nature medicine

Substantial interindividual and limited intraindividual genomic diversity among tumors from men with metastatic prostate cancer

Akash Kumar[1,9], Ilsa Coleman[2,9], Colm Morrissey[3], Xiaotun Zhang[3], Lawrence D True[4], Roman Gulati[5], Ruth Etzioni[5], Hamid Bolouri[2], Bruce Montgomery[6], Thomas White[2], Jared M Lucas[2], Lisha G Brown[3], Ruth F Dumpit[2], Navonil DeSarkar[2], Celestia Higano[6], Evan Y Yu[6], Roger Coleman[2], Nikolaus Schultz[7], Min Fang[4,8], Paul H Lange[3], Jay Shendure[1], Robert L Vessella[3] & Peter S Nelson[1–6,8]

Tumor heterogeneity may reduce the efficacy of molecularly guided systemic therapy for cancers that have metastasized. To determine whether the genomic alterations in a single metastasis provide a reasonable assessment of the major oncogenic drivers of other dispersed metastases in an individual, we analyzed multiple tumors from men with disseminated prostate cancer through whole-exome sequencing, array comparative genomic hybridization (CGH) and RNA transcript profiling, and we compared the genomic diversity within and between individuals. In contrast to the substantial heterogeneity between men, there was limited diversity among metastases within an individual. The number of somatic mutations, the burden of genomic copy number alterations and aberrations in known oncogenic drivers were all highly concordant, as were metrics of androgen receptor (AR) activity and cell cycle activity. AR activity was inversely associated with cell proliferation, whereas the expression of Fanconi anemia (FA)-complex genes was correlated with elevated cell cycle progression, expression of the E2F transcription factor 1 (*E2F1*) and loss of retinoblastoma 1 (*RB1*). Men with somatic aberrations in FA-complex genes or in ATM serine/threonine kinase (*ATM*) exhibited significantly longer treatment-response durations to carboplatin than did men without defects in genes encoding DNA-repair proteins. Collectively, these data indicate that although exceptions exist, evaluating a single metastasis provides a reasonable assessment of the major oncogenic driver alterations that are present in disseminated tumors within an individual, and thus may be useful for selecting treatments on the basis of predicted molecular vulnerabilities.

© 2016 Nature America, Inc. All rights reserved.

Personalized approaches for cancer treatment rely on the premise that defined molecular alterations in tumors will confer susceptibility to specific therapeutics, and that malignancies without the alteration will not respond[1]. Consequently, the identification of subsets of cancer patients with or without a particular oncogenic driver is predicted to enhance treatment efficacy and reduce futile interventions. This concept requires heterogeneity between individuals with respect to the composition of therapy vulnerability characteristics, and it is supported by studies that demonstrate extensive diversity in genomic aberrations and treatment responses among cancers with the same histological classification[2–8]. To achieve full success, precision oncology also requires limited molecular heterogeneity within an individual, such that the targeted vulnerability is represented in most or all tumor cells.

Although striking successes have been achieved with therapies that exploit vital drivers of tumor cell growth[9], deep molecular profiling has uncovered a substantial degree of intratumoral heterogeneity and the existence of divergent cancer clones within a primary tumor[10–12]. The extent of this diversity indicates the potential for misclassification when using current biopsy-based strategies that may under-sample tumors, and it provides clear pathways to rapid treatment resistance. To date, most studies that assess the extent of tumor heterogeneity have characterized primary tumors. However, localized carcinomas are generally managed with surgery or radiation, which are often curative. A major unmet need involves the systemic treatment of metastatic cancer, which usually involves multiple spatially distributed anatomic sites. There is limited information concerning the molecular diversity of the metastatic burden within an individual[13–15]. Such knowledge has considerable implications for the success of precision oncology.

Here we analyzed the intraindividual molecular diversity of metastatic prostate cancer (mPC), a disease that accounts for more than 27,000 deaths annually in the US[16]. The mainstay of initial therapy for mPC, which has remained essentially unchanged for more than 50 years, relies on the suppression of circulating levels of testosterone,

[1]Department of Genome Sciences, University of Washington, Seattle, Washington, USA. [2]Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA. [3]Department of Urology, University of Washington, Seattle, Washington, USA. [4]Department of Pathology, University of Washington, Seattle, Washington, USA. [5]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA. [6]Department of Medicine, University of Washington, Seattle, Washington, USA. [7]Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, New York, USA. [8]Division of Clinical Research, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA. [9]These authors jointly directed this work. Correspondence should be addressed to P.S.N. (pnelson@fhcrc.org).

Received 24 November 2015; accepted 1 February 2016; published online 29 February 2016; doi:10.1038/nm.4053

and consequently, suppression of the AR signaling program in PC cells[17]. This therapy, which was conceptually one of the first precision-medicine approaches to cancer treatment, is highly effective: response rates exceed 90%, and durations of disease control last approximately 24 months. At the time of disease progression—a nearly universal event—metastatic castration-resistant prostate cancers (mCRPC) exhibit a spectrum of recurrent molecular aberrations, some of which are predicted to confer susceptibility to targeted therapeutics[18–21]. Our objective was to determine whether the molecular composition of a single metastatic focus at a particular point in the treatment continuum provides a reasonable assessment of the major oncogenic drivers of the other dispersed metastases in an individual. We used whole-exome sequencing (WES), array CGH and transcript profiling to characterize the composition of metastatic tumors, and we evaluated the consistency and diversity of molecular aberrations within and between individuals.

RESULTS

Genomic landscapes of metastatic prostate cancers

We resected 176 primary or metastatic tumors from 63 men with mCPRC at the time of rapid autopsy, and determined mutations by WES (n = 141 tumors from 56 men), copy number alterations by

