nature medicine

Pan-cancer analysis of the extent and consequences of intratumor heterogeneity

Noemi Andor[1,2], Trevor A Graham[3], Marnix Jansen[3], Li C Xia[1], C Athena Aktipis[4,5], Claudia Petritsch[6–8], Hanlee P Ji[1,9,12] & Carlo C Maley[4,10–12]

Intratumor heterogeneity (ITH) drives neoplastic progression and therapeutic resistance. We used the bioinformatics tools 'expanding ploidy and allele frequency on nested subpopulations' (EXPANDS) and PyClone to detect clones that are present at a ≥10% frequency in 1,165 exome sequences from tumors in The Cancer Genome Atlas. 86% of tumors across 12 cancer types had at least two clones. ITH in the morphology of nuclei was associated with genetic ITH (Spearman's correlation coefficient, $\rho = 0.24–0.41$; $P < 0.001$). Mutation of a driver gene that typically appears in smaller clones was a survival risk factor (hazard ratio (HR) = 2.15, 95% confidence interval (CI): 1.71–2.69). The risk of mortality also increased when >2 clones coexisted in the same tumor sample (HR = 1.49, 95% CI: 1.20–1.87). In two independent data sets, copy-number alterations affecting either <25% or >75% of a tumor's genome predicted reduced risk (HR = 0.15, 95% CI: 0.08–0.29). Mortality risk also declined when >4 clones coexisted in the sample, suggesting a trade-off between the costs and benefits of genomic instability. ITH and genomic instability thus have the potential to be useful measures that can universally be applied to all cancers.

Cancers are a mosaic of clones with varying population sizes, different genetic makeups and distinct phenotypic characteristics[1–4]. This intratumor heterogeneity provides the fuel for the engine of natural selection that drives the processes of carcinogenesis and acquired therapeutic resistance in neoplasms[1,5]. When analyzing genome-sequencing data derived from single-tumor samples, it is important to recognize that, technically, sequences obtained from each tumor sample encode a tumor metagenome, as they represent the aggregate genomes of all of the clones that coexist within the sample[6–10]. Recently, McGranahan *et al.*[6] used exome-sequencing data derived from single-tumor samples to determine the clonal status of known, actionable drivers across nine cancer types and to identify events that trigger clonal expansions, causing ITH. However, the availability of just one sample per tumor and moderate sequencing depth have limited the opportunities for the systematic analysis of both the extent and the clinical consequences of ITH in previous pan-cancer studies[3,7,10–14]. To overcome these limitations, a variety of different algorithms have been developed to deconvolute tumor metagenomes. These algorithms estimate the cellular prevalence of mutations and quantify ITH[15–19]. We leveraged two of these tumor mixture–separation algorithms, EXPANDS[18] and PyClone[17], to quantify ITH from exome-sequencing data in The Cancer Genome Atlas (TCGA) and to validate the robustness of our results.

[1]Division of Oncology, Department of Medicine, Stanford University School of Medicine, Stanford, California, USA. [2]Institute of Bioinformatics and Systems Biology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany. [3]Evolution and Cancer Laboratory, Barts Cancer Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK. [4]Center for Evolution and Cancer, University of California San Francisco, San Francisco, California, USA. [5]Department of Psychology, Arizona State University, Tempe, Arizona, USA. [6]Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, San Francisco, California, USA. [7]Brain Tumor Research Center, Department of Neurological Surgery, University of California San Francisco, San Francisco, California, USA. [8]Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research, University of California San Francisco, San Francisco, California, USA. [9]Stanford Genome Technology Center, Stanford University, Palo Alto, California, USA. [10]Centre for Evolution and Cancer, Institute for Cancer Research, London, UK. [11]Biodesign Institute, Arizona State University, Tempe, Arizona, USA. [12]These authors contributed equally to this work. Correspondence should be addressed to C.C.M. (maley@asu.edu) or H.P.J. (hanleeji@stanford.edu).

Received 24 April; accepted 7 October; published online 30 November 2015; doi:10.1038/nm.3984

RESULTS

Intratumor genetic heterogeneity exists in all tumor types

We measured the number and size of genetically diverse clones of 1,165 primary tumor samples across 12 cancer types from TCGA, using exome sequencing data from paired tumor-normal samples. These samples originated from a single sequencing center (Broad Institute, Cambridge, Massachusetts, USA) and were chosen because they fulfilled established, strict criteria to obtain uniform sequence data quality and depth (**Supplementary Fig. 1**). As clone detection sensitivity is highly dependent on the depth and breadth of genomic sequencing coverage, these criteria are necessary to ensure that measures of ITH that are derived from these sequences are comparable. Detailed inclusion criteria are available in **Supplementary Note 1**.

Somatic single-nucleotide variants (SNVs) and copy-number variants (CNVs) were called using MuTect[20] and ExomeCNV[21], respectively (**Supplementary Fig. 2**). We distinguished nonsynonymous SNVs and splice site– or regulatory-region SNVs (generally referred to as nonsilent) from synonymous SNVs and SNVs within intergenic and intronic regions (also referred to as silent). The incidence of CNVs and somatic nonsilent SNVs varied considerably within and between tumor types (**Fig. 1a–c**), similarly to results obtained from other genome-wide sequencing studies[11,22,23].

EXPANDS was applied to all detected somatic SNVs (including silent SNVs), and loss-of-heterozygosity (LOH) and copy-number estimates to infer the number, size and genetic content of subpopulations

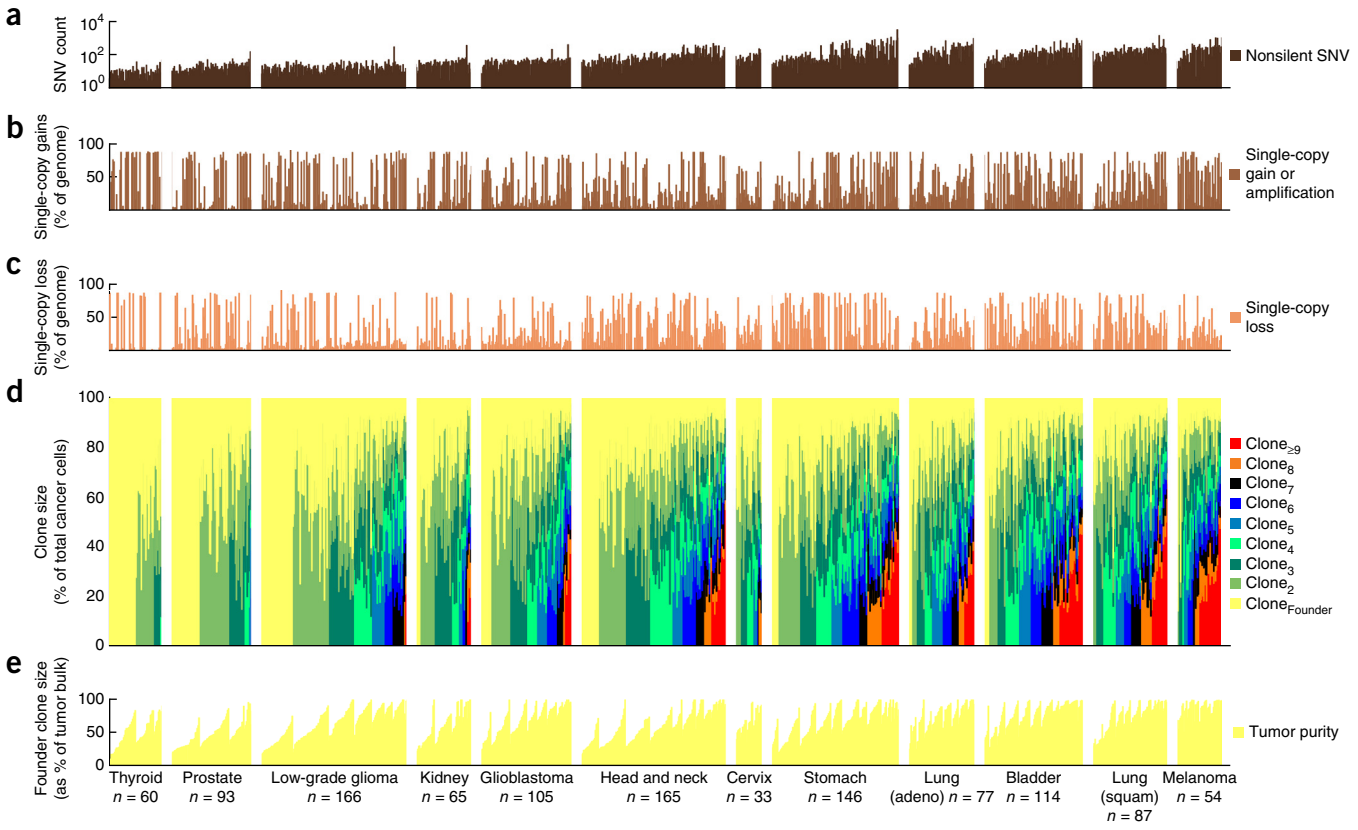

Figure 1 Tumor metagenomes and subclonal genomes in 12 tumor types from TCGA. (**a**) Prevalence of nonsilent somatic SNVs in different tumors. (**b**,**c**) Percentage of the tumor metagenome affected by single-copy gains or amplifications (**b**) and single-copy losses (**c**). (**d**) Clonal composition inferred from SNVs and copy numbers. Every sample contains a founder tumor population (yellow; identified as the largest clone within the sample). Each change in color marks the presence of an additional clone at the indicated size, calculated as 'percentage of the founder-population size' (*y* axis). Color variety within each tumor-type panel reflects the extent of ITH in the corresponding tumor type. The average number of detectable (with ≥10% frequency) clones increases from thyroid carcinoma (left) to melanoma (right). (**e**) Founder clone size as a measure of tumor purity. The exact number of tumors of each type (*n*) is indicated at the bottom of each panel.

of cells that coexist in the tumor (**Fig. 1d**). Briefly, EXPANDS models the cellular prevalence of each SNV as a copy-number–dependent probability distribution. Subsequently, these cellular prevalence distributions are clustered to obtain the genetic content of each subpopulation, i.e., the set of SNVs and CNVs that accumulated in ancestral cells before each clonal expansion. Previous results[18] indicate that the sequencing data available per tumor (on average 5,221 Mb of reads) translates to an accuracy of 50–80% at which EXPANDS detects genetic heterogeneity at a macroscopic resolution[24] (i.e., clones present in ≥10% of the sample). An independent algorithm, PyClone, was used to validate the conclusions derived from EXPANDS. PyClone infers the cellular prevalence of SNVs differently from EXPANDS. In particular, PyClone does not model subclonal CNVs, and it leverages high depth rather than high breadth of sequencing[17] (**Supplementary Note 2**).

