGA CNAs versus the number of events in the same direction. *P* value indicates significance from a chi-squared test. (**b**) Opposite propensities for gain and loss in human tumors and PDX models. Bar plots present the difference between the fraction with gain and the fraction with loss for 12 recurrent TCGA arm-level CNAs. The PDX fractions represent model-acquired CNAs rather than the absolute prevalence of these events. (**c**) Recurrent TCGA arm-level CNAs are more common in early-passage PDXs than in late-passage PDXs. Bar plots present the absolute prevalence of each event in the relevant cancer type. *P* values indicate significance from a Fisher's exact test. (**d**) Evolution of CNA landscapes during tumor progression to advanced disease. Box plots present the progression-acquired CNA fraction in the five tumor types analyzed. Bar, median; box, 25th to 75th percentile; whiskers, data within 1.5 times the IQR; circles, all data points.

human tumors, we compared the CNA dynamics of PDXs *in vivo* to those of cell lines *in vitro*.

We found that the prevalence of model-acquired CNAs in newly derived cell lines was similar to that in PDXs. We analyzed the CNA
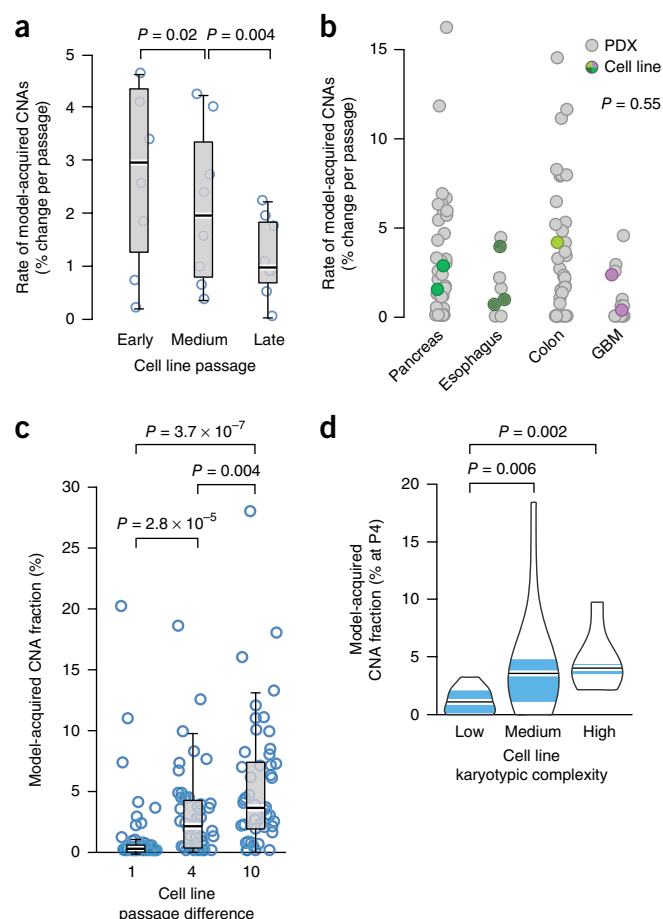
Figure 5 The genomic instability of PDXs is comparable to that of cell lines and CLDXs. (**a**) The rate of CNA acquisition decreases with cell line passaging. Box plots present the rate of CNA acquisition as a function of *in vitro* passage number. Bar, median; box, 25th to 75th percentile; whiskers, data within 1.5 times the IQR; circles, all data points. *P* values indicate significance from a Wilcoxon rank-sum test. (**b**) PDXs and newly derived cell lines have similar rates of CNA acquisition. Dot plots present the distribution of model-acquired CNA rates across four available cancer types. The *P* value indicates lack of significance from a lineage-controlled permutation test. (**c**) Gradual evolution of CNA landscapes throughout CLDX passaging. Box plots present model-acquired CNA fraction as a function of the number of passages between measurements. Box-plot features are as in **a**. *P* values indicate significance from a Wilcoxon rank-sum test. (**d**) The CNA acquisition rate of CLDXs is associated with the numerical karyotypic complexity of the parental cell lines. Violin plots present the fraction of CNAs acquired by P4 as a function of numerical karyotypic complexity. *P* values indicate significance from a Wilcoxon rank-sum test.

landscapes of 38 samples of nine new cell lines derived in our laboratory for five cancer types (colon, GBM, pancreas, esophagus and thyroid; **Supplementary Table 4**). Similar to our observations with PDXs, newly derived cell lines acquired CNAs with passaging, and their CNA landscape gradually shifted away from that of the earliest passage (**Supplementary Fig. 10a**). As seen in PDXs, changes occurred mostly during the first few passages, and the rate of model-acquired CNAs decreased throughout propagation in culture (**Fig. 5a**). Notably, while CNA rates (defined as the fraction of the genome affected by model-acquired CNAs per passage) varied considerably among cell lines, they fell well within the range seen in PDXs, in a lineage-matched comparison (*P* = 0.55, stratified bootstrap test;

Fig. 5b). These results were recapitulated with newly derived cell lines from three independent studies of GBM[42,43], kidney[44], and head and neck[22] cancers (*n* = 31; **Supplementary Fig. 10b** and **Supplementary Table 4**), suggesting that clonal selection is not unique to a particular cell line propagation method.

Next, we compared CNA dynamics between PDXs and CLDXs. To assess CNA dynamics during the *in vivo* propagation of established cancer cell lines, we turned to the National Cancer Institute (NCI) MicroXeno project, which profiled gene expression of 49 human cancer cell lines across multiple *in vivo* passages[45]. We used the same computational algorithms[24–26] that we applied to the PDX models to infer aneuploidy and CNAs from these gene expression profiles, resulting in 823 copy number profiles (**Supplementary Data 3** and **4**). We found that CNAs accumulated with *in vivo* passaging of CLDXs (**Fig. 5c**) and that the DGI of CLDXs correlated with the karyotypic complexity of their parental cell lines (**Fig. 5d** and **Supplementary Fig. 10c**), in line with what we observed in PDXs. However, the rate of CNA acquisition was lower in CLDXs: within four passages, the median model-acquired CNA fraction was 2.2% in CLDXs, as compared to 12.3% in PDXs (*P* = 1.6 × 10⁻⁶, Wilcoxon rank-sum test), likely reflecting the reduced heterogeneity of established cell lines in comparison to primary tumors at the time of xenograft initiation[46].