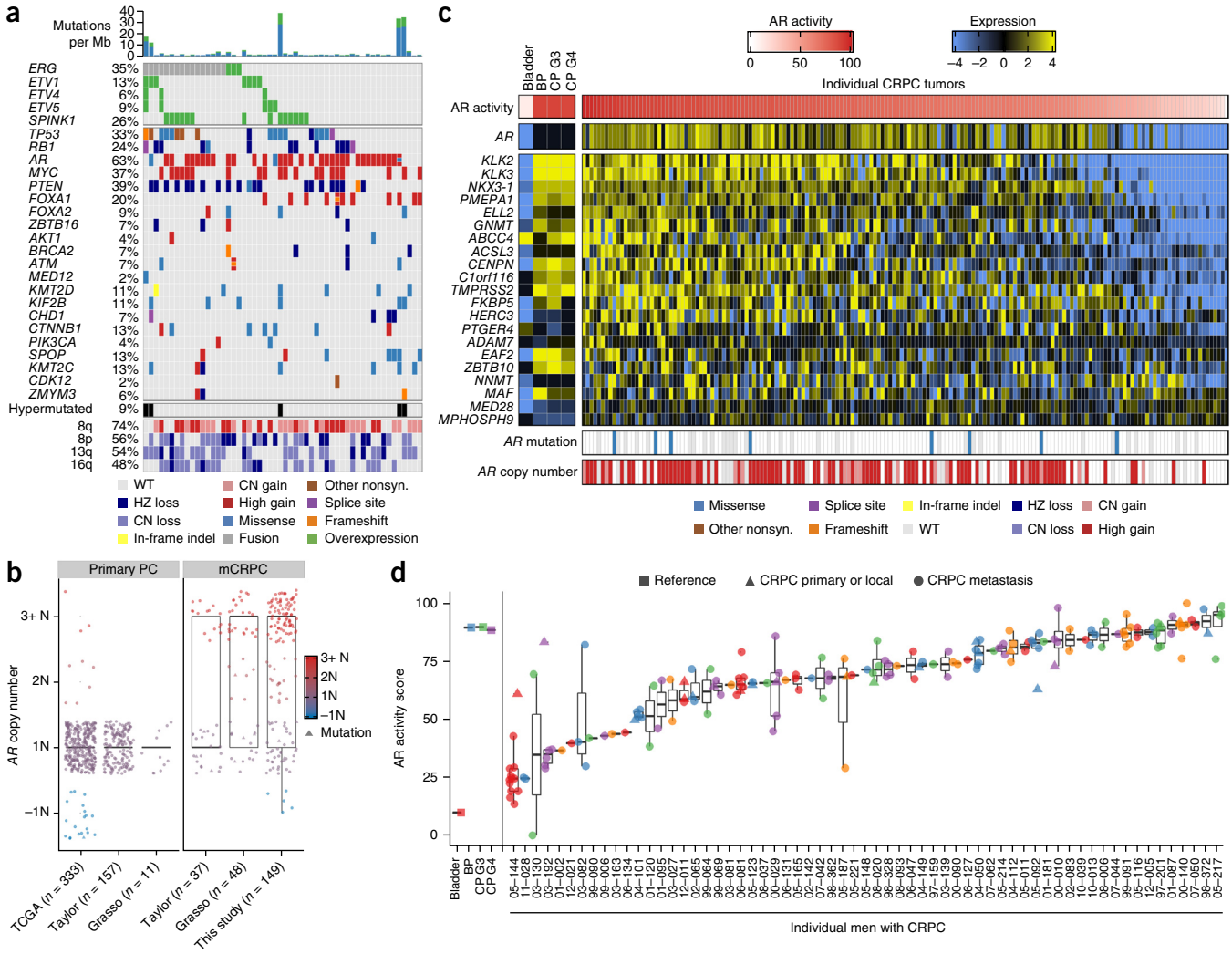

Figure 1 Integrated landscape of somatic aberrations and AR activity in mCRPC. (**a**) Recurrent somatic molecular aberrations from an index metastasis from each of 54 men with mCRPC (columns) ascertained by transcript microarray, array CGH and WES on the same tumor. Columns represent the index tumor from each individual, and rows represent specific genes. Mutations per Mb are shown in the upper histogram, with nonsynonymous mutations in blue and synonymous mutations in green. The frequency of the aberration in the cohort is shown as a percentage (%). Copy number (CN) variations common to mCRPC are shown in the lower matrix; red represents gain, blue represents loss and purple represents no loss or gain. Cases with more than one aberration in a gene are represented by split colors. Except for arm-level CN variations (CNVs), only high-level gain and loss are shown. The 54 individuals with expression, CN and mutation data are shown. HZ, heterozygosity; nonsyn., nonsynonymous mutation. (**b**) *AR* CN quantification in the TCGA[31], Taylor *et al.*[26] and Grasso *et al.*[19] studies of untreated primary PC and mCRPC. (**c**) AR activity as determined by transcript levels of 21 *AR*-regulated genes across CRPC tumors obtained at rapid autopsy. AR activity scores are also shown for microdissected cell populations of bladder urothelium (bladder); BP, benign prostate epithelium; CP, primary prostate cancer; G, Gleason pattern. *AR* somatic mutation and CN status for each tumor sample are shown in the lower matrix. (**d**) Consistency of the AR activity score for metastatic tumors within individuals, and diversity of the AR activity score in tumors between individuals. The y axis is the AR activity score and x axis shows individual men. Scores for individual tumors are plotted over a range of 0–100% as a box and whiskers plot, with the box showing the interquartile range (IQR) and the upper and lower whiskers extending to the values that are within 1.5 × IQR; data beyond the end of the whiskers are outliers and plotted as points. Colors are used to denote different individuals. Circles are metastatic tumors (n = 149) and triangles are primary tumors (n = 22) resected at rapid autopsy.

© 2016 Nature America, Inc. All rights reserved.

array CGH (n = 149 tumors from 60 men) and transcript expression by microarray hybridization (n = 171 tumors from 63 men). A small subset of tumors could not be evaluated by all three methods because of limited tumor quantity. Histological assessments of each metastasis determined that the vast majority (n = 156) were adenocarcinomas, and that 20 tumors from two men were of small-cell neuroendocrine histology. All men were treated with androgen-deprivation therapy (ADT), and after disease progression, most of them also received additional treatments, including at least one additional AR pathway–targeted agent and at least one systemic chemotherapy, most commonly docetaxel (**Supplementary Tables 1** and **2**).

We found recurrent aberrations involving genes previously reported to be altered in mCRPC, including *AR*, *ERG*, *TP53*, *RB1*, *SPOP*, *CHD1*, and *ZBTB16*, as well as gain in regions of chromosome 8q (including the *MYC* locus) and loss in regions of 8p (**Fig. 1a**). We also identified recurrent mutations in the *FOXA2* gene, which encodes a transcription factor expressed in tumors that exhibit neuroendocrine features (**Supplementary Fig. 1a**,**b**)[22,23]. Metastases from five men exhibited hypermutated genomes with complex structural aberrations in the *MSH2* and *MSH6* mismatch-repair genes[24]. All metastases and primary tumors from these men were hypermutated, which suggests that mismatch-repair deficiency occurred early in the genesis of these carcinomas.

Overall, the molecular landscapes of the lethal PCs exhibited a subset of aberrations that were observed recurrently across men, as well as combinations of events that were unique to tumors originating within an individual. When men with hypermutated tumors were excluded, the number of protein-coding mutations averaged 44 events per tumor, and the overall percentage of tumor genomes affected by structural changes (the CNA burden[25]) averaged 38%, although there were wide ranges in these metrics. Notably, most tumors from any given man exhibited similar numbers of mutations and a similar CNA burden, and there were no differences in overall mutation abundance, or in aberrations in any specific genes associated with specific visceral or bone metastatic sites (**Supplementary Fig. 1c–g**).

AR activity is maintained in the majority of CRPC metastases

Amplification or mutation of the *AR* gene occurred in 63% of men (**Fig. 1a**). Such aberrations were extremely rare in untreated primary PCs (Fisher's exact test; $P < 0.0001$) (**Fig. 1b**). By using a gene-expression signature of AR transcriptional output[26], we found that 88% of men had tumors with robust AR activity, despite extensive prior treatment to suppress AR function (**Fig. 1c**). To evaluate this more fully, we assessed *AR* expression and AR output across multiple metastases from individual men. AR activity scores (**Fig. 1d**) and *AR* expression (**Supplementary Fig. 2a**) diverged across individuals, but they were highly concordant among metastases within individuals. *AR* expression was positively associated with the amplification of the *AR* locus ($P \leq 0.05$), but it did not differ among lymph node, visceral or bone metastatic sites (**Supplementary Fig. 2b**,**c**). *AR* genomic aberrations were generally concordant within individuals: in eight men, all tumors were consistent in having no *AR* aberrations, and for 24 men, all tumors were found to have either *AR* copy gain or mutation. In nine men, there was discordance, with one or more tumors exhibiting an *AR* copy gain or mutation and one or more tumors showing no aberration. We identified one man with evidence of convergent evolution, in whom several metastases had *AR* mutation and other metastases had *AR* amplification, each potentially a contributor to treatment resistance. In two men with small-cell carcinomas confirmed by histology and immunohistochemical stains showing chromogranin expression, high neuroendocrine (NE) and low AR activity signatures were concordant across all metastases (**Supplementary Fig. 2d**).

Although *AR* expression was usually associated with AR activity, we identified outliers that showed high AR activity but absent or limited expression of the *AR* gene itself (**Fig. 2a**). Previous reports indicate that other nuclear hormone receptor genes, such as the glucocorticoid receptor (nuclear receptor subfamily 3 group C member 1 gene, *NR3C1*, also known as *GR*), the progesterone receptor (*PGR*) or the estrogen receptor (*ESR1*), could activate canonical AR target genes[27–30]. In a minority of tumors, NR3C1and ESR1 were expressed in the absence of AR (**Fig. 2b**,**c**), and thus could account for the maintenance of at least a component of the AR transcriptome.