In general, the cellular prevalence of SNVs assigned by PyClone and EXPANDS was concordant for SNVs located within genomic regions of clonal copy number (for which all tumor cells have identical ploidy) ($\rho = 0.77$). However, for regions in which CNVs affect only a subset of tumor cells, EXPANDS and PyClone tended to make different inferences for the cellular prevalence of SNVs within those regions ($\rho = 0.25$; **Supplementary Note 2**).

Subpopulations detected within the same tumor sample may have sizes that cumulatively exceed 100%, as a subpopulation may be nested in a parental population that carries earlier mutations. Both algorithms detect such nested subpopulation compositions. We refer to these inferred subpopulations as clones and to the cellular prevalence of a subpopulation within the tumor sample as its clone size. We denote the total number of clones detected within a tumor sample as the clone number. As noted previously, we define the term 'tumor metagenome' as the aggregate genomes of all coexistent clones within a tumor.

Assuming a monoclonal tumor origin, the largest inferred clone in each sample harbors the founder mutations (also referred to as 'trunk mutations') and corresponds to the first (founder) clonal expansion. This holds true regardless of the fitness difference between the founder and descending clones. The cellular prevalence of founder mutations will always be greater than or equal to the cellular prevalence of mutations acquired by descendant subclones, even if these later subclones proliferate faster than the founder cells. This implies that the size of the largest clone (i.e., the founder clone) is also a measure of tumor purity (**Fig. 1e**); this was confirmed by an independent study that compared the performance of EXPANDS to four other methods that predict tumor purity[25]. The size of the largest clone was correlated to tumor purity as predicted from expression profiling using ESTIMATE[26] (EXPANDS: Pearson coefficient of correlation, $r = 0.43$; $P < 1 \times 10^{-6}$; PyClone: $r = 0.63$; $P < 1 \times 10^{-6}$; **Supplementary Fig. 3a**).

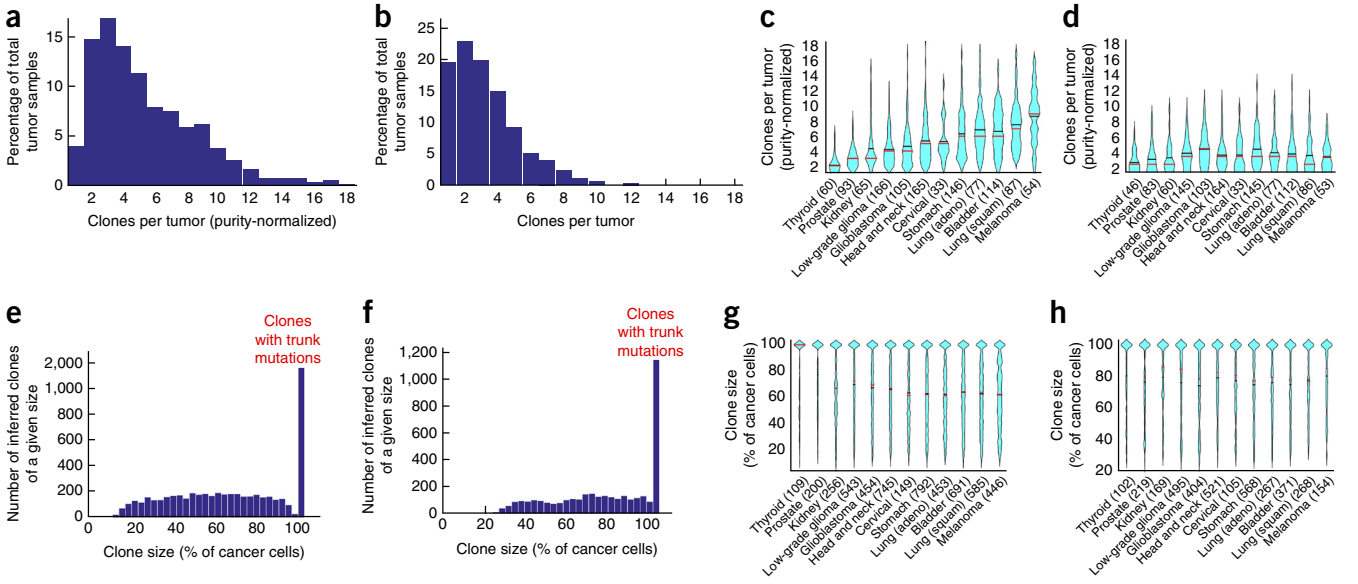

Figure 2 Intratumor genetic heterogeneity in 12 tumor types. (**a**,**b**) Clone number distribution as predicted by EXPANDS (of 1,165 tumor samples) (**a**) and PyClone (of 1,107 tumor samples) (**b**) across tumor types. (**c**,**d**) Violin plots of clone number distribution as predicted by EXPANDS (**c**) and PyClone (**d**) within different tumor types. The exact number of tumors of each type is indicated in parentheses. (**e**,**f**) Clone size distribution as predicted by EXPANDS (**e**) and PyClone (**f**) across tumor types. (**g**,**h**) Violin plots of clone size distribution as predicted by EXPANDS (**g**) and PyClone (**h**) within different tumor types. The number of clones cumulatively detected in all samples of a tumor type is indicated in parentheses for each type. EXPANDS-derived clone numbers (**a**,**c**) and all clone sizes (**e**–**h**) have been normalized by tumor purity. For PyClone-derived clone numbers (**b**,**d**), normalization by tumor purity was not necessary. Violin plots contain marks for the means (black lines) and medians (red lines). 'Trunk mutations' denotes the set of SNVs accumulated by a founder cell prior to its clonal expansion. Thyroid, thyroid carcinoma; prostate, prostate adenocarcinoma; kidney, kidney renal clear cell carcinoma; head and neck, head and neck squamous cell carcinoma; cervical, cervical squamous cell carcinoma and endocervical adenocarcinoma; stomach, stomach adenocarcinoma; lung (adeno), lung adenocarcinoma; bladder, bladder urothelial carcinoma; lung (squam), lung squamous cell carcinoma; melanoma, skin cutaneous melanoma.

We observed that the number of somatic SNVs in large clones correlated with age at diagnosis ($\rho = 0.3$; $P < 1 \times 10^{-6}$), a result previously reported for chronic lymphocytic leukemia (CLL)[8]. In addition, the number of SNVs in small clones also correlated with age ($\rho = 0.18$; $P = 5 \times 10^{-6}$; **Supplementary Fig. 3b**,**c**).

We compared the extent of genetic ITH across and within tumor types (**Fig. 2a**–**d**). The difference in the number of clones harbored between tumor types was similar before and after correcting for tumor purity (**Fig. 2c**; see Online Methods). On average, four clones were estimated to coexist in a tumor at the time of biopsy or surgical resection (the median estimated number of clones was five by using EXPANDS and three by using PyClone; **Fig. 2a**,**b**). There was a median of 10 (EXPANDS estimate) to 16 (PyClone estimate) nonsilent somatic SNVs per clone, and the distribution of clone sizes across tumor types was relatively uniform. Notably, reduced detection sensitivity (because of low tumor purity) was not sufficient to explain the smaller number of clones that were observed in low-purity tumors (**Fig. 2e**–**h** and **Supplementary Note 3**).

In 14% (EXPANDS estimate) to 20% (PyClone estimate) of the analyzed tumor samples, only a single, genetically homogeneous cell population was detected. Even for thyroid carcinoma, the least heterogeneous tumor type, two or more clones were predicted to coexist in >50% of the samples (EXPANDS estimate: 52%; PyClone estimate: 65%). Therefore, we concluded that genetic ITH occurs in the vast majority of cancers represented among the 12 types that we included in this study.

Driver genes are mutated in clones of characteristic sizes

To investigate the influence of the incidence of driver gene mutations on genetic ITH, we analyzed 259 cancer-driver genes (hereafter referred to as 'CAN genes'; **Supplementary Table 1**). A gene was deemed to be substantially associated with a given cancer type on the basis of: (i) prior experimental evidence, (ii) the frequency of gene mutations in our sample cohort and (iii) mutation deleteriousness (see Online Methods). We found 124 non-private CAN genes (48%) that are mutated in a minimum of two cancer types (**Fig. 3a**). We refer to a clone with one or more nonsilent mutations in any given CAN gene as a 'CAN clone'.

Next, we tested whether CAN clones differed in size depending on which CAN genes are mutated in the corresponding clones. The size of a clone depends on its selective fitness (how fast it expands relative to the other clones within a tumor) and on its formation time (when the underlying clonal expansion started). CAN-gene SNVs specific to a given clone may therefore have a direct impact on its size. To test this possibility, we first normalized clone sizes by tumor purity. We then calculated the mean and variance in clone size among all CAN clones and compared them to the variance calculated from random samples of clone sizes from our data (**Supplementary Fig. 4**).