Taken together, our data from three types of cancer models (PDXs, cell lines and CLDXs) demonstrate that switching the environment in which cancer cells are propagated results in CNA dynamics that gradually alter the CNA landscape. All cancer models are subject to such clonal selection. PDXs do not appear to be spared.

CNA dynamics in PDXs and cell lines may affect drug response

It is conceivable that, while PDXs undergo selection in the mouse, such selection is unimportant with respect to modeling therapeutic response. To address this possibility, we turned to a data set of PDXs with accompanying data on responses to both genotoxic chemotherapies and targeted therapeutics[3].

Both very low and very high levels of aneuploidy have been associated with response to genotoxic drugs and improved patient survival[28,30,47,48]. Notably, CNA acquisition rate (DGI), rather than absolute levels of aneuploidy, determines sensitivity to further perturbation of chromosome segregation[49]. We therefore determined the DGI of PDX models and asked whether it similarly predicts response to chemotherapies. For three of five chemotherapies tested, extreme (either very low or very high) levels of DGI—but not overall aneuploidy levels—were associated with favorable therapeutic response (**Fig. 6a**): dacarbazine in skin cancer PDXs, paclitaxel in lung cancer PDXs, and combined abraxane and gemcitabine therapy in pancreatic cancer PDXs (*P* = 0.04, 0.014, and 0.006, respectively, Wilcoxon rank-sum test). The biological activity and clinical efficacy of these drugs were previously linked to chromosomal instability[50–54]. PDXs thus recapitulate the correlation observed in patients between genomic instability and response to cytotoxic chemotherapies.

We next asked whether particular model-acquired CNAs might affect PDX responses to targeted therapies, given that specific recurrent arm-level or whole-chromosome aberrations have been reported to alter the cellular response to certain drugs[55–57]. We evaluated the association between PDX response to targeted therapies and the presence or absence of individual arm-level CNAs, focusing on the 12 driver CNAs found to be selected against during PDX passaging. We identified three statistically significant drug response–CNA associations (**Fig. 6b**): chromosome 4 loss was associated with increased response of colon cancer PDXs to the TNKS inhibitor LCJ049 (*P* = 0.005, *q* = 0.04 for 4p loss and *P* = 0.00002, *q* = 0.0003 for 4q loss);
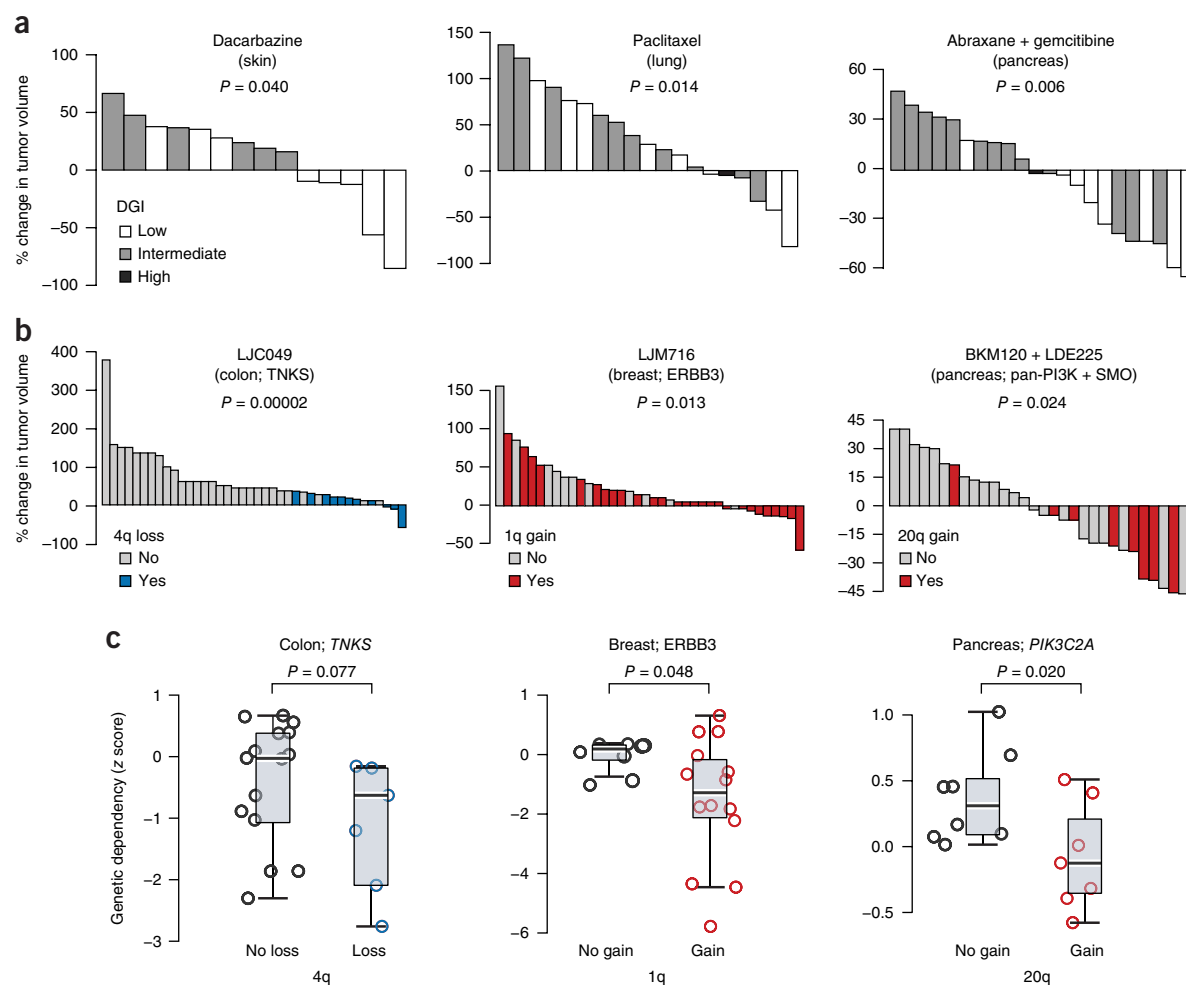
Figure 6 CNA dynamics affect PDX drug response. (**a**) Extreme levels of genomic instability are associated with better therapeutic response to chemotherapies. Waterfall plots present the response to dacarbazine ($n = 14$), paclitaxel ($n = 19$), and the combination of abraxane and gemcitabine ($n = 22$) in skin, lung and pancreatic cancer PDXs, respectively. P values indicate significance from a Wilcoxon rank-sum test. (**b**) The status of recurrent arm-level CNAs is associated with response to targeted therapies. Waterfall plots present the response to the TNKS inhibitor LCJ049 ($n = 40$), the ERBB3 inhibitor LJM716 ($n = 38$), and the combination of the PI3K inhibitor BKM120 and the SMO inhibitor LDE225 ($n = 31$). P values indicate significance from a Wilcoxon rank-sum test. (**c**) The status of recurrent arm-level CNAs is associated with genetic depletion of the genes targeted by the identified drugs. Box plots present dependency scores from RNA interference (RNAi)-mediated knockdown of the indicated genes. Colon cancer cell lines with chromosome 4q loss are more sensitive to knockdown of *TNKS*, breast cancer cell lines with chromosome 1q gain are more sensitive to knockdown of *ERBB3* and pancreatic cancer cell lines with chromosome 20q gain are more sensitive to knockdown of multiple PI3K genes, including *PIK3C2A*. Bar, median; box, 25th to 75th percentile; whiskers, data within 1.5 times the IQR; circles, all data points. P values indicate significance from a Wilcoxon rank-sum test.