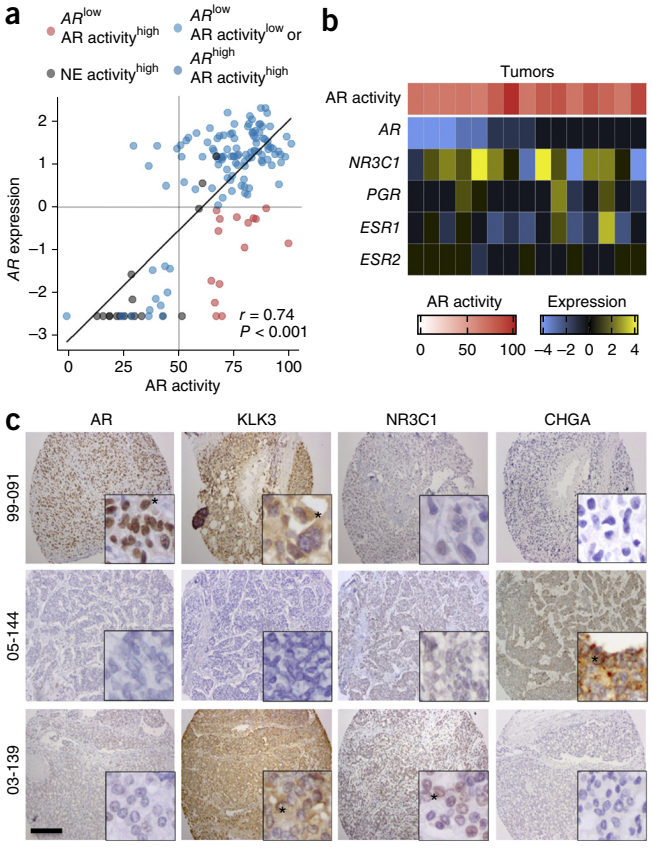

Figure 2 Relationships between AR activity and the expression of *AR* and other nuclear hormone receptor genes. (**a**) The relationship between *AR* transcript expression, plotted as mean centered $\log_2$ ratio, to the AR activity score for each tumor. Blue points represent adenocarcinomas, in which AR activity levels are correlated with *AR* level; black points are neuroendocrine tumors; red circles are tumors with high AR activity and relatively low AR expression. There was a positive overall correlation between *AR* transcript levels and AR activity score (Pearson's product moment correlation analysis, $r = 0.74$; $P < 0.001$). (**b**) Transcript levels of nuclear hormone receptors for the 15 tumors with high AR activity and low *AR* expression. *NR3C1*, nuclear receptor subfamily 3 group C member 1; *PGR*, progesterone receptor; *ESR1*, estrogen receptor 1; *ESR2*, estrogen receptor 2. (**c**) Immunohistochemical assessment of AR (also known as NR3C4), NR3C1, kallikrein-related peptidase 3 (KLK3) and chromogranin A (CHGA) proteins in mCRPC tumors from three men with different *AR* expression and AR activity relationships: 99-091 expresses AR and the AR-regulated protein KLK3, whereas 03-139 lacks AR expression but expresses NR3C1 and KLK3. Asterisk (*), cells with positive staining. Scale bar, 200 μm.

© 2016 Nature America, Inc. All rights reserved.

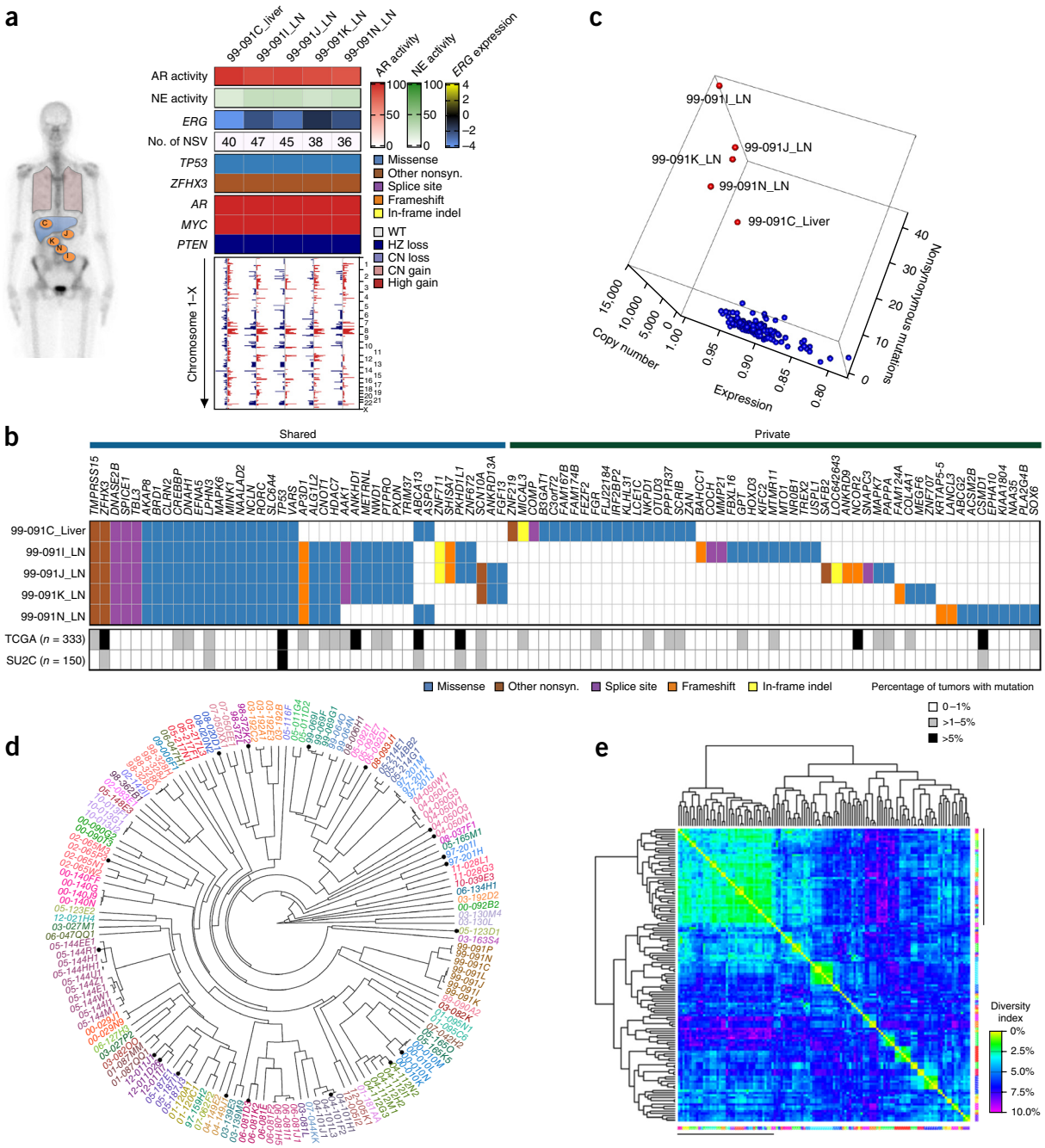

Figure 3 Molecular aberrations are shared between metastases within individuals with mCRPC. (**a**) Gene expression pathways and putative somatic driver aberrations are consistent across metastases within an individual. Shown are selected features from five metastatic tumors from patient 99-091. Activity scores comprise transcripts of genes regulated by AR (AR activity score), genes expressed in neuroendocrine carcinoma (NE activity score). Genes shown are putative PC-driver genes. *ERG* transcript levels are relative to the mean-centered ratio for the cohort. NSV, number of somatic nonsynonymous (nonsyn.) nucleotide variants. Genome-wide copy number losses (blue) and gains (red) are shown for each tumor, ordered by chromosome. Approximate physical locations of the metastatic tumors are shown in orange overlaid on the skeletal bone scan (C = liver metastases; I, J, K, N = individual LN metastases). (**b**) Distribution of all 91 nonsynonymous point mutations and indels identified in the five tumors from patient 99-091. Individual tumors comprise rows, with columns comprising individual mutations color coded by the type of mutation or its absence (white) in each tumor. The gray scale bars indicate the frequency of a mutation in the specified gene in the TCGA (primary tumors) or the SU2C-PCF (metastatic tumors) data sets. (**c**) Relationships of tumors derived from a single person (red points), relative to tumors from all other individuals (blue points), based on a calculated sample similarity score that comprised single-nucleotide variants, CNAs and gene expression, using the 133 tumors with data on all three platforms. (**d**) Unsupervised clustering of 149 tumors by correlation of genome-wide CNVs. Tumors from the same individual are designated with the same color. Black circles denote primary tumors. (**e**) Heat map of the percentage of the total of 984 distinct copy number losses and gains identified in the entire cohort of tumors that differ between samples, sorted on the basis of complete linkage dendrograms. Margin line denotes tumors with very low copy number aberration burden. Tumors from the same individual are denoted by the same color, shown on the bottom and right figure margins. The diversity index is defined as the proportion of corresponding aberrations that differed between samples: limited intraindividual diversity is reflected by green areas located at the intersection of tumors from the same individual.