The size of clones harboring CAN-gene mutations varied across the CAN genes (**Fig. 3a**) and was correlated to both the relative order of driver-gene mutations reported in earlier studies[11,27–30] and the clone sizes predicted by PyClone ($0.43 < r < 0.94$; $3.4 \times 10^{-18} < P < 0.12$; **Supplementary Fig. 5**). Across tumor samples, and even across tumor types, CAN genes were often mutated in clones of a characteristic size, i.e., the variance in clone size was significantly lower than that expected by chance (one-sided *t*-test: $P < 0.05$; **Fig. 3a**). For example, *TP53* SNVs were found in larger clones (EXPANDS: 0.811 average cancer-cell fraction; PyClone: 0.746 average cancer-cell fraction) in all nine cancers that are significantly associated with genetic aberrations

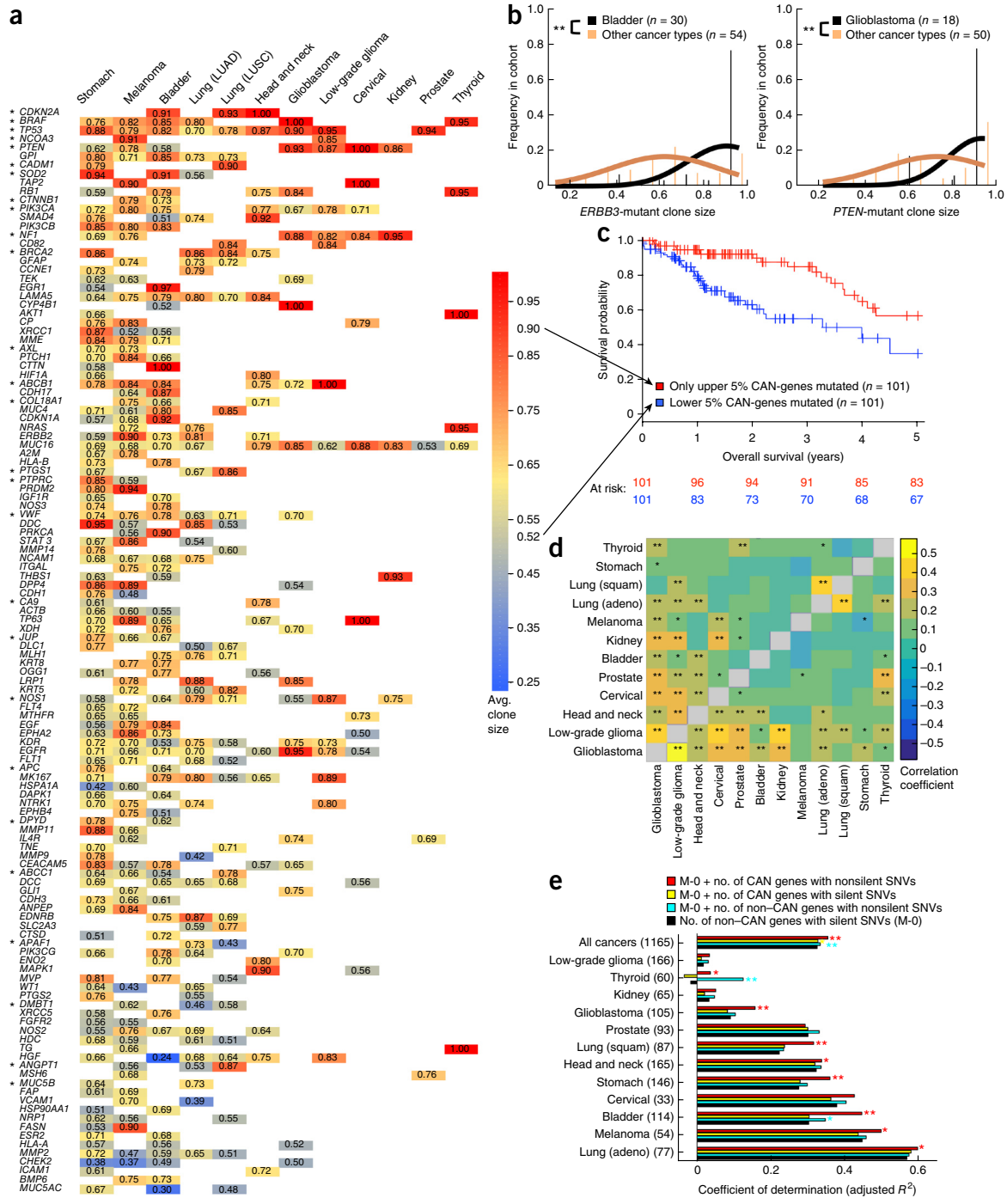

Figure 3 Association of driver-gene mutations with clone size and clone number. (**a**) Clone size distribution for mutations in 124 CAN genes, as predicted by EXPANDS. Clone size is calculated as the fraction of cancer cells. For each cancer type CAN (*x* axis) and gene *G* (*y* axis), the average clone size was calculated across all CAN clones that harbored nonsilent SNVs in *G*. Blank entries denote that *G* was not significantly associated with a cancer type. SNVs in CAN genes often had the tendency to occur in clones of characteristic sizes, independently of cancer type (* $P < 0.05$; one-sided *t*-test). (**b**) Distribution of *ERBB3* mutations in bladder carcinoma (left) and *PTEN* mutations in glioblastoma (right) (** $P < 1 \times 10^{-4}$, one-sided *t*-test). (**c**) Prognosis for survival of individuals with SNVs in CAN genes that are mutated in clones of the top 5% (red line; $n = 101$) or bottom 5% (blue line; $n = 101$) of all CAN clone sizes; CAN genes mutated in clones that characteristically remain smaller predicted poor prognosis across tumor types (log-rank test, $P = 2.9 \times 10^{-4}$; HR = 2.72). (**d**) Comparison of clones with CAN-gene mutations showing similar size distributions across certain tumor types, suggesting that the order and/or selective advantage of CAN-gene mutations is often not tissue specific. Pairwise similarity between tumor types was calculated as a Spearman correlation (** $P < 0.01$, * $P < 0.05$) on the basis of results from EXPANDS (above diagonal) and PyClone (below diagonal) analyses. (**e**) Partial dependency of clone number on SNV incidence. The number of clones identified in a sample depends on SNV incidence, but not all SNV categories are equally associated with the resulting number of clones. Nonsilent SNV incidence in CAN genes (red; mean = 2 genes) explains the variability in clone number (measured as coefficient of determination (R^2) ; *x* axis) better than the incidence of silent SNVs in CAN genes (yellow; mean = 1 gene) or of nonsilent SNVs in non–CAN genes (cyan; mean = 128 genes). ** $P < 0.01$, * $P < 0.05$; log-likelihood test.

in *TP53*. In contrast, somatic SNVs in *DMBT1* were found in smaller clones (EXPANDS: 0.641 average cancer-cell fraction; PyClone: 0.652 average cancer-cell fraction) of the three cancers in which *DMBT1* was among the drivers (**Fig. 3a**).

For a subset of CAN genes, however, we found substantial differences in the dominance of mutated clones, depending on cancer type. For example, clones with *ERBB3* mutations were larger in bladder cancer than they were in any other cancer type, whereas clones with *PTEN* mutations grew particularly large in glioblastoma (**Fig. 3b**). Clones with SNVs in CAN genes that are druggable ($n = 98$) did not have significantly different sizes as compared to clones with SNVs in the remaining CAN genes ($n = 161$; $P = 0.77$, Student's *t*-test).

Furthermore, clone size inferiority of a mutated CAN gene (as shown in **Fig. 3a**) was correlated to the propensity of the CAN gene as a risk factor (EXPANDS: $P = 3.2 \times 10^{-7}$, HR = 2.86; PyClone: $P = 0.005$, HR = 1.67; univariate Cox model). For instance, mutations in CAN genes that occurred in clones with average clone sizes in the bottom 5% of all CAN clone sizes were associated with poor outcome (**Fig. 3c**). This relation was also significant in low-grade gliomas (EXPANDS: $P = 1.60 \times 10^{-3}$, HR = 4.04; PyClone: $P = 0.24$, HR = 1.77), kidney carcinoma (EXPANDS: $P = 0.02$, HR = 5.92; PyClone: $P = 0.04$, HR = 3.48) and glioblastoma (EXPANDS: $P = 0.04$, HR = 2.27; PyClone: $P = 1.62 \times 10^{-3}$, HR = 3.84; **Supplementary Table 2**). Several cellular functions and pathways, including tyrosine protein kinase activity ($P = 1.39 \times 10^{-12}$) and