chromosome 20q gain was associated with increased response of pancreatic cancer PDXs to combination of the PI3K inhibitor BKM120 and the SMO inhibitor LDE225 ($P = 0.024$, $q = 0.19$); and chromosome 1q gain was associated with increased response of breast cancer PDXs to the ERBB3 inhibitor LJM716 ($P = 0.013$, Wilcoxon rank-sum test, $q = 0.23$). These results indicate that it is not unusual for CNAs (and presumably other genomic events) that undergo negative selection in the mouse host to be associated with changes in sensitivity to specific targeted agents. Such associations may affect the stability of PDX drug response.

To further assess the potential functional relevance of model-acquired chromosomal changes, we turned to the Cancer Cell Line Encyclopedia (CCLE) project and its associated data sets of genomic features, genetic dependencies and drug responses[58–60]. For the 12 recurrent CNAs, we compared cell lines with and without the aberration with regard to gene expression profiles, genetic

dependencies and drug sensitivity (controlling for cell lineage). In line with the PDX drug response data, colon cancer cell lines with chromosome 4q loss were more sensitive to knockdown of *TNKS* ($P = 0.077$), breast cancer cell lines with chromosome 1q gain were more sensitive to knockdown of *ERBB3* ($P = 0.048$) and pancreatic cancer cell lines with chromosome 20q gain were more sensitive to knockdown of multiple PI3K genes ($P = 0.020$, 0.076, 0.005 and 0.014 for *PIK3C2A*, *PIK3CD*, *PIK3CG* and *PIK3R2*, respectively, Wilcoxon rank-sum test; **Fig. 6c** and **Supplementary Fig. 11a**). The analysis of cell lines also showed that arm-level CNAs were commonly associated with significant up- or downregulation of genes residing within the aberrant arm and that these changes were significantly associated with cell line genetic dependencies and pharmacological responses (**Supplementary Figs. 11** and **12**, **Supplementary Tables 5** and **6**, and **Supplementary Note**). We conclude that recurrent arm-level CNAs are associated with drug response, and their

gradual disappearance throughout PDX propagation may therefore be functionally important.

DISCUSSION

The ability to directly transfer human tumors into mice and propagate them for multiple passages *in vivo* offers unique opportunities for cancer research and drug discovery, making PDXs a valuable cancer model. Like any other model system, however, understanding its limitations—and the ways in which it differs from human tumors in their natural environment—is required for optimal application. Our findings suggest that the genomic instability of PDXs has been underappreciated: the CNA landscapes of PDXs change continuously, and so their propagation distances them from the primary tumors from which they were derived (**Supplementary Note**). Indeed, comparison of PDXs to newly derived cell lines showed that PDXs do not necessarily capture the genomic landscape of primary tumors better than cell lines, in contrast to common belief[1] (**Supplementary Note**). The similar CNA rates suggest that multiple cell line models from a single primary tumor may capture more of the original genomic landscape and its heterogeneity than a single PDX model. Moreover, the acquisition of genetic alterations throughout model propagation is unlikely to be restricted to CNAs (**Supplementary Fig. 13** and **Supplementary Note**).

As our analysis was based on bulk-population measurements, the cellular origin of each model-acquired event could not be definitively determined. Our study strongly suggests that preexisting alterations have a major role in model-acquired CNAs, especially at the early stages of PDX derivation and propagation, but that *de novo* events also contribute to genomic instability (**Supplementary Fig. 14** and **Supplementary Note**). Regardless of their origin, we found that CNAs often quickly became fixed in the population; a single *in vivo* passage sometimes rendered an undetected chromosomal aberration readily detectable at the population level. Such strong clonal dynamics suggest that distinct selection pressures between patients and animal models result in divergent tumor evolution trajectories (**Supplementary Note**).

Recent genomic analyses have shown that metastases evolve independently from primary tumors, often representing common ancestral subclones that are not detected in individual biopsies of the primary tumor. In contrast to the considerable heterogeneity between primary tumors and metastases, distinct metastatic sites tend to be relatively homogeneous[33,61,62]. Our findings from PDXs echo those from metastasis. The dominant clones in PDXs often come from minor subclones of the primary tumor, and PDXs that originate from the same primary tumor (the equivalence of multiple metastatic sites) tend to evolve in similar trajectories. It has been suggested that caution is required when inferring the genetic composition of metastatic disease from a primary tumor biopsy, and vice versa[33,61,62]; similarly, we propose that the genetic composition of a PDX tumor may differ from that of its matched primary tumor, potentially in therapeutically meaningful ways (**Supplementary Note**). This should be considered when using PDXs as avatar models for personalized medicine or to identify biomarkers of drug sensitivity.