© 2016 Nature America, Inc. All rights reserved.

Limited genomic diversity within patients

The substantial concordance of AR- or NE-associated gene expression across metastases within an individual prompted a further assessment of intraindividual molecular diversity. Presumptive driver aberrations identified in a representative tumor from each man were generally shared by all tumors in that individual. For example, patient 99-091 had *AR* copy gain, high AR activity, mutation in tumor protein p53 (*TP53*) and lacked an v-ets avian erythroblastosis virus E26 oncogene homolog (*ERG*) rearrangement (**Fig. 3a**). The other four evaluable tumors from 99-091 shared these and other features. CNAs were also highly concordant (**Fig. 3a**). At the mutation level, a subset of non-synonymous single nucleotide variants (NSVs) were shared, although each tumor had private mutations, of which few were predicted to be deleterious (**Fig. 3b**). To further assess the potential for these metastasis-private NSVs to influence pathogenesis, we examined WES data from 333 primary PCs from The Cancer Genome Atlas (TCGA)[31] and 150 mCRPC tumors from the Stand Up to Cancer and Prostate Cancer Foundation (SU2C-PCF) study[20]. Of 51 genes with NSVs exclusive to an individual metastasis from patient 99-091, only two, CUB and Sushi multiple domains 1 (*CSMD1*) and nuclear

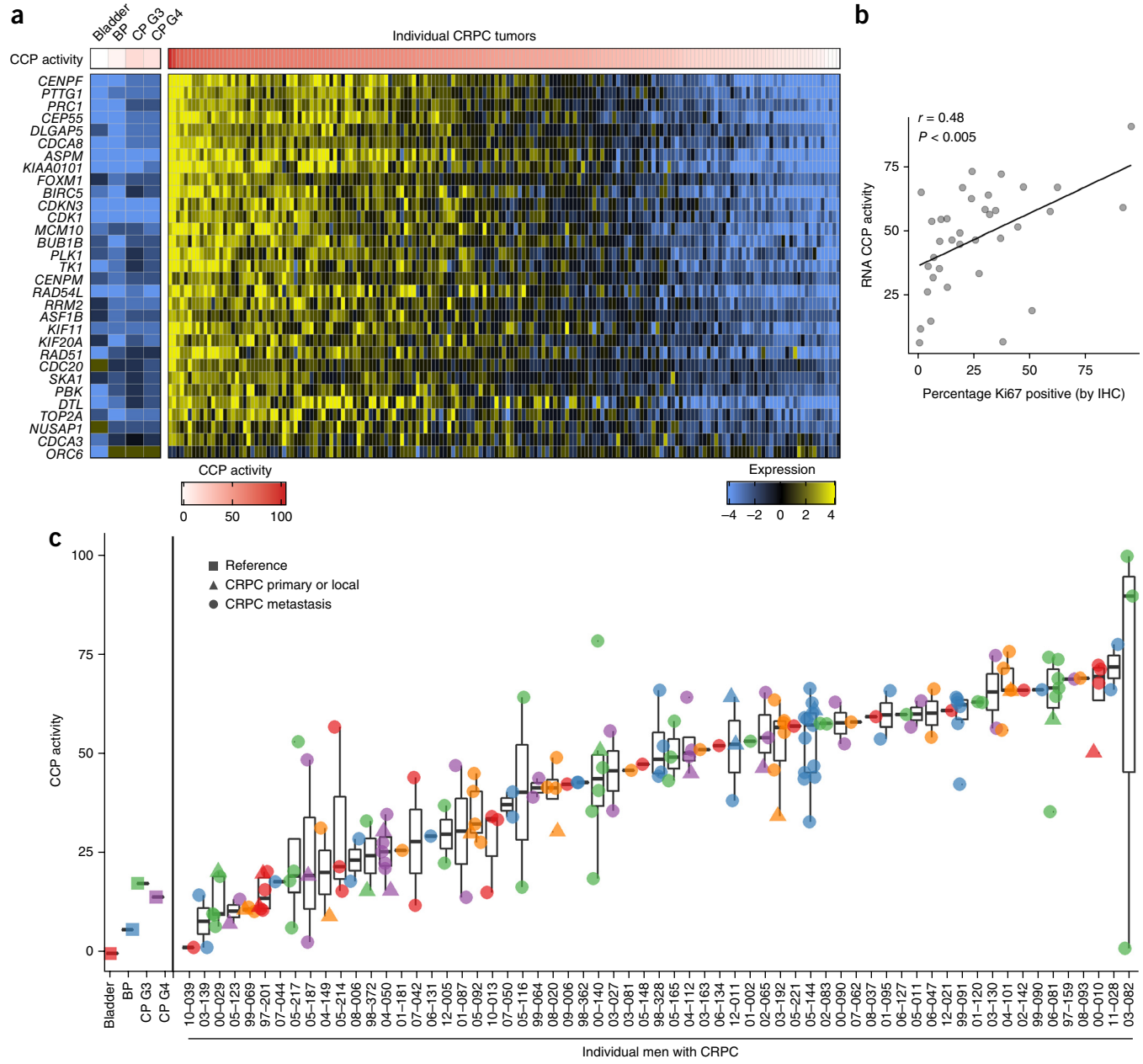

Figure 4 Tumor cell cycle activity within and between individuals. (**a**) CCP scores are based on the expression of 31 cell cycle–associated transcripts and vary across tumors. The genes comprising the CCP score are shown with expression levels colored to reflect high (yellow) or low (blue) transcript abundance. Columns represent individual tumors, excepting bladder, benign prostate epithelium (BP), primary PC with Gleason grade 6 (CP G3) and primary PC with Gleason pattern ≥ 7 (CP G4), which are composites of 10–20 samples. The 171 CRPC tumors from 63 men with expression data are shown. (**b**) Correlation between CCP score and Ki67-specific immunohistochemistry (IHC) (Pearson's product moment correlation analysis, $r = 0.48$; $P < 0.005$) in 36 tumors with matching protein and RNA expression data. (**c**) CCP scores (*y* axis) for each of 171 tumors grouped by patient (*x* axis) are plotted as a box and whiskers, of IQR and 1.5 × IQR, respectively; data beyond the end of the whiskers are outliers and plotted as points.

© 2016 Nature America, Inc. All rights reserved.