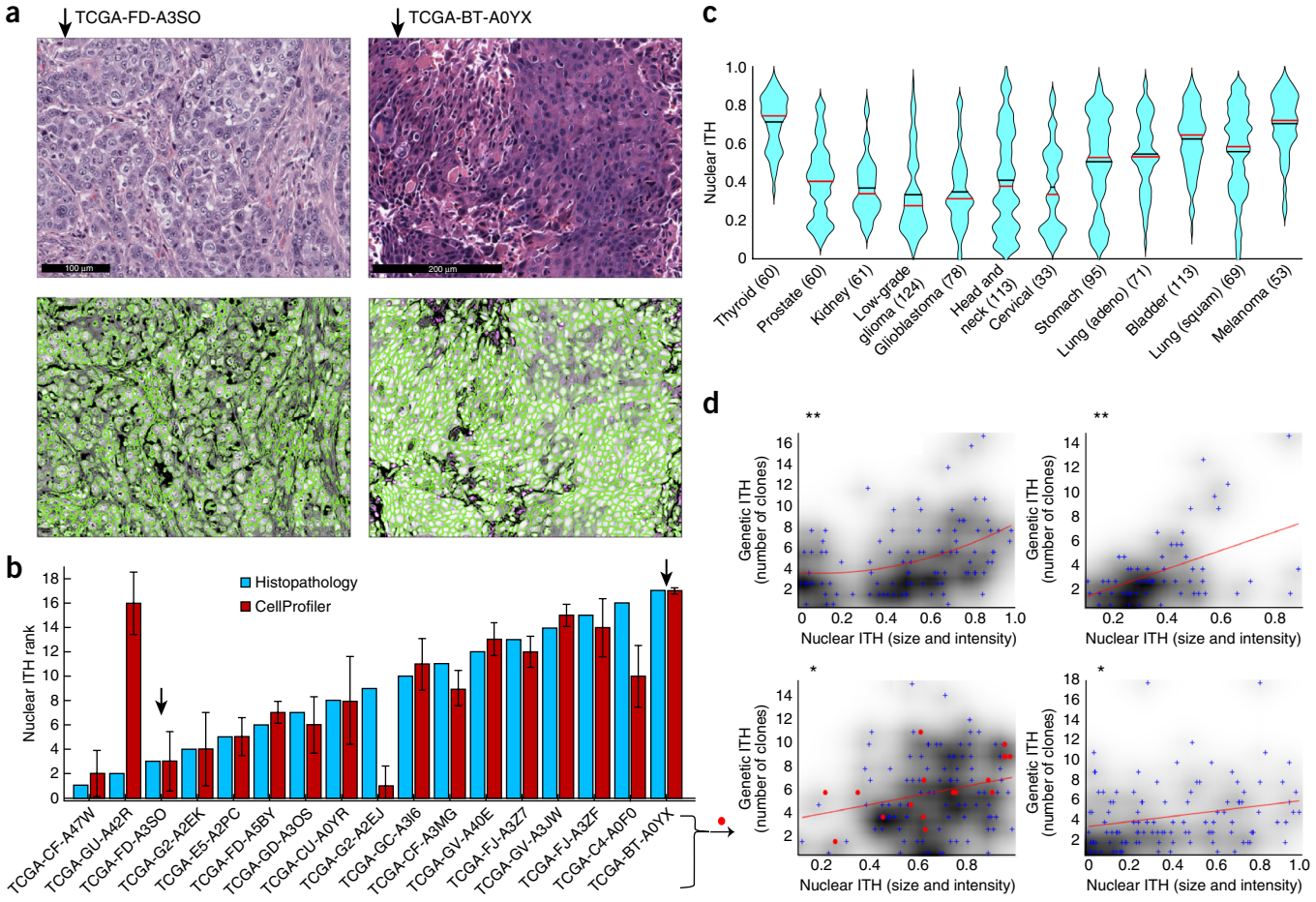

Figure 4 Intratumor nuclear diversity accompanies intratumor genetic diversity. (**a**,**b**) Representative images (**a**) and quantification (**b**) of nuclear ITH from whole-slide images of H&E-stained tumor samples. Representative staining regions are shown for two bladder cancer specimens (**a**; top). The lesion on the left (TCGA-GD-A3SO) demonstrates monomorphic high-grade nuclei with open chromatin and prominent nucleoli, whereas the lesion on the right (TCGA-BT-A0YX) demonstrates nuclei that vary from small, with condensed chromatin, to very large, with open chromatin (anisochromasia). Representative images with outlined nuclei (**a**; bottom) and quantification of nuclear variability from the H&E-stained images (**b**) using CellProfiler (red bars) are shown. Black arrows represent the bladder cancer specimens shown in **a**. Error bars show the median absolute deviation of nuclear diversity rankings on the basis of measurements of three different descriptive aspects of nuclear size and staining intensity. Independent ranking of intratumor nuclear diversity across these 17 bladder cancer specimens by an expert histopathologist (blue bars) validates the automated nuclear-diversity measures ($\rho = 0.62$; $P = 0.009$). In **a**, scale bars, 100 µm (left), 200 µm (right). (**c**) Violin plots of nuclear diversity within tumor types. Nuclear diversity was normalized to account for differences in tumor purity. Tumor types were ordered according to the extent of genetic ITH (**Fig. 2b**). (**d**) Relationship between nuclear and genetic ITH in four different tumor types. Nuclear diversity per stomach adenocarcinoma (top left; $n = 95$), kidney renal clear cell carcinoma (top right; $n = 61$), bladder urothelial carcinoma (bottom left; $n = 113$) and head and neck squamous cell carcinoma (bottom right; $n = 113$) tumor sample (*x* axis; quantified as diversity in nuclear intensity and size) showing increase with increasing clone number per tumor (*y* axis; * denotes $\rho > 0.25$, $P < 0.01$; ** denotes $\rho > 0.4$, $P < 0.001$). This was also true for all of the cancers combined ($\rho = 0.243$; $P = 6.30 \times 10^{-14}$). The *P* values have not been corrected for multiple-hypothesis testing.

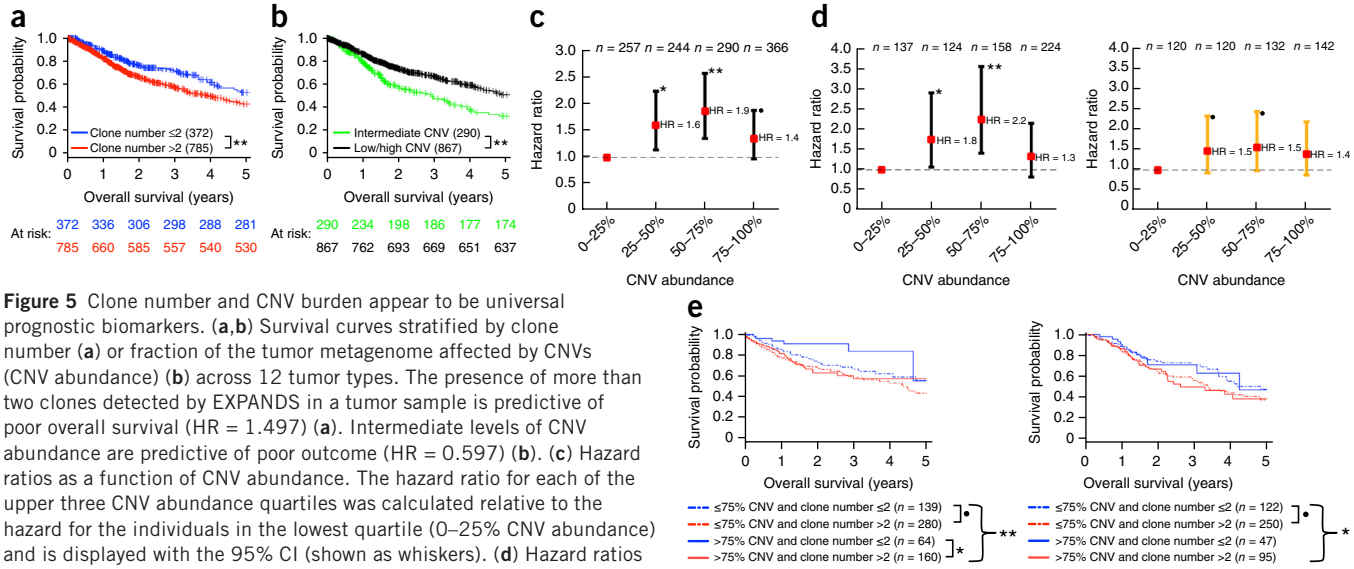

Figure 5 Clone number and CNV burden appear to be universal prognostic biomarkers. (**a**,**b**) Survival curves stratified by clone number (**a**) or fraction of the tumor metagenome affected by CNVs (CNV abundance) (**b**) across 12 tumor types. The presence of more than two clones detected by EXPANDS in a tumor sample is predictive of poor overall survival (HR = 1.497) (**a**). Intermediate levels of CNV abundance are predictive of poor outcome (HR = 0.597) (**b**). (**c**) Hazard ratios as a function of CNV abundance. The hazard ratio for each of the upper three CNV abundance quartiles was calculated relative to the hazard for the individuals in the lowest quartile (0–25% CNV abundance) and is displayed with the 95% CI (shown as whiskers). (**d**) Hazard ratios as a function of CNV abundance for untreated individuals (left) or those treated with chemo- or radiotherapy (right). Individuals with low (<25%) or high (>75%) CNV abundance progressed more slowly than individuals with intermediate CNV abundance levels (25–75%), especially within the group that did not receive adjuvant chemo- or radiotherapy. (**e**) Survival curves for untreated individuals (left) or those treated with chemo- or radiotherapy (right). Untreated and treated individuals with few clones in their tumors (blue lines) survived longer than individuals with a large number of clones detected in their tumors (red lines). Untreated individuals (left) had an especially favorable outcome when these few clones shared a large CNV burden (blue continuous line). This was not the case for treated individuals (right). All hazard ratios were calculated with log-rank tests (** $P < 0.005$; * $P < 0.05$; • $P < 0.1$). For each stratum in panels **a**,**b**,**e**, at least 50% of the 12 analyzed tumor types were represented at >5% frequency.

positive regulation of locomotion ($P = 3.39 \times 10^{-9}$), were associated exclusively with smaller-sized and medium-sized clones, but not with larger-sized clones (**Supplementary Tables 1** and **3**).

Next, we used the size rankings of CAN clones to compare cancer types (**Fig. 3d**). We measured whether a given CAN gene that is substantially associated to >1 cancer type is mutated in clones of similar sizes, regardless of cancer type. Head and neck cancer, low-grade glioma and glioblastoma showed significant similarities in CAN clone sizes with those of most of the other cancer types ($0.12 \leq \rho \leq 0.43$; $P < 0.05$), although all cancers shared their CAN clone–size signature with at least one other cancer type (**Fig. 3d**; PyClone and EXPANDS: $0.12 \leq \rho \leq 0.58$; $P \leq 0.05$).

Finally, we tested whether distinct SNV categories differ in how well they model the number of detected clones per tumor. For each cancer type, silent SNVs in non-CAN genes accounted for 0–57% of the variability in the number of clones (on average 25%). Including silent SNVs in CAN genes as predictors of clone number did not improve the model. In contrast, including nonsilent SNVs in CAN genes improved the predictions, accounting for 30% of the variability in the number of clones (log-likelihood test: $P < 0.05$; **Fig. 3e**). These results suggest that mutations driving clonal expansions are more common among nonsilent SNVs in CAN genes than among other SNV categories (see Online Methods and **Supplementary Note 2**).

Histologic ITH and proliferation rate reflect genetic ITH

A standard histomorphologic metric of tumor differentiation is the variability in nuclear size and H&E staining, which facilitates comparisons across cancers independently of tissue origin. A total of 2,231 whole-slide images of H&E-stained tumor samples were available in TCGA for 930 (80%) of the analyzed tumor samples (**Supplementary Fig. 2**). To quantify histologic ITH from these images, we measured the variabilities in nuclear size and staining intensity (**Supplementary Note 4**). For each tumor, the established image-analysis software CellProfiler[31] was used to measure the size and staining intensity of every nucleus detected in the tumor's H&E-stained images[32,33]. A histopathologist conducted an independent scoring in a blinded fashion of a subset of 17 images of H&E-stained tumor samples (**Supplementary Fig. 6**), which confirmed the accuracy of the nuclear diversity as scored by CellProfiler ($\rho = 0.62$, $P = 0.009$; **Fig. 4a**,**b**).

The extent of nuclear ITH varied between tumor types (**Fig. 4c**); greater nuclear diversity was observed with a greater number of clones detected in a tumor sample (**Fig. 4d**) in kidney cancer ($\rho = 0.413$; false-discovery rate (FDR)-adjusted $P = 0.004$), stomach cancer ($\rho = 0.406$; FDR-adjusted $P = 2.97 \times 10^{-4}$), head and neck cancer ($\rho = 0.278$; FDR-adjusted $P = 0.009$), bladder cancer ($\rho = 0.246$; FDR-adjusted $P = 0.022$) as well as across all 12 cancer types ($\rho = 0.243$; FDR-adjusted $P = 8.15 \times 10^{-13}$). Increased nuclear diversity with increasing clone number was observed for both PyClone- and EXPANDS-based clone number predictions, as well as after normalizing nuclear and genetic ITH measures to account for tumor purity (**Supplementary Table 4** and **Supplementary Fig. 7**).