METHODS

Methods, including statements of data availability and any associated accession codes and 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.

AUTHOR CONTRIBUTIONS

U.B.-D. conceived the project, collected the data and carried out the analyses. G.H., N.F.G. and J.M.M. assisted with computational analyses. Y.-Y.T. and J.S.B. provided cell line data. C.O. assisted with the copy number analysis of cell lines. B.W. assisted with figure design and preparation. J.S. assisted with the copy number analysis of TCGA samples. R.B. and T.R.G. directed the project. U.B.-D., R.B. and T.R.G. wrote the manuscript.

1. Tentler, J.J. *et al.* Patient-derived tumour xenografts as models for oncology drug development. *Nat. Rev. Clin. Oncol.* **9**, 338–350 (2012).
2. Siolas, D. & Hannon, G.J. Patient-derived tumor xenografts: transforming clinical samples into mouse models. *Cancer Res.* **73**, 5315–5319 (2013).
3. Gao, H. *et al.* High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response. *Nat. Med.* **21**, 1318–1325 (2015).
4. Hidalgo, M. *et al.* Patient-derived xenograft models: an emerging platform for translational cancer research. *Cancer Discov.* **4**, 998–1013 (2014).
5. Zhang, X. *et al.* A renewable tissue resource of phenotypically stable, biologically and ethnically diverse, patient-derived human breast cancer xenograft models. *Cancer Res.* **73**, 4885–4897 (2013).
6. Hodgson, J.G. *et al.* Comparative analyses of gene copy number and mRNA expression in glioblastoma multiforme tumors and xenografts. *Neuro-oncol.* **11**, 477–487 (2009).
7. Eirew, P. *et al.* Dynamics of genomic clones in breast cancer patient xenografts at single-cell resolution. *Nature* **518**, 422–426 (2015).
8. Bruna, A. *et al.* A biobank of breast cancer explants with preserved intra-tumor heterogeneity to screen anticancer compounds. *Cell* **167**, 260–274.e22 (2016).
9. Kresse, S.H., Meza-Zepeda, L.A., Machado, I., Llombart-Bosch, A. & Myklebost, O. Preclinical xenograft models of human sarcoma show nonrandom loss of aberrations. *Cancer* **118**, 558–570 (2012).
10. 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).
11. Lin, D. *et al.* High fidelity patient-derived xenografts for accelerating prostate cancer discovery and drug development. *Cancer Res.* **74**, 1272–1283 (2014).
12. Huynh, H. *et al.* Bevacizumab and rapamycin induce growth suppression in mouse models of hepatocellular carcinoma. *J. Hepatol.* **49**, 52–60 (2008).
13. Zhang, X. *et al.* SRRM4 expression and the loss of REST activity may promote the emergence of the neuroendocrine phenotype in castration-resistant prostate cancer. *Clin. Cancer Res.* **21**, 4698–4708 (2015).
14. Oakes, S.R. *et al.* Sensitization of BCL-2-expressing breast tumors to chemotherapy by the BH3 mimetic ABT-737. *Proc. Natl. Acad. Sci. USA* **109**, 2766–2771 (2012).
15. Jäger, W. *et al.* Patient-derived bladder cancer xenografts in the preclinical development of novel targeted therapies. *Oncotarget* **6**, 21522–21532 (2015).
16. Monsma, D.J. *et al.* Using a rhabdomyosarcoma patient-derived xenograft to examine precision medicine approaches and model acquired resistance. *Pediatr. Blood Cancer* **61**, 1570–1577 (2014).
17. Wong, N.C. *et al.* Stability of gene expression and epigenetic profiles highlights the utility of patient-derived paediatric acute lymphoblastic leukaemia xenografts for investigating molecular mechanisms of drug resistance. *BMC Genomics* **15**, 416 (2014).
18. Gu, Q. *et al.* Genomic characterization of a large panel of patient-derived hepatocellular carcinoma xenograft tumor models for preclinical development. *Oncotarget* **6**, 20160–20176 (2015).
19. Smith, K.B. *et al.* Novel dedifferentiated liposarcoma xenograft models reveal *PTEN* down-regulation as a malignant signature and response to PI3K pathway inhibition. *Am. J. Pathol.* **182**, 1400–1411 (2013).
20. Ma, C.X. *et al.* Targeting Chk1 in p53-deficient triple-negative breast cancer is therapeutically beneficial in human-in-mouse tumor models. *J. Clin. Invest.* **122**, 1541–1552 (2012).
21. Morisot, S. *et al.* High frequencies of leukemia stem cells in poor-outcome childhood precursor-B acute lymphoblastic leukemias. *Leukemia* **24**, 1859–1866 (2010).

22. Peng, S. *et al.* Tumor grafts derived from patients with head and neck squamous carcinoma authentically maintain the molecular and histologic characteristics of human cancers. *J. Transl. Med.* **11**, 198 (2013).

23. Chou, J. *et al.* Phenotypic and transcriptional fidelity of patient-derived colon cancer xenografts in immune-deficient mice. *PLoS One* **8**, e79874 (2013).

24. Ben-David, U., Mayshar, Y. & Benvenisty, N. Virtual karyotyping of pluripotent stem cells on the basis of their global gene expression profiles. *Nat. Protoc.* **8**, 989–997 (2013).

25. Ben-David, U. *et al.* The landscape of chromosomal aberrations in breast cancer mouse models reveals driver-specific routes to tumorigenesis. *Nat. Commun.* **7**, 12160 (2016).

26. Fehrmann, R.S. *et al.* Gene expression analysis identifies global gene dosage sensitivity in cancer. *Nat. Genet.* **47**, 115–125 (2015).

27. Zack, T.I. *et al.* Pan-cancer patterns of somatic copy number alteration. *Nat. Genet.* **45**, 1134–1140 (2013).

28. Carter, S.L., Eklund, A.C., Kohane, I.S., Harris, L.N. & Szallasi, Z. A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. *Nat. Genet.* **38**, 1043–1048 (2006).