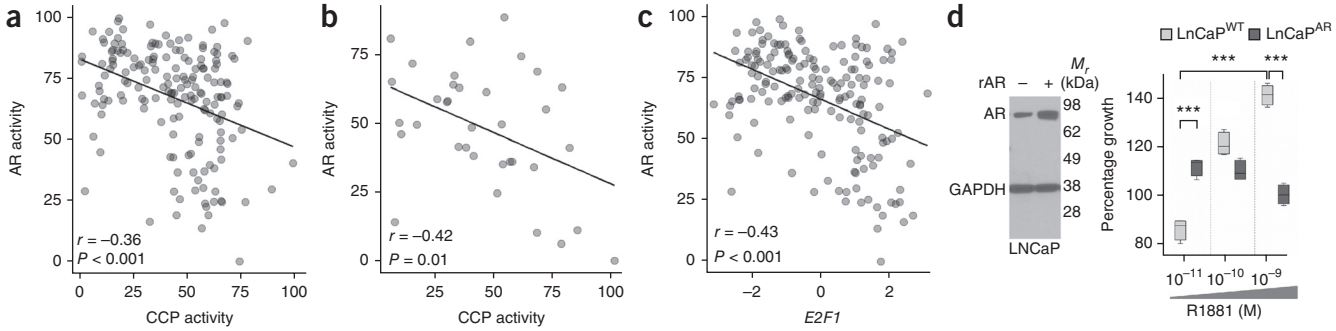

Figure 5 CCP activity and *E2F1* expression are inversely associated with AR activity. (**a**) CCP scores are inversely associated with AR activity scores, as determined by the expression levels of 21 *AR*-regulated genes in 171 tumors (Pearson's product moment correlation analysis, $r = -0.36$; $P < 0.001$). (**b**) Inverse association between CCP scores and AR activity scores was also observed in a published data set[19] of 35 CRPC tumors (Pearson's product moment correlation analysis, $r = -0.42$; $P = 0.01$). (**c**) AR activity is inversely associated with *E2F1* transcript expression in 171 tumors (Pearson's product moment correlation analysis, $r = -0.43$; $P < 0.001$). (**d**) *AR* expression level influences cellular responses to the AR ligand R1881. Left, representative western blot ($n = 2$) of AR and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) protein levels in LNCaP^{WT} cells and in LNCaP^{AR} cells engineered to overexpress AR (rAR). Right, LNCaP^{WT} and LNCaP^{AR} cell growth measured 72 h after exposure to the indicated R1881 concentration; $n = 4$ in each group, box and whiskers plots showing IQR and min to max. *$P < 0.0001$ by two-sample t test.

© 2016 Nature America, Inc. All rights reserved.

receptor corepressor 2 (*NCOR2*), were altered at frequencies that exceeded 5% in either the TCGA or the SU2C-PCF tumors. Similar results were observed across multiple tumors from other men (**Supplementary Figs. 3** and **4**). Collectively, these data suggest that most mutations unique to an individual tumor are not drivers of tumor behavior.

By using an integrated analysis that incorporated CNAs, mutations and gene expression, we found that all tumors from a given individual congregated closely with other tumors from the same individual, and that they were distinct from tumors in other men (**Fig. 3c** and **Supplementary Figs. 3** and **4**). Unsupervised clustering based on CNA profiles grouped tumors from the same individual together with nearly 100% accuracy (**Fig. 3d**). Of note, when the primary tumor removed at the time of rapid autopsy was analyzed, it grouped with the metastatic tumors from the same individual, rather than with primary tumors from different men.

We next evaluated the intraindividual consistency of a recurrent driver aberration, the *TMPRSS2-ERG* gene fusion, which is present in approximately 40% of primary PCs[32]. Fluorescence *in situ* hybridization (FISH) using *TMPRSS2* and *ERG* probes on 53 tumors from 13 men found 100% concordance, all positive or all negative, between all metastases within an individual (**Supplementary Table 3**). By using array CGH to identify chromosomal loss between the *TMRPSS2* and *ERG* genes, we found that 32 of 34 men, including the 13 men with FISH results, had 100% concordance across all metastases (94%), with divergence in two men. For one individual, 00-029, FISH demonstrated that all tumors were positive for a *TMPRSS2-ERG* rearrangement.

Clustering tumors based on expression profiles also grouped tumors by person of origin, excepting bone metastasis (**Supplementary Fig. 5a**). This may reflect inherent tumor biology, microenvironment effects, the overall lower tumor cell content of bone metastasis or technical factors that are associated with processing for transcript quantitation. Although we and others have reported differences in the expression of specific genes between bone and visceral sites[33], we did not identify mutations or CNAs associated with metastasis location.

We next determined the overall molecular similarity and diversity of tumors by using a statistical framework. We focused on a subset of 38 men from whom multiple tumors per individual were successfully sequenced. Of the 123 tumors from this cohort, 5,389 genes harbored a somatic NSV, of which 2,843 were from hypermutated tumors, leaving 2,546 NSVs derived from 34 men and a total of 112 tumors. Collectively, there were 984 regions with defined CNAs in these tumors. We compared the NSVs and CNAs in each tumor with every other tumor, and we determined that the median percent of corresponding aberrations that differed between tumors from the same individual was 6.7%, compared to 22.1% from different individuals. The median interindividual versus intraindividual tumor difference in aberrations was 15.7% (95% CI 15.5–15.9%; $P < 0.001$; two-sided two-sample Wilcoxon rank-sum test). Heat maps with complete linkage dendrograms using CNAs (**Fig. 3e**) or mutations (**Supplementary Fig. 5b**) clustered tumors according to the person of origin.

Cell cycle progression inversely correlates with AR activity

The concordance of AR activity or NE activity in tumors derived from the same person prompted further analyses to determine whether a relationship existed between these programs, as well as a key metric of treatment resistance: tumor cell proliferation. We evaluated a 31-gene signature of cell cycle progression (CCP) that was shown to be associated with PC mortality when assessed in primary tumors[34]. A wide-range of CCP activity was evident across the CRPC tumors (**Fig. 4a**). We confirmed that CCP activity was associated with cell proliferation, as determined by Ki67 immunohistochemistry on the same tumor samples ($r = 0.48$; $P < 0.005$)(**Fig. 4b** and **Supplementary Fig. 5c**). CCP activity was more concordant among tumors in an individual than tumors between different individuals (**Fig. 4c**). There were modest positive associations between CCP activity and mutation burden ($r = 0.21$; $P = 0.023$) (**Supplementary Fig. 5d**,**e**).

Contrary to our expectations, AR activity was inversely associated with CCP activity ($r = -0.33$; $P < 0.001$) (**Fig. 5a**), a finding confirmed in an independent data set comprising mCRPC tumors (**Fig. 5b**)[19]. AR activity was also inversely associated with *E2F1* expression, a key regulator of cell proliferation ($P < 0.001$) (**Fig. 5c**). To determine the functional relevance of this correlation, we compared the proliferation of wild-type (WT) cells of the androgen-responsive prostate-cancer cell line LNCaP, expressing endogenous AR (LNCaP^{WT}), to that of LNCaP cells engineered to overexpress AR (LNCaP^{AR}). Proliferation of LNCaP^{AR} cells was significantly greater than that of LNCaP^{WT} cells at very low AR ligand concentrations of 10^{-11}M of