We used the mRNA expression level of the proliferation marker Ki67, which was available for 854 (73%) of the samples, as a measure of a tumor's average proliferation rate[34]. Clone number was significantly correlated to Ki67 expression within low-grade glioma ($\rho = 0.18$; $P = 0.021$) and prostate cancer ($\rho = 0.21$; $P = 0.046$) as well as across cancers ($\rho = 0.31$; $P = 2.69 \times 10^{-20}$). However, tumor type–specific P values did not remain significant after FDR correction for multiple-testing ($P > 0.05$). For a subset of cancer types (the three squamous cell carcinomas of the head and neck, lung and cervix, respectively), very heterogeneous tumors (with >8 clones) had low Ki67 expression (**Supplementary Fig. 8**).

Overall, these results show that nuclear and cellular features typically associated with aggressive disease correlate with greater genetic ITH across cancer types.

Table 1 Pan-cancer multivariate Cox model of overall survival

	P value	Hazard ratio	Standard error (coefficient)	*z*-score
4 clones (Ref)	NA	1.000	NA	NA
1 or 2 clones versus Ref	0.006	0.442	0.300	−2.723
3 clones versus Ref	0.076	0.618	0.271	−1.776
5 clones versus Ref	0.007	0.450	0.296	−2.703
6 or 7 clones versus Ref	0.014	0.503	0.279	−2.463
8 or 9 clones versus Ref	0.014	0.489	0.290	−2.469
≥10 clones versus Ref	0.003	0.389	0.314	−3.011
Age at diagnosis	0.002	5.938*	0.579	3.078
Low/high CNV abundance	1.81×10^{-4}	0.129	0.548	−3.744
Pathologic stage	2.90×10^{-8}	3.339	0.217	5.548
MKI67 mRNA expression	2.21×10^{-4}	5.236	0.448	3.694
Lymphocytes (%)	0.141	0.310	0.796	−1.473
Model summary	Likelihood-ratio test = 92 on 11 degrees of freedom; $P = 6.88 \times 10^{-15}$; $n = 610$; number of events = 157			

*The hazard ratio for 'age at diagnosis' may not be reliable (test of proportional hazards: $P = 0.007$). Ref, reference (samples relative to which the risk was calculated); NA, not applicable.

Prognostic value of genomic instability and genetic ITH

We tested whether measures of genomic instability and genetic ITH (**Supplementary Table 5**) could predict overall and progression-free survival. We constructed univariate Cox models for each individual cancer type as well as pan-cancer Cox models (**Supplementary Tables 2** and **6**). Prostate adenocarcinoma and thyroid carcinoma were excluded from the cancer type–specific survival analysis owing to insufficient availability of uncensored survival information (**Supplementary Table 7**).

When considering each cancer type separately, no significant monotonic association between clone number and survival was evident ($P > 0.05$; **Supplementary Table 2**), except for gliomas (EXPANDS: $P = 0.03$, HR = 3.25; PyClone: $P = 0.04$, HR = 2.34). Across cancer types, the presence of more than two clones was associated with worse overall survival as compared to tumors in which either one or two clones were detected (EXPANDS: $P = 8.6 \times 10^{-4}$, HR = 1.49; PyClone: $P = 0.09$, HR = 1.21; log-rank test; **Fig. 5a** and **Supplementary Fig. 9**).

The association between clone number and survival was nonlinear. An increased risk with increasing clone number was observed only for tumors with up to four clones. Additional diversification (>4 clones) did not impart further risk. In fact, a tendency for reduced risk was observed among highly diverse tumors. This risk reduction did not reach significance in the univariate analysis (**Supplementary Fig. 10a**), although it was significant in the multivariate analysis (described below). The nonlinear relationship between ITH and survival was also apparent when using alternative measures of ITH (such as nuclear diversity or accounting for differential tumor purity; **Supplementary Fig. 10b–d**).

A measurement of genomic instability is the fraction of the tumor metagenome that is affected by CNVs (CNV abundance)[35]. Because genomic instability correlates with ITH (**Supplementary Fig. 10e**), we hypothesized that the increased genomic instability necessary to produce a high level of ITH (i.e., >4 detected clones) could adversely affect tumor-cell fitness after the generation of deleterious CNVs. We therefore analyzed the impact of somatic CNV abundance in the tumor metagenomes and its relation to ITH. We found that, across cancer types, low or high CNV abundance—i.e., CNVs affecting either a very low or a very high fraction of the tumor metagenome—was predictive of improved survival ($P = 5 \times 10^{-6}$, HR = 0.15; log-rank test; **Fig. 5b**). This was not the case for low or high somatic SNV abundance (FDR-adjusted $P > 0.05$; **Supplementary Table 6**). We validated this result with CNVs measured using genome-wide SNP arrays from the same tumor samples and from an independent data set consisting of 2,010 tumor samples across seven distinct cancer types. Both validation analyses confirmed that intermediate CNV abundance is associated with poor survival (**Supplementary Fig. 11** and **Supplementary Table 8**).

The highest risk was observed among individuals in whom 50–75% of the tumor metagenome was affected by CNVs in both the original and the independent data sets (**Fig. 5c** and **Supplementary Fig. 11a**,**b**). In fact, tumors with 50–75% CNV abundance did represent the highest risk group among individuals with bladder cancer, head and neck cancer, lung adenocarcinoma, stomach adenocarcinoma, cervical cancer and low-grade gliomas (**Supplementary Fig. 12**). These observations suggest the existence of an optimal degree of genomic instability that is independent of tumor type.

Of the 12 tumor types, glioblastoma was the only cancer for which >75% CNV abundance was associated with the worst prognosis (**Supplementary Fig. 12g**). Notably, 85% of individuals diagnosed with glioblastoma undergo chemo- and/or radiotherapy; thus, DNA-damaging therapy is administered more frequently for glioblastoma than for any of the other analyzed tumor types (**Supplementary Table 7**). Therefore, we verified whether or not adjuvant chemo- or radiotherapy affected the nonlinear association between CNV abundance and survival. In contrast to the 643 individuals who did not undergo chemo- or radiotherapy, the association between intermediate CNV abundance and poor survival was not significant among the 514 individuals treated with DNA-damaging agents (**Fig. 5d**). This finding was confirmed in the independent SNP-array data set, in which tumors with intermediate CNV abundance did represent the highest-risk group among untreated individuals but not among those treated with chemo- or radiotherapy (**Supplementary Fig. 11c**,**d**).

A tumor with a critical level of >75% CNV abundance per tumor metagenome may either be composed of many clones with low levels of CNV abundance per clone or of a few clones, each with high levels of CNV abundance (**Supplementary Fig. 13**). We used a clone number of 2 and a CNV abundance of 75% as thresholds to stratify untreated individuals into four groups as follows: (i) CNV abundance ≤75% and a maximum of two clones; (ii) CNV abundance ≤75% and a minimum of three clones; (iii) CNV abundance >75% and a maximum of two clones; and (iv) CNV abundance >75% and a minimum of three clones. The overall survival between individuals of these four groups was significantly different (EXPANDS: $P = 0.0015$; HR = 1.4; log-rank test). In general, and consistent with our previous observations, a low clone number was associated with a good outcome. In particular, the best outcome among the four groups was observed when a high CNV burden was shared among ≤2 clones (group (iii); **Fig. 5e** and **Supplementary Fig. 9**). After similarly stratifying individuals who had undergone chemo- and/or radiotherapy, we observed that differences in clone number, rather than CNV burden, were associated with differences in overall survival between individuals of the four groups (EXPANDS: $P = 0.038$, HR = 1.4; log-rank test; **Fig. 5e**). Stratification of tumors on the basis of PyClone-derived clone numbers also supported these conclusions, although with borderline significance ($P \leq 0.07$; **Supplementary Fig. 9a**).

To account for factors that may confound the associations observed between clinical outcome and genetic ITH, the prognostic significance of clone number and low or high CNV abundance was evaluated using multivariate Cox models (see Online Methods). All tumor types were included in a pan-cancer Cox model, except for gliomas—as the staging system is not applicable to gliomas. Across cancers, both

genomic instability and genetic ITH remained significantly associated to survival in the multivariate setting (**Table 1**). As concluded from the univariate analysis (**Supplementary Fig. 10a**,**b**), the relationship between clone number and survival was nonlinear: an increased risk with increasing clone number was only observed for up to four clones. Additional diversity (>4 clones) was associated with an increase in overall CNV burden and a significant decrease in the risk of mortality (**Supplementary Fig. 10e**,**f**). A similar scenario was observed within eight of the ten analyzed cancer types, in which the highest hazard was associated with an intermediate number (between three and five) of clones. ITH levels above or below an intermediate number of clones were associated with significantly reduced risk (multivariate Cox: HR = 0.01–0.21, $P \leq 0.05$) relative to that of the intermediate groups for head and neck cancer, melanoma and kidney cancer (**Supplementary Fig. 10g** and **Supplementary Table 9**).

DISCUSSION

Quantification of ITH is a key measure of tumor evolution. We performed a cross-sectional analysis of ITH in 1,165 cancers from 12 cancer types, revealing the extent of ITH and supporting its potential as a universal, although perhaps nonlinear, prognostic biomarker. Evidence from two tumor mixture–separation algorithms and from H&E-imaging analysis collectively indicates that ITH is a feature of the vast majority of diagnosed cancers.

To our knowledge, this is the first report of a cross-cancer correlation between genetic ITH and histopathologic ITH, suggesting that measurements from H&E-stained tumor sections can provide a proxy for genetic ITH. Currently, single-tumor samples provide the only opportunity to study genetic ITH in a large pan-cancer cohort. Measuring ITH from single-tumor samples benefits from high depth and high genomic breadth of coverage. At TCGA, the best trade-off between these two sequencing-design parameters becomes manifest in exome-sequencing data. Using exome-sequencing to quantify ITH implies that clone distinction is confined to coding regions, and thus two clones differing in only noncoding regions would be indistinguishable using this approach. Whole genomes sequenced at a higher depth and the availability of multiple samples from different geographical regions of a tumor will further improve the sensitivity in detecting small clones and will increase our ability in resolving the clonal composition of tumors and its variability across cancer types.

Our results showed that mutations in particular driver genes are associated with clones of a characteristic size, often independently of tumor type. This observation suggests that there are constraints on the order in which neoplastic cells acquire driver events[36] or that these events differ in the magnitude of the fitness advantages they provide to neoplastic cells (**Fig. 3a**). Small clones may be fit but evolve late in tumor progression. Alternatively, these clones may be less fit but can function as a 'cornucopia of evolution' from which new clones frequently emerge. Both alternatives explain why these clones are so small and why their presence is associated with poor outcome (**Fig. 3c**). Notably, the number of mutated CAN genes did not predict outcome, suggesting that the relation between the survival and presence of small clones was not confounded by the incidence of mutations in the CAN genes.