29. Clappier, E. *et al.* Clonal selection in xenografted human T cell acute lymphoblastic leukemia recapitulates gain of malignancy at relapse. *J. Exp. Med.* **208**, 653–661 (2011).

30. Andor, N. *et al.* Pan-cancer analysis of the extent and consequences of intratumor heterogeneity. *Nat. Med.* **22**, 105–113 (2016).

31. Bakker, B. *et al.* Single-cell sequencing reveals karyotype heterogeneity in murine and human malignancies. *Genome Biol.* **17**, 115 (2016).

32. Mroz, E.A., Tward, A.D., Hammon, R.J., Ren, Y. & Rocco, J.W. Intra-tumor genetic heterogeneity and mortality in head and neck cancer: analysis of data from the Cancer Genome Atlas. *PLoS Med.* **12**, e1001786 (2015).

33. Gibson, W.J. *et al.* The genomic landscape and evolution of endometrial carcinoma progression and abdominopelvic metastasis. *Nat. Genet.* **48**, 848–855 (2016).

34. Kim, T.M. *et al.* Subclonal genomic architectures of primary and metastatic colorectal cancer based on intratumoral genetic heterogeneity. *Clin. Cancer Res.* **21**, 4461–4472 (2015).

35. Bai, H. *et al.* Integrated genomic characterization of *IDH1*-mutant glioma malignant progression. *Nat. Genet.* **48**, 59–66 (2016).

36. Hedberg, M.L. *et al.* Genetic landscape of metastatic and recurrent head and neck squamous cell carcinoma. *J. Clin. Invest.* **126**, 1606 (2016).

37. Um, S.W. *et al.* Molecular evolution patterns in metastatic lymph nodes reflect the differential treatment response of advanced primary lung cancer. *Cancer Res.* **76**, 6568–6576 (2016).

38. Sveen, A. *et al.* Intra-patient inter-metastatic genetic heterogeneity in colorectal cancer as a key determinant of survival after curative liver resection. *PLoS Genet.* **12**, e1006225 (2016).

39. Lee, S.Y. *et al.* Comparative genomic analysis of primary and synchronous metastatic colorectal cancers. *PLoS One* **9**, e90459 (2014).

40. Lim, B. *et al.* Genome-wide mutation profiles of colorectal tumors and associated liver metastases at the exome and transcriptome levels. *Oncotarget* **6**, 22179–22190 (2015).

41. Daniel, V.C. *et al.* A primary xenograft model of small-cell lung cancer reveals irreversible changes in gene expression imposed by culture *in vitro*. *Cancer Res.* **69**, 3364–3373 (2009).

42. Günther, H.S. *et al.* Glioblastoma-derived stem cell–enriched cultures form distinct subgroups according to molecular and phenotypic criteria. *Oncogene* **27**, 2897–2909 (2008).

43. Schulte, A. *et al.* A distinct subset of glioma cell lines with stem cell–like properties reflects the transcriptional phenotype of glioblastomas and overexpresses CXCR4 as therapeutic target. *Glia* **59**, 590–602 (2011).

44. Cifola, I. *et al.* Renal cell carcinoma primary cultures maintain genomic and phenotypic profile of parental tumor tissues. *BMC Cancer* **11**, 244 (2011).

45. Hollingshead, M.G. *et al.* Gene expression profiling of 49 human tumor xenografts from *in vitro* culture through multiple *in vivo* passages—strategies for data mining in support of therapeutic studies. *BMC Genomics* **15**, 393 (2014).

46. Gillet, J.P., Varma, S. & Gottesman, M.M. The clinical relevance of cancer cell lines. *J. Natl. Cancer Inst.* **105**, 452–458 (2013).

47. Birkbak, N.J. *et al.* Paradoxical relationship between chromosomal instability and survival outcome in cancer. *Cancer Res.* **71**, 3447–3452 (2011).

48. Zhang, W. *et al.* Centromere and kinetochore gene misexpression predicts cancer patient survival and response to radiotherapy and chemotherapy. *Nat. Commun.* **7**, 12619 (2016).

49. Silk, A.D. *et al.* Chromosome missegregation rate predicts whether aneuploidy will promote or suppress tumors. *Proc. Natl. Acad. Sci. USA* **110**, E4134–E4141 (2013).

50. Zasadil, L.M. *et al.* Cytotoxicity of paclitaxel in breast cancer is due to chromosome missegregation on multipolar spindles. *Sci. Transl. Med.* **6**, 229ra43 (2014).

51. Janssen, A., Kops, G.J. & Medema, R.H. Elevating the frequency of chromosome mis-segregation as a strategy to kill tumor cells. *Proc. Natl. Acad. Sci. USA* **106**, 19108–19113 (2009).

52. Burrell, R.A. *et al.* Targeting chromosomal instability and tumour heterogeneity in HER2-positive breast cancer. *J. Cell. Biochem.* **111**, 782–790 (2010).

53. Tempest, H.G. *et al.* Sperm aneuploidy frequencies analysed before and after chemotherapy in testicular cancer and Hodgkin's lymphoma patients. *Hum. Reprod.* **23**, 251–258 (2008).

54. Khan, F., Sherwani, A.F. & Afzal, M. Analysis of genotoxic damage induced by dacarbazine: an *in vitro* study. *Toxin Rev.* **29**, 130–136 (2010).

55. Ben-David, U. *et al.* Aneuploidy induces profound changes in gene expression, proliferation and tumorigenicity of human pluripotent stem cells. *Nat. Commun.* **5**, 4825 (2014).

56. Cai, Y. *et al.* Loss of chromosome 8p governs tumor progression and drug response by altering lipid metabolism. *Cancer Cell* **29**, 751–766 (2016).

57. Chen, G. *et al.* Targeting the adaptability of heterogeneous aneuploids. *Cell* **160**, 771–784 (2015).

58. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).

59. Rees, M.G. *et al.* Correlating chemical sensitivity and basal gene expression reveals mechanism of action. *Nat. Chem. Biol.* **12**, 109–116 (2016).