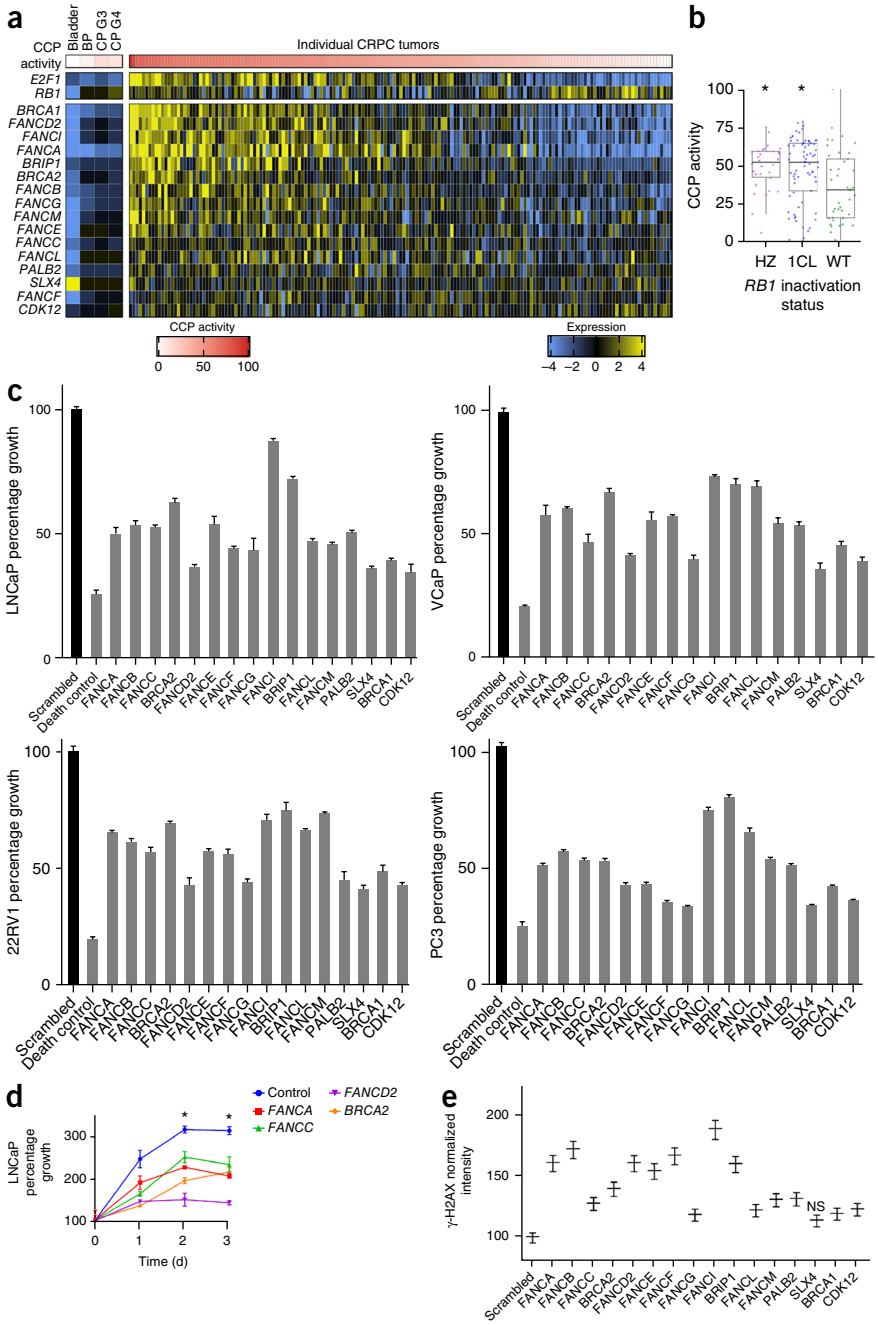

Figure 6 Expression levels of FA-complex genes are associated with CCP and *E2F1* expression. (**a**) CCP scores for each tumor are ordered from high (left) to low (right). Corresponding FA-complex genes expressed in each tumor are colored to reflect high (yellow) or low (blue) transcript abundance. Columns represent individual tumors, excepting bladder; BP, benign prostate; CP G3, primary PC with Gleason pattern 3; CP G4, primary PC with Gleason pattern 4, which are composites of 10–20 samples. The 171 CRPC tumors from 63 men with expression data are shown. (**b**) Higher CCP score is associated with heterozygous and homozygous *RB1* inactivation by copy loss and/or mutation. $^*P < 0.01$ by pairwise *t* test to wild-type (WT) group. HZ, homozygous loss; 1CL, heterozygous loss. (**c**) Suppression of individual FA-complex genes reduces the proliferation of multiple prostate cancer cell lines. Cell numbers were measured 5 d after introducing siRNAs that target the specified genes, relative to scrambled control siRNAs. All gene knockdowns produced a significant reduction in growth compared to scrambled control (two-sample *t* test; $P < 0.05$; error bars are mean of $n = 3$ biological replicates per gene and standard deviation of measurements). (**d**) Time course of LNCaP cell growth after the suppression of individual FA-complex genes. Growth curves measured over 3 d after introducing siRNAs targeting the specified genes. $^*P < 0.05$ in percentage growth compared to control for each time point using two-sample *t* tests (error bars are mean of $n = 3$ biological replicates per gene and standard deviation of measurements). (**e**) Assessment of DNA damage in LNCaP cells by γ-H2AX assays after the knockdown of individual FA-complex genes by siRNA and exposure to 50 μM carboplatin. All FA knockdowns produced significantly greater γ-H2AX than did scrambled siRNA controls ($P < 0.05$ by two-sample *t* test), excepting SLX4 structure-specific endonuclease subunit (*SLX4*), NS = not significant. Plots box and whiskers plots showing median and min to max of three replicates. (**f**) Kaplan-Meier plot comparing the duration of carboplatin treatment for 21 men with mCRPC with or without a germline or somatic alteration in genes involved in DNA repair: *BRCA2*, *PALB2*, or *ATM* ($P = 0.02$ by log-rank test).

the AR agonist metribolone (hereafter known as R1881) ($P < 0.0001$; **Fig. 5d**) that approximate castrate serum testosterone concentrations achieved with ADT[35]. However, at higher ligand levels (10^{-9} M), approximating concentrations in a noncastrated male, the growth of $LNCaP^{WT}$ cells substantially exceeded that of $LNCaP^{AR}$ cells, whose proliferation was suppressed by the higher androgen concentration (**Fig. 5d**). After the introduction of AR into AR-null human prostate cancer PC3 cells, 10^{-9} M R1881 also significantly inhibited their growth ($P = 0.02$) (**Supplementary Fig. 5f**). These findings support the concept that in addition to oncogenic roles, the AR may have activities that suppress aspects of tumor progression[36,37].

CCP correlates with FA gene expression

We next sought to identify genes with expression levels that are positively associated with elevated cell proliferation. We computed genome-wide correlations of CCP score to each gene, and found that elevated expression of FA-complex components, including *FANCA*, *FANCI*, *FANCD2*, *BRCA1* and *BRCA2*, were associated with tumors that exhibit high CCP activity (**Fig. 6a**). Genes comprising the FA pathway were expressed at low levels in untreated primary PCs, which also

© 2016 Nature America, Inc. All rights reserved.

had low CCP scores (**Fig. 6a**). For several FA genes we found significant relationships between copy number and transcript levels, including *FANCA* ($P < 0.006$), *BRCA1* ($P < 0.005$) and *PALB2* ($P < 0.005$), although most tumors did not have alterations in genes encoding FA components (**Supplementary Fig. 6a–c**). A subset of FA genes is positively regulated by E2F transcription factors, primarily by E2F1 (refs. 38,39). We found a significant relationship between high CCP activity and loss of the *RB1* tumor-suppressor locus ($P < 0.01$) (**Fig. 6b**), which was also associated with E2F1 expression and which remained significant after removing the highly proliferative NE metastases[40]. Tumors with high E2F1 expression also had elevated CCP scores ($r = 0.8$; $P < 0.001$) and elevated activity of a signature comprising 15 FA-pathway genes ($r = 0.78$; $P < 0.001$) (**Supplementary Fig. 6d–g**).

To determine whether FA components can influence cell proliferation, we suppressed the expression of individual FA genes by using short interfering RNAs (siRNAs) in the LNCaP, 22Rv1, VCaP and PC3 human PC cell lines. Suppression of each FA gene reduced the proliferation of each respective cell line (**Fig. 6c**,**d** and **Supplementary Fig. 6h**,**i**). Consistently with the known role of the FA complex in facilitating DNA repair, the suppression of individual FA-complex genes significantly increased the phosphorylation of H2AX (γ-H2AX), a marker of DNA double-strand breaks, in LNCaP cells after exposure to the genotoxic drug carboplatin ($P < 0.05$ for all FA genes tested, excepting *SLX4*) (**Fig. 6e**). Of the men in this rapid autopsy cohort, 20 received carboplatin chemotherapy during their clinical course. Those with a somatic aberration in a DNA repair–pathway gene—defined as a homozygous deleterious event in any FA-pathway component or a heterozygous inactivating event in the *ATM* gene, events both recently shown to be associated with responses to the poly (ADP-ribose) polymerase (PARP) inhibitor olaparib[41]—had significantly greater durations of response to carboplatin chemotherapy, as determined by time on drug treatment ($P = 0.02$) (**Fig. 6f**).