The significant association between a high clone number and the poor survival detected in the combined analysis of low-grade glioma and glioblastoma is interesting in the context of the highly variable clinical behavior of low-grade gliomas. A recent study found that histopathologic classification may overlook a subset of glioblastoma tumors and label them as low-grade gliomas[37]. Thus, knowing the extent of ITH may help improve the differential diagnosis between glioblastoma and low-grade glioma.

As previously observed in ovarian, gastric, non–small cell lung cancers and estrogen receptor (ER)-negative breast cancers[38,39], individuals with intermediate CNV burdens detected in their primary, untreated tumor had the worst overall survival. We find this association to be present across several tumor types; however, the strength of the association varies with the type of adjuvant therapy the individuals received subsequently to surgery. Our results suggest a potential advantage when tumors with intermediate levels of CNVs are treated with adjuvant chemo- and radiotherapy (**Fig. 5d**). Chemo- and radiotherapy may be particularly effective against tumors with intermediate CNV burdens by pushing them past the limit of 'tolerable' genomic instability. Our results from two distinct high-throughput technologies that measured CNV abundance in two independent pan-cancer cohorts suggest that this tolerance limit is exceeded when >75% of a tumor's metagenome is affected by CNVs, independently of cancer type. Because >37% of cancers have been shown to undergo whole-genome doubling events[10], it will be of interest to see whether these tumors have the same phenotype as tumors with a >75% CNV burden.

In light of recent evidence supporting a stronger role for CNVs rather than SNVs in developing and maintaining ITH[9], this upper limit of tolerable genomic instability may be responsible for the nonlinear association that we observed between genetic ITH and survival. Previous studies have shown that low ITH is predictive of favorable prognosis in head and neck cancer[9], leukemia[8,40] as well as in Barrett's esophagus[41,42], a premalignant condition. Consistent with these studies we found that the presence of only one or two clones is, in general, prognostic of a favorable outcome, especially when these few clones share a high CNV burden. However, diversification beyond four clones was also associated with a favorable outcome.

Individuals with high ITH may have a favorable outcome because a large number of clones can attract more immune cells. Alternatively, the observation may be, in part, because of the technical difficulty in distinguishing between ITH and genomic instability, in particular as both measures increase. Finally, the favorable outcome may be a consequence of a trade-off that exists between the chance of acquiring an advantageous alteration that initiates a new clonal expansion and the risk of generating inviable daughter cells. The observed synchronous increase in ITH and CNV burden suggests that efforts aimed at modulating this trade-off may represent a new therapeutic avenue to slow tumor evolution and improve clinical outcomes.

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

This work was supported in part by the US National Institutes of Health (NIH) (grant no. P01 CA91955 (C.C.M.), R01 CA149566 (C.C.M.), R01 CA170595 (C.C.M.), R01 CA185138 (C.C.M.), R01 CA140657 (C.C.M.), P01 HG000205 (H.P.J.), U01CA151920 (H.P.J.), U01CA17629901 (H.P.J.), R01 HG006137 (H.P.J.), R01 CA164746 (C.P.), R01 NS08061904 (C.P.) and R01 HG006137 (L.C.X.)). Additional support to C.C.M. came from the Breast Cancer Research Program Breakthrough Award (award no. BC132057), a Congressionally Directed Medical Research Program (CDMRP). Additional support to H.P.J. came from the Doris Duke Clinical Foundation Clinical Scientist Development Award, a Research Scholar Grant from the American Cancer Society (award no. RSG-13-297-01-TBG) and a Howard Hughes Medical Institute Early Career Grant. N.A. was supported

by awards from the Don and Ruth Seiler Fund and the National Cancer Institute (NCI) Cancer Target Discovery and Development (CTDD) Consortium (grant no. U01CA17629901). T.A.G. was supported by the Higher Education Founding Council for England (HEFCE). We are grateful to W. Mewes for advice on the presentation of our results and for insightful discussions about their implications; S.T. Jensen for advice on statistical data analysis; and C.W. Turck and M. Oft for reviewing the manuscript. The results presented here are in part based upon data generated by TCGA Research Network. We thank Hoffmann H. (University of Bonn, Germany) for the availability of the MATLAB function 'violin' that we used to generate the violin plots for the distribution of clone numbers and clone sizes.

AUTHOR CONTRIBUTIONS

N.A. developed analytic methods, analyzed data and wrote the manuscript. T.A.G. developed analytic methods, gave technical support and conceptual advice, and wrote the manuscript. M.J. analyzed the histopathology images and provided advice on data visualization and interpretation. L.C.X. provided advice on the choice of statistical methods and the design of the statistical analysis. C.A.A. gave technical support and conceptual advice. C.C.M. developed analytic methods, wrote the manuscript and supervised the project. H.P.J. wrote the manuscript and supervised the project. C.P. supervised the project. All authors edited the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Marusyk, A. & Polyak, K. Tumor heterogeneity: causes and consequences. *Biochim. Biophys. Acta* **1805**, 105–117 (2010).
2. Bonavia, R., Inda, M.-M., Cavenee, W.K. & Furnari, F.B. Heterogeneity maintenance in glioblastoma: a social network. *Cancer Res.* **71**, 4055–4060 (2011).
3. Wang, Y. *et al.* Clonal evolution in breast cancer revealed by single-nucleus genome sequencing. *Nature* **512**, 155–160 (2014).
4. Nowell, P.C. The clonal evolution of tumor cell populations. *Science* **194**, 23–28 (1976).
5. Greaves, M. & Maley, C.C. Clonal evolution in cancer. *Nature* **481**, 306–313 (2012).
6. McGranahan, N. *et al.* Clonal status of actionable driver events and the timing of mutational processes in cancer evolution. *Sci. Transl. Med.* **7**, 283ra54 (2015).
7. Nik-Zainal, S. *et al.* The life history of 21 breast cancers. *Cell* **149**, 994–1007 (2012).
8. Landau, D.A. *et al.* Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell* **152**, 714–726 (2013).
9. Mroz, E.A., Tward, A.D., Hammon, R.J., Ren, Y. & Rocco, J.W. Intratumor genetic heterogeneity and mortality in head and neck cancer: analysis of data from the Cancer Genome Atlas. *PLoS Med.* **12**, e1001786 (2015).
10. Zack, T.I. *et al.* Pan-cancer patterns of somatic copy number alteration. *Nat. Genet.* **45**, 1134–1140 (2013).
11. Kandoth, C. *et al.* Mutational landscape and significance across 12 major cancer types. *Nature* **502**, 333–339 (2013).
12. Ciriello, G. *et al.* Emerging landscape of oncogenic signatures across human cancers. *Nat. Genet.* **45**, 1127–1133 (2013).
13. Almendro, V. *et al.* Inference of tumor evolution during chemotherapy by computational modeling and *in situ* analysis of genetic and phenotypic cellular diversity. *Cell Rep.* **6**, 514–527 (2014).
14. Gerlinger, M. *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366**, 883–892 (2012).
15. Oesper, L., Satas, G. & Raphael, B.J. Quantifying tumor heterogeneity in whole-genome and whole-exome sequencing data. *Bioinformatics* **30**, 3532–3540 (2014).
16. Li, B. & Li, J.Z. A general framework for analyzing tumor subclonality using SNP array and DNA sequencing data. *Genome Biol.* **15**, 473 (2014).
17. Roth, A. *et al.* PyClone: statistical inference of clonal population structure in cancer. *Nat. Methods* **11**, 396–398 (2014).
18. Andor, N., Harness, J.V., Müller, S., Mewes, H.W. & Petritsch, C. EXPANDS: expanding ploidy and allele frequency on nested subpopulations. *Bioinformatics* **30**, 50–60 (2014).
19. Ha, G. *et al.* TITAN: inference of copy number architectures in clonal cell populations from tumor whole-genome sequence data. *Genome Res.* **24**, 1881–1893 (2014).
20. Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013).
21. Sathirapongsasuti, J.F. *et al.* Exome sequencing–based copy-number variation and loss of heterozygosity detection: ExomeCNV. *Bioinformatics* **27**, 2648–2654 (2011).
22. Alexandrov, L.B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
23. Vogelstein, B. *et al.* Cancer genome landscapes. *Science* **339**, 1546–1558 (2013).
24. Barber, L.J., Davies, M.N. & Gerlinger, M. Dissecting cancer evolution at the macro-heterogeneity and micro-heterogeneity scale. *Curr. Opin. Genet. Dev.* **30**, 1–6 (2015).
25. Yadav, V.K. & De, S. An assessment of computational methods for estimating purity and clonality using genomic data derived from heterogeneous tumor tissue samples. *Brief. Bioinform.* **16**, 232–241 (2015).
26. Yoshihara, K. *et al.* Inferring tumor purity, and stromal and immune cell admixture from expression data. *Nat. Commun.* **4**, 2612 (2013).
27. Tajiri, R. *et al.* Intratumoral heterogeneous amplification of *ERBB2* and subclonal genetic diversity in gastric cancers revealed by multiple ligation-dependent probe amplification and fluorescence *in situ* hybridization. *Hum. Pathol.* **45**, 725–734.
28. Sakurada, A., Lara-Guerra, H., Liu, N., Shepherd, F.A. & Tsao, M.-S. Tissue heterogeneity of *EGFR* mutation in lung adenocarcinoma. *J. Thorac. Oncol.* **3**, 527–529 (2008).
29. Imielinski, M. *et al.* Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* **150**, 1107–1120 (2012).
30. Vitale, M. Intratumor $BRAF^{V600E}$ heterogeneity and kinase inhibitors in the treatment of thyroid cancer: a call for participation. *Thyroid* **23**, 517–519 (2013).
31. Carpenter, A.E. *et al.* CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7**, R100 (2006).
32. Wang, W., Ozolek, J.A. & Rohde, G.K. Detection and classification of thyroid follicular lesions based on nuclear structure from histopathology images. *Cytometry A* **77**, 485–494 (2010).
33. Hartwell, K.A. *et al.* Niche-based screening identifies small-molecule inhibitors of leukemia stem cells. *Nat. Chem. Biol.* **9**, 840–848 (2013).
34. Yamamoto, S. *et al.* Clinical relevance of Ki67 gene expression analysis using formalin-fixed paraffin-embedded breast cancer specimens. *Breast Cancer* **20**, 262–270 (2013).
35. Cazier, J.-B. *et al.* Whole-genome sequencing of bladder cancers reveals somatic *CDKN1A* mutations and clinicopathological associations with mutation burden. *Nat. Commun.* **5**, 3756 (2014).
36. Swanton, C. Cancer evolution constrained by mutation order. *N. Engl. J. Med.* **372**, 661–663 (2015).
37. Cancer Genome Atlas Research Network. *et al.* Comprehensive, integrative genomic analysis of diffuse lower-grade gliomas. *N. Engl. J. Med.* **372**, 2481–2498 (2015).
38. Roylance, R. *et al.* Relationship of extreme chromosomal instability with long-term survival in a retrospective analysis of primary breast cancer. *Cancer Epidemiol. Biomarkers Prev.* **20**, 2183–2194 (2011).
39. Birkbak, N.J. *et al.* Paradoxical relationship between chromosomal instability and survival outcome in cancer. *Cancer Res.* **71**, 3447–3452 (2011).
40. Bochtler, T. *et al.* Clonal heterogeneity as detected by metaphase karyotyping is an indicator of poor prognosis in acute myeloid leukemia. *J. Clin. Oncol.* **31**, 3898–3905 (2013).
41. Merlo, L.M.F. *et al.* A comprehensive survey of clonal diversity measures in Barrett's esophagus as biomarkers of progression to esophageal adenocarcinoma. *Cancer Prev. Res. (Phila.)* **3**, 1388–1397 (2010).
42. Maley, C.C. *et al.* Genetic clonal diversity predicts progression to esophageal adenocarcinoma. *Nat. Genet.* **38**, 468–473 (2006).