60. Tsherniak, A. *et al.* Defining a cancer dependency map. *Cell* **170**, 564–576.e16 (2017).

61. Brastianos, P.K. *et al.* Genomic characterization of brain metastases reveals branched evolution and potential therapeutic targets. *Cancer Discov.* **5**, 1164–1177 (2015).

62. Liu, W. *et al.* Copy number analysis indicates monoclonal origin of lethal metastatic prostate cancer. *Nat. Med.* **15**, 559–565 (2009).

ONLINE METHODS

PDX data assembly and processing. CGH array, SNP array and gene expression microarray data were obtained from the Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) and EMBL-EBI (http://www.ebi.ac.uk/) repositories. RNA sequencing data were obtained from the NCBI Sequence Read Archive (SRA; https://www.ncbi.nlm.nih.gov/sra/). Accession numbers are provided in **Supplementary Table 1**. Normalized matrix files were downloaded, and samples were curated manually to identify the cancer tissue type and the PDX passage number. Arrays were analyzed for quality control, and outliers were removed. The final database consisted of 1,110 PDX tumor samples, from 543 unique PDX models across 24 cancer types. The analysis was performed in batches, and normal tissue samples included in each study served as internal controls, whenever available. Data were processed using R statistical software (http://www.r-project.org/). For all platform types, probe sets were organized by their chromosomal location and $\log_2$-transformed values were used. Probe sets without an annotated chromosomal location were removed. For gene expression data, all probe sets for each gene were averaged (as well as their chromosomal location), to obtain one intensity value per gene. A threshold expression value was set, and genes with lower expression values were collectively raised to that level: flooring values were 6–7 for the Affymetrix and Illumina platforms and −0.5 for the Agilent platforms. Probe sets not expressed in >20% of the samples within a batch were removed. The 10% of the probe sets with the most variable expression levels were also excluded, to reduce expression noise.

Generation of CNA profiles. CNA profiles from SNP arrays were generated using the Copy Number Workflow of the Partek Genomics Suite software (http://www.partek.com/pgs), as reported by the original studies. CNA profiles from CGH arrays were generated using CGH-Explorer software (http://heim.ifi.uio.no/bioinf/Projects/CGHExplorer/), using the program's piecewise constant fit (PCF) algorithm with the following set of parameters: least allowed deviation = 0.3; least allowed aberration size = 30; winsorize at quantile = 0.001; penalty = 12; threshold = 0.01. CNA profiles from gene expression data were generated using the protocols developed by Ben-David *et al.*[24] and by Fehrman *et al.*[26]. For all gene expression platforms, the e-karyotyping method was applied[24]: whenever normal tissue samples were available, the median expression value for each gene across the normal samples was subtracted from the expression value for that gene in the tumor samples, to obtain comparative values. These relative gene expression data were then subjected to a CGH-PCF analysis with the following set of parameters: least allowed deviation = 0.25; least allowed aberration size = 30; winsorize at quantile = 0.001; penalty = 12; threshold = 0.01. For Affymetrix gene expression platforms Human Genome U133A and U133Plus2.0, the functional genomic mRNA profiling (FGMP) method[26] was also applied: gene expression data were corrected for the first 25 previously identified transcriptional components, and the corrected data were subjected to the same processing steps and CGH-PCF analysis described above with the following set of parameters: least allowed deviation = 0.15; least allowed aberration size = 30; winsorize at quantile = 0.001; penalty = 12; threshold = 0.01. CNA profiles from DNA sequencing were obtained in a processed table form from the publication by Eirew *et al.*[7]. CNA profiles from SNP arrays were obtained in a processed form from the publication by Gao *et al.*[3] and compared to CNA profiles from RNA sequencing of the same PDX models. For visualization purposes, moving-average plots were generated using the CGH-Explorer moving average fit tool.

Identification of model-acquired CNAs. To identify CNAs emerging during the generation and propagation of PDXs, 342 PDX models for which data were available from multiple time points were analyzed. These PDX models were compared either to the primary tumors from which they were derived or to their earliest available passage. For gene expression data, model-acquired CNAs were identified by e-karyotyping. For each probe set, a relative value was obtained by subtracting the early time point value from the late time point value. CGH-PCF analysis was then performed with the same parameters described above. For visualization purposes, moving-average plots were generated using the CGH-Explorer moving average fit tool.

CNA recurrence analysis. For each tissue type, arm-level CNA recurrence was computed and compared between the PDX data set and the human patient TCGA data set (http://cancergenome.nih.gov/). Chromosome arm-level events in TCGA samples were called using a new approach to be described in a separate publication (Taylor, J.S. and G.H. *et al.,* unpublished data). Briefly, segments of CNAs identified by ABSOLUTE[63] were determined as loss, neutral or gain relative to each sample's predicted tumor ploidy. Consecutive segments were iteratively joined such that the combined segment was no less than 80% altered in a given direction (gain or loss, but not both). For each combination of arm/chromosome and direction of alteration within each TCGA tumor type, the start coordinates, end coordinates and proportion of the chromosome arm altered (based on the longest joined segment) were clustered across samples using a 3D Gaussian mixture model. The optimal clustering solution was chosen on the basis of the Bayesian information criterion. Clusters whose mean fraction altered in either specific direction was ≥80% were considered 'aneuploid', and those whose mean fraction altered (in both directions) was ≤20% were considered 'non-aneuploid'. Chromosome arm-level events in PDX samples were determined using the CNA status of the largest overlapping segment from the e-karyotyping analysis. Comparisons of absolute CNA landscapes were performed using the FGMP-derived CNA profiles, and comparisons of model-acquired CNAs were performed using the e-karyotyping-derived CNA profiles. Comparisons between early- and late-passage PDXs were performed using FGMP-derived CNA profiles: samples from P1 or earlier were defined as early passage, and those from P3 or later were defined as late passage. Heat maps were generated using the pheatmap R package, and clustering was performed using Euclidean distance and complete linkage.