DISCUSSION

Comprehensive molecular assessments of PC have identified numerous recurrent genomic and epigenomic aberrations in genes that are implicated in driving cancer development, of which a substantial fraction have clinical implications for directing treatment[12,19,20,42–45]. These findings provide a rationale for tumor sequencing to identify the facets of tumor biology that prioritize and individualize therapeutics. However, deep sequencing of tumor genomes has also identified mutational heterogeneity, particularly within primary tumors. Branching evolution and the co-existence of multiple cancer lineages occur in a spectrum of carcinomas and have implications with respect to accurately sampling tumors and also for evolutionary mechanisms that may rapidly promote treatment resistance[11,12,46].

In the present study of mCRPC, we determined that constellations of genomic events were unique to the individual from which a tumor was derived, which demonstrates substantial interindividual heterogeneity. In contrast, the molecular aberrations identified in multiple, spatially distinct metastases derived from a single individual and sampled at a single point in time were highly concordant, although a limited number of unique events specific to any particular tumor was evident, and very few represented known oncogenic drivers. We found that overall mutation burden, CNA burden, AR activity signature and cell cycle output were all generally concordant across tumors within an individual. Our results provide further evidence for monoclonal seeding of metastatic sites, although our methods precluded assessments of their polyclonal composition or intermixing of metastases, as has been reported by others[15]. It is important to recognize that the ability of tumor clones from one site to seed other metastatic foci and to intermix would result in a degree of homogenization, leading to the practical consequence that sampling one site has the potential to interrogate the extent of mutational heterogeneity present within an individual. Further studies are required to ascertain the influence of intratumoral heterogeneity with respect to treatment responses of tumors with variable frequencies of clones that possess druggable alterations.

Although the analysis of a single metastatic tumor site clearly does underestimate the total burden of molecular aberrations found in the totality of all metastases, most drivers and actionable features are either represented as common roots across all tumors or result from convergent evolution conferred by therapeutic pressure[15], and most molecular differences between metastases do not seem to influence tumor phenotypes. Although there are exceptions, these findings suggest that clinical decision-making on the basis of a biopsy from a single metastatic site is reasonable. This assumption may not be relevant when sampling a primary tumor, in which greater molecular diversity seems to occur than in metastases[12], and it must be tempered given that only three men in our study were treated with the potent AR-pathway antagonists abiraterone or enzalutamide. These antagonists could induce diverse resistance mechanisms, including *AR* amplification, *AR* mutation and expression of *AR* splice variants that may differ across metastases within an individual. In some situations, simply determining the activity status of a particular pathway will be sufficient for directing therapy, whereas in other instances, the mechanism by which the pathway is activated—for example, a specific *AR* ligand binding site mutation—will have treatment implications.

Our findings are supported by previous studies reporting substantially greater similarities in the genomic and epigenomic compositions of metastases, relative to the more expansive diversity that is found within regions of primary tumors[10–12,45–49]. In addition to the severe bottlenecks imposed by the metastatic process, the time scales involved in the evolutionary stages of cellular transformation to invasion and metastasis may underlie the disparity in molecular heterogeneity within primary tumors and between metastases. It has been estimated that decades of chronological time are required for cells with initiating events to acquire sufficient additional mutations to endow subclones with the characteristics that enable invasion and metastatic spread[10]. By contrast, the intervals between the acquisition of metastatic capability and the ability to grow in distant organs generally range from only months to a few years[10]. Indeed, prostatic intraepithelial neoplasia and primary PCs are readily evident in men aged 30–40 years, decades before metastatic PC is commonly diagnosed[50]. These years of intraorgan tumor growth allow for substantial evolutionary diversity, clonal expansion, clonal mixing and the genesis of completely independent primary tumors.

In the present study, another bottleneck limiting tumor diversity may be treatment pressures that probably eliminate sensitive tumor clones and thus enable the expansion of resistant clones. Extreme diversity among metastases would probably result in the rapid expansion of one or a few metastases and the regression or stasis of others. Our finding that most metastatic sites within a given individual exhibited concordant proliferation rates suggests that there are limited differences across metastases with respect to treatment responses to ADT or taxane chemotherapy. This finding may reflect the fact that most successful therapies for mPC are directed toward *AR*, which functions as a core lineage oncogene[51], or toward fundamental cellular processes, such as microtubule dynamics[6], that are involved in cell division. In this context, the observation that FA-pathway activity is associated with a subset of tumors with high proliferation rates, which probably reflects

© 2016 Nature America, Inc. All rights reserved.

important roles in replication stress[52], coupled with the high frequency of somatic events involving DNA-repair proteins, supports a therapeutic approach that targets this pathway in selected individuals[41].

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Gene Expression Omnibus: Molecular profiling data have been deposited with the accession number GSE77930.

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

ACKNOWLEDGMENTS

We thank all of the men and their families who participated in this study. We thank A. McKenna and other members of the Nelson, Shendure and Vessella laboratories for helpful advice and assistance. We thank T. Taniguchi for helpful discussions. LNCaP^{AR} cells were a kind gift from C. Sawyers (Memorial Sloan Kettering Cancer Center). We thank M. Roudier, J. Noteboom, J. Kho and all other members of the tissue-acquisition necropsy team for their hard work and dedication. This work was supported by a Stand Up To Cancer–Prostate Cancer Foundation Prostate Dream Team Translational Cancer Research Grant. Stand Up To Cancer is a program of the Entertainment Industry Foundation administered by the American Association for Cancer Research (SU2C-AACR-DT0712) (B.M., N.S. and P.S.N.), and also by awards from the US National Institutes of Health: Pacific Northwest Prostate Cancer SPORE grant P50 CA097186 (P.S.N., B.M., L.D.T., R.L.V. and C.M.), P01 CA163227 (P.S.N., R.L.V. and C.M.), P01 CA85859 (R.L.V., C.M., M.F. and P.S.N.), US Department of Defense awards W81XWH-15-1-0562 and W81XWH-15-1-0430 (P.S.N., R.F.D., I.C., R.C. and N.S.) a Fred Hutchinson Cancer Research Center Solid Tumor Translational Research award (P.S.N., H.B. and B.M.) and the Prostate Cancer Foundation. J.S. and N.S. are supported by Prostate Cancer Foundation Young Investigator awards.

AUTHOR CONTRIBUTIONS

A.K., T.W. and I.C. coordinated overall sequencing and bioinformatics analysis. I.C. and N.S. coordinated data deposition, assembly, figures and tables. R.G., R.E., I.C. and H.B. performed statistical analyses. A.K., I.C., R.C., R.F.D., T.W., N.D. and J.S. performed sequencing and analyses. L.D.T., M.F. and X.Z. coordinated central pathology review, FISH and immunohistochemistry studies. J.M.L. conducted gene-manipulation studies and growth assays. R.L.V., L.G.B. and C.M. coordinated clinical enrollment and tissue procurement. P.H.L., C.H., E.Y.Y., P.S.N. and B.M. enrolled patients and provided clinical insights. P.S.N., I.C., J.S. and A.K. wrote the manuscript, which all authors reviewed.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