ONLINE METHODS

Copy-number and somatic point-mutation profiles. To calculate copy-number segments we used a circular binary segmentation approach optimized for exome-sequencing input data[21] and excluded regions of low sequencing depth (≤30-fold). Somatic point mutations for each tumor sample were obtained with Mutect[43], and LOH was evaluated with JointSNVMix[44], on the basis of the tumor-derived BAM file and the matched normal BAM file. Genomic point mutations were annotated with Oncotator version 1.2. The settings that were chosen for all of the programs are noted in **Supplementary Figure 2**.

The choice of estimating copy number from exome-sequencing rather than SNP-array data was guided by: (i) the guaranteed consistency of clonal proportions between SNV and CNV data sources and (ii) the priority of having high resolution of copy numbers within exonic regions over noncoding regions. For (i), as digital signals from SNVs and CNVs are integrated to infer clonal compositions, the consistency of clonal proportions across these two data sources is critical. Estimating both copy number and SNVs from sequencing data rules out the presence of any experimental variability in the proportion of clones between the CNV and SNV data sources. For (ii), the vast majority of clonal analysis algorithms rely only on the copy-number information at loci with detected SNVs, i.e., for exome-sequencing data, within exonic regions. Although estimating copy numbers from exome sequencing leaves us blind to breakpoints in-between exons, within exonic regions it provides more 'dense' information than does SNP data[45]. For our analysis, exome-CNV calls were better suited and more precise for determining copy-number–overlapping somatic SNVs.

Inference of genetic ITH using EXPANDS. The EXPANDS version 1.4 software was used with default parameters to infer clonal subpopulations of tumor cells[46]. SNVs located within allosomes or within mitochondrial genomes were excluded. SNVs that could not be explained by a clone present in 10% or more of the sample, at a ploidy of six or less, were also excluded. For example, a SNV measured at an allelic frequency of 0.04 cannot be explained by a clone present in ≥10% of the sample, unless we assume that the high-level amplification of the wild-type allele dilutes the signal of the mutated allele. Finally, in 40.5% of tumors we found homozygous deletions affecting on average 1.5% (95% CI: 0–6.5%) of the tumor metagenome. Because a tumor cell cannot harbor both a point mutation and a homozygous deletion within the same genomic region and because EXPANDS requires both signals, these regions could not be used for inferring clonal composition. The remaining detected copy-number segments and SNVs were used to predict the number of clones that coexisted in each tumor sample, the clone size, clone-specific mutations and tumor phylogeny. Tumor phylogeny predictions were unavailable for tumor samples in which EXPANDS inferred only one single clone ($n = 160$) and unsuccessful for samples with insufficient CNVs in the tumor metagenome ($n = 120$). All other prediction aspects were available for all 1,165 tumor samples.

Inference of genetic ITH using PyClone. PyClone version 0.12.7 was used as an alternative and complementary approach to infer the clonal composition of each tumor sample. The same copy-number segments and SNVs described above were used as input, with the only modification being that the copy number of each segment was rounded to the closest integer value. Integer copy numbers are required by PyClone because the method does not model subclonal CNVs[47]. We took advantage of PyClone's flexibility in specifying prior probability estimates of possible mutational genotypes to model the uncertainty of the number of DNA copies affected by SNVs for each mutated locus. We modeled the number of SNV-harboring DNA copies for a given locus as a Poisson variable with the parameter $\lambda_S = 0.5$ for somatic SNVs and $\lambda_{LOH} = 0.7$ for germline SNVs within LOH regions. The probability that the SNV will affect a given DNA copy is higher for LOH, as an over-representation of the nonreference allele in the cancer cell relative to a normal cell is inherent to the definition of LOH[4] (see also **Supplementary Note 2**).

Quantification of nuclear ITH. CellProfiler was used to detect nuclei in representative H&E images for each tumor sample. Images processed by TCGA's centralized Biospecimen Core Resource were available for 930 tumor samples. Briefly, the nuclear diversity measure was obtained by ranking tumor samples according to the median absolute deviation (MAD) of nuclear radius and the MAD of nuclear staining intensity across detected nuclei. Image quality control and normalization steps (to account for variability in storage, handling and stain up-take) are described in **Supplementary Note 4**.

A subset of 17 individuals with bladder cancer who are representative of the full interval of measured nuclear diversity was randomly selected to validate the accuracy of nuclear-diversity measures obtained from automated image analysis. A total of 47 scanned H&E images were available for these 17 individuals at TCGA and were used to quantify nuclear diversity for each individual using CellProfiler.

Of these 47 images, 17 (one image per individual; **Supplementary Fig. 6**) were ranked according to ITH in nuclear intensity by a histopathologist (M.J.) who was blinded to the nuclear-diversity ranks estimated from the automated image analysis. Finally, the nuclear diversity rankings derived from histopathology analysis were compared to the diversity rankings obtained from the automated image analysis. Representative regions from H&E slides were visualized with OpenSlide[48].

Code availability. The CellProfiler pipeline employed to detect and measure nuclei from H&E-stained images is available at http://dna-discovery.stanford.edu/projects/completed-projects/pan-cancer-ith.html.

Normalization of ITH measures to account for differences in tumor purity. Two linear regression analyses were used to model clone number and nuclear diversity, respectively, as a function of the percentage of tumor nuclei that were estimated from histopathological examination ('tumor purity'). The clone number and nuclear diversity measures were then normalized by subtracting the tumor purity–associated effects quantified in the linear regression analyses (see also **Supplementary Fig. 7**).

Selection of cancer-driver genes. Cancer-driver gene candidates (CAN genes) were selected by using three sources: (i) the mutation frequency analysis in our cohort; (ii) the Cancer Gene Index (CGI) resource that is provided by the National Cancer Institute (NCI), is derived from literature mining and estimates how strongly a gene is associated to a given type of cancer; and (iii) the Combined Annotation–Dependent Depletion (CADD) framework resource[49], which quantifies the deleteriousness of SNVs by integrating various annotation resources into a so-called Phred C-Score. The Phred C-score is a ranking score assigned to each SNV that has been shown to strongly correlate with the pathogenicity of the SNV and with its experimentally measured regulatory effects. Altogether, a gene *G* was included as significantly associated with a given cancer type if three or more CGI entries confirmed that *G* has been experimentally associated with the cancer type and if at least three tumors of the respective type harbored somatic SNVs within the gene that have a deleteriousness score (Phred C-Score) ≥3. This integrated analysis yielded 259 CAN genes. The results presented in **Figure 3** were robust to changes in the thresholds used in the various sources (i–iii) to define the CAN genes.

CAN gene stratification and functional annotation. We classified the 259 CAN genes on the basis of their mutation occurrence in small-sized (lower quartile), medium-sized (interquartile range) and large-sized (upper quartile) clones (**Supplementary Table 1**). We used the functional annotation clustering program DAVID[50,51] on the gene sets defined by clone size. All reported *P* values were FDR-adjusted. The Drug-Gene Interaction Database[52] (DGIdb) was used to classify CAN genes as either 'druggable' ($n = 98$) or 'currently not druggable' ($n = 161$). The size of clones with CAN-gene mutations was calculated as an average across all 12 cancer types.

Modeling clone number from SNV accumulation. Linear regressions were fit to model the number of detected clones as a function of SNV accumulation. Four categories of somatic SNVs were analyzed as independent predictors of clone number: (i) nonsilent SNVs in CAN genes (total of 259 genes); (ii) silent SNVs in CAN genes; (iii) nonsilent SNVs in genes not included in the CAN-gene set (total of 23,782 genes); and (iv) silent SNVs in non–CAN genes. Four linear regression models were fitted for each cancer type as follows: the first model (M-0) used silent SNVs in non–CAN genes as an independent variable and number of clones as the dependent variable. For the other three models (M-1, M-2 and M-3), nonsilent SNV incidence in CAN genes, silent SNV incidence in CAN genes and nonsilent SNV incidence in non–CAN genes were, respectively,

doi:10.1038/nm.3984

added as independent variables to M-0. Finally, likelihood-ratio-test statistics measuring the improvement of M-1, M-2 and M-3 over M-0 were calculated. MATLAB functions 'fitlm' and 'lratiotest' were used for model-fitting and for the likelihood-ratio tests, respectively. We note that clone detection by EXPANDS and PyClone does not rely on differences between SNVs in their potential to affect selective fitness (for example, by distinguishing between silent and nonsilent SNVs or by taking into account the identity of the target gene).

Quantitative measure of accumulation of mutations in CAN genes: the progression index. Because previous evidence suggests that driver genes are mutated in a predetermined sequential order during tumor growth[53,54], the incidence of driver -gene mutations detected in an individual can predict the degree of tumor progression.