DGI comparison across passages and tissue types. The DGI of each sample was determined in three ways: (i) the fraction of the genome affected by model-acquired CNAs per passage, (ii) the number of discrete events per passage, and (iii) the fraction of altered genes per passage. For each cancer type, the median number of model-acquired arm-level CNAs across all PDX samples was determined and compared to several TCGA statistics: the percentage of samples with WGD taken from the publication by Zack *et al.*[27], the median number of arm-level CNAs per sample and the median number of clones per sample taken from the publication by Andor *et al.*[30].

Similarity analysis. PDX samples derived from the same primary tumors but propagated in different mice starting from their initial transplantation (transplanted into different P0 mice) were defined as 'siblings'. PDX samples derived from the same primary tumors and propagated in the same mice at some point during PDX propagation were excluded from the analysis. PDX samples from distinct primary tumors were defined as 'non-siblings'. Similarity scores were calculated for each pair of samples on the basis of the arm-level events that occurred during their *in vivo* passaging (model-acquired CNAs), using a modified Jaccard similarity coefficient. This similarity coefficient was inversely weighted to account for the observed prevalence of each CNA in each PDX tissue type. Therefore, the similarity score was calculated using the following equation

$$\text{Similarity} = \frac{A}{A + B + C}$$

Where

$$A = \sum_{k=0}^{n = \text{Number of shared events}} 1/\text{freq}(k)$$

$$B = \sum_{k=0}^{n = \text{Number of non-shared events}} 1$$

$$C = 2 * \sum_{k=0}^{n = \text{Number of opposite events}} 1$$

and freq(k) is the frequency of event k in that tumor type.

Loss-of-heterozygosity analysis. Allelic copy number data were obtained from the publication by Eirew *et al.*[7]. Using 10-Mb windows along the genome, we identified the following scenarios: (i) the minor allele was 0 (LOH) in a primary tumor but >0 (presence of both alleles) in the tumor-derived PDX model and (ii) the minor allele was 0 (LOH) in an early-passage PDX but >0 in a later passage of the same PDX model. These instances of apparent reversion of LOH were visualized using the Integrative Genomics Viewer (IGV; https://www.broadinstitute.org/igv/) and replotted in **Figure 2f**.

Gene expression signature scores. The apoptosis and proliferation gene sets were derived from the Molecular Signature Database (http://software.broad-institute.org/gsea/msigdb) using the Hallmark_Apoptosis[64] and Benporath_Proliferation[65] gene sets, respectively. The CIN70 gene set was derived from the publication by Carter *et al.*[13]. Signature scores were generated for all PDX models reported by Gao *et al.*[3]. For each gene set, genes not expressed at all in the PDX data set were removed and the remaining gene expression values were $\log_2$ transformed and scaled by subtracting the gene expression means. The signature score was defined as the sum of these scale-normalized gene expression values.

Advanced disease data assembly, processing and CNA profiling. CGH array and SNP array data were obtained from the GEO and EMBL-EBI repositories. CNA profiles from DNA sequencing were obtained in a processed table form from the authors of the respective publications. Accession numbers are provided in **Supplementary Table 4**. The final database consisted of 306 tumor samples across five cancer types. Chromosome arm-level events in advanced-disease samples were determined as described above. Progression-acquired CNAs were determined as events that were identified in the primary tumor but not in its advanced-disease sample (metastasis, recurrence or progressed sample), or vice versa.

CLDX data assembly, processing and CNA profiling. Gene expression microarray data from the NCI MicroXeno project were downloaded from the GEO repository under accession GSE48433 (ref. 45). Data were processed as described above. CNA profiles were generated using the FGMP method, and model-acquired CNAs were identified by e-karyotyping, as described above. The gene expression values for *in vitro*–cultured (P0) cell lines were used as the reference in the e-karyotyping analysis. The numerical karyotypic complexity categorization of the cell lines was obtained from the publication by Roschke *et al.*[66].

Cell line data assembly, processing and CNA profiling. Whole-exome sequencing data from nine newly derived cell lines were obtained from Tseng *et al.* (Y.-Y.T., unpublished data). CNA profiles were generated from these data using the ReCapSeg program (http://gatkforums.broadinstitute.org/gatk/categories/recapseg-documentation), from the ratio of tumor read depth to the expected read depth (as determined from a panel of normal samples). Gene expression microarray data were obtained from the GEO repository. Accession numbers are provided in **Supplementary Table 3**. Data were processed as described above. CNA profiles were generated using the FGMP method, and model-acquired CNAs were identified by e-karyotyping, as described above. Renal cancer CNA data were obtained directly from Cifola *et al.*[44], and model-acquired CNAs were identified as described above. For comparison of model-acquired CNA rates across passages, samples were compared to the earliest available passage (P0 or P1). Samples from P7 or earlier were defined as early passage, samples from P10 were defined as medium passage and samples from P19 or later were defined as late passage.

PDX drug response association analyses. PDX drug response data were obtained from the publication by Gao *et al.*[3]. For analysis of the association between chemotherapy response and absolute levels of aneuploidy, the CNA fraction was determined according to the FGMP-derived CNA profiles of the sample from the latest passage available for each model. Low CNA levels were determined as CNA fraction < 0.3; intermediate CNA levels were determined as 0.3 < CNA fraction < 0.7; and high CNA levels were determined as CNA fraction > 0.7. For analysis of the association between chemotherapy response and DGI, the DGI level of each model was determined as the number of discrete model-acquired CNAs per passage, using the latest passage

available for each model: low DGI was determined as DGI = 0; intermediate DGI was determined as 0 < DGI < 4; and high DGI was determined as DGI > 4. BestAvgResponse values were used to make response calls. Association tests were conducted in each available tissue type independently, yielding a total of six drug–tissue association tests (representing five chemotherapies in five tissue types). For analysis of the association between targeted therapy response and the existence of specific arm-level events, the arm-level copy number status of each model was set according to the FGMP-derived CNA profiles of the sample from the latest passage available for that model. BestAvgResponse values were used to make response calls. For each of the 12 recurrent TCGA events that tended to disappear throughout PDX passaging, the association of the event with PDX drug response was evaluated in the relevant cancer type. All targeted drugs that were used as single agents and yielded at least partial response in at least one mouse were evaluated. Drug combinations were also evaluated if one or both of the drugs in the combination was not tested as a single agent. A total of 54 association tests were performed (representing 15 single-agent drugs and 5 drug combinations in three tissue types).