Let $G = \{g_1 \ldots g_n\}$ be the n driver genes previously derived for a given type of cancer, f_g be the mean clone size among all clones with mutations in g across all tumors and $G^I \subset G$ be the subset of genes mutated in the tumor of an individual I. The progression index of an individual I diagnosed with a given cancer type was then calculated from the incidence of nonsilent mutations in driver genes (G^I):

$$\text{Progression index} = \begin{cases} 1 - \min_{g \in G^I}(f_g), & \text{if}(G^I \neq \varnothing) \\ 0, & \text{otherwise} \end{cases}$$

The measure focuses on those particular SNVs with the least amount of representation in the data set that are the most difficult to model accurately. However, because the inference of clonal composition is restricted to regions with adequate read coverage (**Supplementary Note 1**), the measure is less prone to errors due to low sequencing coverage.

Testing the size variance of clones with nonsilent mutations in CAN genes. For every CAN gene g, let k_g be the number of clones harboring mutations in g and V_g be the variance in the size of those clones. We further calculated the variance in clone size of k_g clones randomly selected from the entire set of all detected clones and repeated this process 2,000 times, obtaining $V_R = \{V_1 \ldots V_{2000}\}$. Finally we tested whether V_g is lower than the 2,000 variances obtained from random samplings of clones (V_R) using a one-sided students t-test (MATLAB function 'ttest2'). Clone sizes were always calculated relative to the size of the founder clone for the purpose of this analysis. Variances V_g were calculated across all tumors with mutations in g, independently of cancer type.

Quantitative measures of mutation burden: SNV abundance and CNV abundance. A tumor's mutation burden was calculated for each tumor from somatic SNVs and CNVs. First, SNV abundance was calculated as the count of nonsilent somatic SNVs. In addition, we quantified low and high SNV abundance measures for each tumor, T, as:

$$|\text{SNV abundance}(T) - M_{snv}|$$

where M_{snv} is the mean SNV abundance observed among all 1,165 analyzed tumors.

Second, the fraction of the genome affected by copy-number changes, a surrogate for chromosomal instability, was calculated from the copy-number segments, s, obtained by circular binary segmentation in the following manner.

Let C_S and L_S denote the absolute copy number and the total length of segment s, respectively. Furthermore, let f_X be binary functions on C_S, defined to calculate: (i) X = amplification abundance, (ii) X = deletion abundance and (iii) X = CNV abundance as follows:

$$\text{(i)} \quad f_{\text{amplif}}(C_S) = \begin{cases} 1, & \text{if}(C_S - 2) > t \\ 0, & \text{otherwise} \end{cases};$$

$$\text{(ii)} \quad f_{\text{del}}(C_S) = \begin{cases} 1, & \text{if}(2 - C_S) > t \\ 0, & \text{otherwise} \end{cases};$$

$$\text{(iii)} \quad f_{\text{cnv}}(C_S) = \begin{cases} 1, & \text{if}\,|2 - C_S| > t \\ 0, & \text{otherwise} \end{cases}$$

where t is a threshold reflecting the minimal deviation from the diploid copy-number status above which a segment is considered to be amplified or deleted. $t = 0.25$ was chosen, which corresponded to a cellular resolution on single-copy changes affecting ≥25% of the sequenced sample. The choice of the threshold was a compromise between low false-positive rates and high resolution on subclonal CNVs and corresponded to a resolution on subclonal single-copy changes present in at least 25% of the sample. For high-level amplifications the cellular resolution increases accordingly (for example, ≥13% for twofold amplifications). Copy changes affecting a smaller fraction of cells are increasingly difficult to distinguish from noise during circular binary segmentation (see also **Supplementary Note 2**).

Then, the fractions of the tumor metagenome affected by amplifications, deletions and CNVs in general, are calculated as:

$$\text{Abundance}_f(T) = \frac{1}{\text{Genome size}} \cdot \sum_{s \in S_f} L_S$$

where $f \in \{f_{\text{amplif}}, f_{\text{del}}, f_{\text{cnv}}\}$ and $S_f = \{s \mid f(C_S) = 1\}$, is the set of non-overlapping segments for which f yields 1.

In addition, we quantified low and high CNV abundance measures for each tumor, T, as:

$$\text{Log}\,|\,\text{Abundance}\,f\_cnv(T)\,| - M_{cnv}$$

where M_{cnv} is the mean CNV abundance observed among all 1,165 tumors analyzed.

Because CNVs often cause loss or overrepresentation of one germline allele over the other, we compared the incidence of LOH between tumors with different CNV abundance levels to verify the authenticity of the CNV abundance measurements (**Supplementary Fig. 13c**).

Univariate Cox models. The prognostic significance of SNV abundance, amplification abundance, deletion abundance, CNV abundance, low or high SNV abundance, low or high CNV abundance, clone number, nuclear diversity, *MKI67* expression and the progression index was retrospectively tested. For each variable, values were scaled between 0 and 1 across individuals, to facilitate intervariable comparison of hazard ratios. The follow-up period ranged from 0–15 years (median = 1 year) and was censored at 5 years (**Supplementary Table 7**). Of the 1,165 individuals, 8 (0.7%) had missing overall and progression-free survival data, without any follow-up times available. Two clinical endpoints were evaluated: overall survival time (defined as the time between diagnosis and death due to any cause) and progression-free survival time (defined as the time between diagnosis and either the first incidence of disease progression (including tumor recurrence) or death due to any cause). The number of events that occurred during the 5-year period and the number of censored individuals are: 446 events, with 711 censored, for progression-free survival and 346 events, with 811 censored, for overall survival.

Univariate Cox models were calculated for each tumor type separately as well as in a combined analysis that included all tumor types. Missing values were omitted. Resulting P values were corrected for multiple-hypothesis testing using the FDR method (R function 'p.adjust' with method set to 'BH').

Multivariate Cox models. Multivariate Cox models of overall survival were built across and within tumor types and included the following measures as covariates: clone number, age at diagnosis, low or high CNV abundance, pathological stage, *MKI67* mRNA expression and percentage of lymphocytes in tumor sample. Clone number was normalized to account for tumor purity and was included as a categorical variable to model the nonlinear relationship between the clone number and survival. Data from individuals with an intermediate number of clones were used as reference, relative to which point-wise estimates of hazard ratios and their corresponding standard errors were calculated at: ≤2 clones, 3 clones, 4 clones, 5 clones, 6 or 7 clones, 8 or 9 clones, and ≥10 clones. All other covariates were normalized between 0 and 1 to make the hazard ratios comparable.

All features fulfilled the proportional-hazards assumption (chi-squared tests: $P \geq 0.05$), except for age at diagnosis in low-grade gliomas ($P = 0.003$) and across cancers ($P = 0.01$), *MKI67* mRNA expression in bladder cancer ($P = 0.04$) and

pathologic stage in lung adenocarcinoma ($P = 0.01$). These exceptions have been flagged in **Table 1** and **Supplementary Table 9**.

The R package 'survival' was used to build the Cox models (function 'coxph'). The R package 'smoothHR' was used to plot the hazard ratios and to test the proportional-hazards assumption (function 'smoothHR').

Tumor stratification on the basis of ITH and CNVs. Four cut points were chosen for four distinct measures: (i) the average size of clones with mutations in CAN genes (see section on progression index); (ii) the CNV burden; (iii) the clone number; and (iv) the purity-normalized clone number. Each stratum resulting from the four chosen cut points was represented by at least 47 members. Results obtained by comparing the groups were robust to variable stratification thresholds. For measures (ii)–(iv), the magnitudes of the individual cut points were selected such that they marked the transition (in CNV burden and clone number, respectively), which is accompanied by the greatest change in hazard (**Fig. 5a–d**).

Statistical analyses. The chi-squared goodness-of-fit test (MATLAB function 'chi2gof') was used to verify whether the data had a normal distribution. The two-sample *F*-test (MATLAB function 'vartest2') was used to verify whether the compared groups had equal variances. When the groups did meet the assumption of normality but did not meet the assumption of equal variances (as in **Fig. 3b**), the test statistic under the null hypothesis had an approximate Student's *t*-distribution with a number of degrees-of-freedom given by Satterthwaite's approximation (MATLAB function 'ttest2' with 'Vartype' option set to 'unequal'). If the compared groups did not meet either the assumption of normality or the assumption of equal variances, then the Kolmogorov-Smirnov test (MATLAB function 'kstest2') was used to test the null hypothesis that two samples were from the same continuous distribution.

Correlation coefficients. The Spearman correlation coefficient (ρ) was reported when at least one of the assumptions underlying the Pearson correlation (i.e., normal distribution, homoscedasticity or linearity) was not met. Otherwise we reported the Pearson correlation coefficient. Both correlation coefficients were calculated using the MATLAB function 'corr'.

43. Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013).
44. Roth, A. *et al.* JointSNVMix: a probabilistic model for accurate detection of somatic mutations in normal-tumor–paired next-generation sequencing data. *Bioinformatics* **28**, 907–913 (2012).
45. Sathirapongsasuti, J.F. *et al.* Exome sequencing–based copy-number variation and loss of heterozygosity detection: ExomeCNV. *Bioinformatics* **27**, 2648–2654 (2011).
46. Andor, N., Harness, J.V., Müller, S., Mewes, H.W. & Petritsch, C. EXPANDS: expanding ploidy and allele frequency on nested subpopulations. *Bioinformatics* **30**, 50–60 (2014).
47. Roth, A. *et al.* PyClone: statistical inference of clonal population structure in cancer. *Nat. Methods* **11**, 396–398 (2014).
48. Goode, A., Gilbert, B., Harkes, J., Jukic, D. & Satyanarayanan, M. OpenSlide: a vendor-neutral software foundation for digital pathology. *J. Pathol. Inform.* **4**, 27 (2013).
49. Kircher, M. *et al.* A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **46**, 310–315 (2014).
50. Huang, D.W., Sherman, B.T. & Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
51. Huang, D.W., Sherman, B.T. & Lempicki, R.A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).
52. Griffith, M. *et al.* DGIdb: mining the druggable genome. *Nat. Methods* **10**, 1209–1210 (2013).
53. Swanton, C. Cancer evolution constrained by mutation order. *N. Engl. J. Med.* **372**, 661–663 (2015).
54. Kandoth, C. *et al.* Mutational landscape and significance across 12 major cancer types. *Nature* **502**, 333–339 (2013).

doi:10.1038/nm.3984