Cell line gene expression, genetic dependency and drug response association analyses. Chromosome arm-level events in cell line samples were called using the same approach used to call arm-level events in TCGA samples (described above). Gene expression was measured for CCLE cell lines using RNA-seq ($n = 936$ cell lines with arm-level CNA calls) and was normalized to RPKM values for each gene[60]. Gene essentiality profiles for each cell line were derived from genome-wide RNAi screens across 503 cell lines ($n = 446$ with arm-level CNA calls), using the DEMETER algorithm to isolate the effects of gene knockdown from off-target effects[60]. Drug sensitivity measurements were taken from the Cancer Therapeutics Response Portal (CTRP v2) data set (downloaded from https://ocg.cancer.gov/programs/ctd2/data-portal/). These data represented dose–response curve AUC measures for 887 cell lines ($n = 804$ with arm-level CNA calls) across 545 compounds. Comparisons of gene expression, RNAi-based gene essentiality and compound sensitivity for cell lines with and without particular arm-level CNAs were made using standard linear modeling tools developed for differential expression analysis: the R package limma was used to derive P values from empirical Bayes moderated t statistics[67]. In all cases, we tested the effect of arm-level CNAs while controlling for between-lineage differences by including lineage as a covariate in the model[67]. Gene set testing was performed using a parametric approximation to permutation-based testing, implemented in the R package npGSEA[68]. Between-lineage differences were controlled for by regressing lineage out of both the arm-level CNA calls and the variable of interest[68].

Mutation allelic fraction analysis. The allelic fraction (AF) and predicted effects of point mutations were obtained from the publication by Eirew *et al.*[7]. Shifts in allelic fraction were determined as $|\text{AF}_{(primary)} - \text{AF}_{(PDX)}| > 0.2$. Missense and nonsense coding mutations were considered separately from the rest of the mutations, and their shifts in allelic fraction are plotted in **Supplementary Figure 4**.

Comparison of CNA-based and mutation-based phylogenetic trees. Copy number and mutation data were obtained from Gibson *et al.*[33] and Bi *et al.*[69]. To construct single-nucleotide variant (SNV)-based trees, SNVs present at low allelic fractions and SNVs from regions of low sequencing coverage were first excluded. To construct arm-level CNA–based trees, copy number data from the tumors were converted into arm-level calls. Phylogenetic trees were then generated separately. In each case, branch points in the tree were assigned on the basis of the overall similarity in shared events and branch lengths were set to be proportional to the number of shared events. Patients for whom trees could not be generated owing to insufficient information (for example, no branch points identified) were excluded from the analysis. CNA-based trees were then compared to SNV-based trees to determine their sensitivity (the proportion of SNV-based branch points identified in CNA-based trees) and specificity (the proportion of CNA-based branch points identified in SNV-based trees).

Statistical analyses. The significance of differences in prevalence and rates of absolute CNAs and model-acquired CNAs between PDX passages, between

primary and metastatic PDXs, between PDXs with wild-type and mutated or deleted p53, between PDXs from the most stable (upper quartile) and least stable (lower quartile) tissue types, between cell line passages, between CLDX passages and between CLDXs from cell lines of distinct numerical karyotypic complexity was determined using the two-tailed Wilcoxon rank-sum test. The significance of differences in similarity scores between siblings and non-siblings, the significance of differences in CNA rates between PDXs and cell lines, and the significance of differences between PDX-acquired and progression-acquired CNA fractions were determined using a stratified bootstrap test, permuting the data 100,000 times within each tissue type. The significance of the gene expression signature trends observed throughout PDX passaging was determined by the Kruskal–Wallis rank-sum test. The significance of correlations between PDX and TCGA data was determined using a Spearman's correlation test. To evaluate the tendency to acquire or lose recurrent TCGA CNAs during PDX propagation and during disease progression in patients, recurrent CNAs were defined for each tissue type as those that recurred in over 40% of the samples and the number of events that involved these CNAs were computed in the lineage-matched PDX cohorts; the significance of differences between the emergence frequency and the loss frequency in PDXs and in advanced disease was determined using McNemar's test, and the significance of differences between the two groups was determined using a chi-squared test for equality of proportions (using the proportion observed in humans as the expected proportion). The significance of the differences in CNA prevalence between early- and late-passage PDX samples was evaluated using the one-tailed Fisher's exact test. The significance of the association between chromosome arms and drug response was determined using the Wilcoxon rank-sum test, with false discovery rate (FDR) multiple-test correction performed for each tissue type independently. Box plots show the median, 25th and 75th percentiles, lower whiskers show data within the 25th percentile − 1.5 times

the IQR, upper whiskers show data within the 75th percentile + 1.5 times the IQR, and circles show the actual data points. Violin plots show the combination of a box plot and a kernel density plot, in which the width is proportional to the relative frequency of the measurements. All of the statistical tests were performed using R statistical software, and the box plots and violin plots were generated using the boxplot and vioplot R packages, respectively. A **Life Sciences Reporting Summary** is available.

Code availability. The codes used to generate and/or analyze the data during the current study are publically available or are available from the authors upon request.

Data availability. The data sets generated during and/or analyzed during the current study are available within the article and its supplementary information files or are available from the authors upon request.

63. Carter, S.L. *et al.* Absolute quantification of somatic DNA alterations in human cancer. *Nat. Biotechnol.* **30**, 413–421 (2012).
64. Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
65. Ben-Porath, I. *et al.* An embryonic stem cell–like gene expression signature in poorly differentiated aggressive human tumors. *Nat. Genet.* **40**, 499–507 (2008).
66. Roschke, A.V. *et al.* Karyotypic complexity of the NCI-60 drug-screening panel. *Cancer Res.* **63**, 8634–8647 (2003).
67. Ritchie, M.E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
68. Larson, J.L. & Owen, A.B. Moment based gene set tests. *BMC Bioinformatics* **16**, 132 (2015).
69. Bi, W.L. *et al.* Genomic landscape of high-grade meningiomas. *NPJ Genom. Med.* **2**, 15 (2017).