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

1. National Research Council (U.S.). Committee on A Framework for Developing a New Taxonomy of Disease. *Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease*. (National Academies Press, 2011).
2. Perou, C.M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747–752 (2000).
3. Frattini, V. *et al.* The integrated landscape of driver genomic alterations in glioblastoma. *Nat. Genet.* **45**, 1141–1149 (2013).
4. Sato, Y. *et al.* Integrated molecular analysis of clear-cell renal cell carcinoma. *Nat. Genet.* **45**, 860–867 (2013).
5. Verhaak, R.G. *et al.* Prognostically relevant gene signatures of high-grade serous ovarian carcinoma. *J. Clin. Invest.* **123**, 517–525 (2013).
6. Tannock, I.F. *et al.* Docetaxel plus prednisone or mitoxantrone plus prednisone for advanced prostate cancer. *N. Engl. J. Med.* **351**, 1502–1512 (2004).
7. de Bono, J.S. *et al.* Abiraterone and increased survival in metastatic prostate cancer. *N. Engl. J. Med.* **364**, 1995–2005 (2011).
8. Scher, H.I. *et al.* Increased survival with enzalutamide in prostate cancer after chemotherapy. *N. Engl. J. Med.* **367**, 1187–1197 (2012).
9. Druker, B.J. *et al.* Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells. *Nat. Med.* **2**, 561–566 (1996).
10. Yachida, S. *et al.* Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature* **467**, 1114–1117 (2010).
11. Gerlinger, M. *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366**, 883–892 (2012).
12. Cooper, C.S. *et al.* Analysis of the genetic phylogeny of multifocal prostate cancer identifies multiple independent clonal expansions in neoplastic and morphologically normal prostate tissue. *Nat. Genet.* **47**, 367–372 (2015).
13. Liu, W. *et al.* Copy number analysis indicates monoclonal origin of lethal metastatic prostate cancer. *Nat. Med.* **15**, 559–565 (2009).
14. Haffner, M.C. *et al.* Tracking the clonal origin of lethal prostate cancer. *J. Clin. Invest.* **123**, 4918–4922 (2013).
15. Gundem, G. *et al.* The evolutionary history of lethal metastatic prostate cancer. *Nature* **520**, 353–357 (2015).
16. Siegel, R.L., Miller, K.D. & Jemal, A. Cancer statistics, 2015. *CA Cancer J. Clin.* **65**, 5–29 (2015).
17. Huggins, C. & Hodges, C.V. Studies on prostate cancer 1: the effect of castration, of estrogen and of androgen injection on serum phosphatases in metastatic carcinoma of the prostate. *Cancer Res.* **1**, 293–297 (1941).
18. Kumar, A. *et al.* Exome sequencing identifies a spectrum of mutation frequencies in advanced and lethal prostate cancers. *Proc. Natl. Acad. Sci. USA* **108**, 17087–17092 (2011).
19. Grasso, C.S. *et al.* The mutational landscape of lethal castration-resistant prostate cancer. *Nature* **487**, 239–243 (2012).
20. Robinson, D. *et al.* Integrative clinical genomics of advanced prostate cancer. *Cell* **161**, 1215–1228 (2015).
21. Mosquera, J.M. *et al.* Concurrent *AURKA* and *MYCN* gene amplifications are harbingers of lethal treatment-related neuroendocrine prostate cancer. *Neoplasia* **15**, 1–10 (2013).
22. Qi, J. *et al.* Siah2-dependent concerted activity of HIF and FoxA2 regulates formation of neuroendocrine phenotype and neuroendocrine prostate tumors. *Cancer Cell* **18**, 23–38 (2010).
23. Gupta, A. *et al.* Mash1 expression is induced in neuroendocrine prostate cancer upon the loss of Foxa2. *Prostate* **73**, 582–589 (2013).
24. Pritchard, C.C. *et al.* Complex MSH2 and MSH6 mutations in hypermutated microsatellite unstable advanced prostate cancer. *Nat. Commun.* **5**, 4988 (2014).
25. Hieronymus, H. *et al.* Copy number alteration burden predicts prostate cancer relapse. *Proc. Natl. Acad. Sci. USA* **111**, 11139–11144 (2014).
26. Taylor, B.S. *et al.* Integrative genomic profiling of human prostate cancer. *Cancer Cell* **18**, 11–22 (2010).
27. Sahu, B. *et al.* FoxA1 specifies unique androgen and glucocorticoid receptor binding events in prostate cancer cells. *Cancer Res.* **73**, 1570–1580 (2013).
28. Arora, V.K. *et al.* Glucocorticoid receptor confers resistance to antiandrogens by bypassing androgen receptor blockade. *Cell* **155**, 1309–1322 (2013).
29. Lai, J., Myers, S.A., Lawrence, M.G., Odorico, D.M. & Clements, J.A. Direct progesterone receptor and indirect androgen receptor interactions with the kallikrein-related peptidase 4 gene promoter in breast and prostate cancer. *Mol. Cancer Res.* **7**, 129–141 (2009).
30. Setlur, S.R. *et al.* Estrogen-dependent signaling in a molecularly distinct subclass of aggressive prostate cancer. *J. Natl. Cancer Inst.* **100**, 815–825 (2008).
31. Cancer Genome Atlas Research Network. The molecular taxonomy of primary prostate cancer. *Cell* **163**, 1011–1025 (2015).
32. Tomlins, S.A. *et al.* Recurrent fusion of *TMPRSS2* and ETS transcription factor genes in prostate cancer. *Science* **310**, 644–648 (2005).
33. Morrissey, C. *et al.* Differential expression of angiogenesis associated genes in prostate cancer bone, liver and lymph node metastases. *Clin. Exp. Metastasis* **25**, 377–388 (2008).
34. Cuzick, J. *et al.* Prognostic value of an RNA expression signature derived from cell cycle proliferation genes in patients with prostate cancer: a retrospective study. *Lancet Oncol.* **12**, 245–255 (2011).
35. Sedelaar, J.P. & Isaacs, J.T. Tissue culture media supplemented with 10% fetal calf serum contains a castrate level of testosterone. *Prostate* **69**, 1724–1729 (2009).
36. Kokontis, J.M., Hay, N. & Liao, S. Progression of LNCaP prostate tumor cells during androgen deprivation: hormone-independent growth, repression of proliferation by androgen, and role for p27Kip1 in androgen-induced cell cycle arrest. *Mol. Endocrinol.* **12**, 941–953 (1998).
37. Schweizer, M.T. *et al.* Effect of bipolar androgen therapy for asymptomatic men with castration-resistant prostate cancer: results from a pilot clinical study. *Sci. Transl. Med.* **7**, 269ra2 (2015).
38. Tategu, M., Arauchi, T., Tanaka, R., Nakagawa, H. & Yoshida, K. Systems biology-based identification of crosstalk between E2F transcription factors and the Fanconi anemia pathway. *Gene Regul. Syst. Bio.* **1**, 1–8 (2007).
39. Meier, D. & Schindler, D. Fanconi anemia core complex gene promoters harbor conserved transcription regulatory elements. *PLoS One* **6**, e22911 (2011).
40. Tan, H.L. *et al.* Rb loss is characteristic of prostatic small cell neuroendocrine carcinoma. *Clin. Cancer Res.* **20**, 890–903 (2014).
41. Mateo, J. *et al.* DNA-repair defects and olaparib in metastatic prostate cancer. *N. Engl. J. Med.* **373**, 1697–1708 (2015).
42. Lindberg, J. *et al.* The mitochondrial and autosomal mutation landscapes of prostate cancer. *Eur. Urol.* **63**, 702–708 (2013).
43. Barbieri, C.E. *et al.* Exome sequencing identifies recurrent *SPOP*, *FOXA1* and *MED12* mutations in prostate cancer. *Nat. Genet.* **44**, 685–689 (2012).
44. Robbins, C.M. *et al.* Copy number and targeted mutational analysis reveals novel somatic events in metastatic prostate tumors. *Genome Res.* **21**, 47–55 (2011).

© 2016 Nature America, Inc. All rights reserved.

45. Aryee, M.J. *et al.* DNA methylation alterations exhibit intraindividual stability and interindividual heterogeneity in prostate cancer metastases. *Sci. Transl. Med.* **5**, 169ra10 (2013).
46. de Bruin, E.C. *et al.* Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science* **346**, 251–256 (2014).
47. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
48. Wu, X. *et al.* Clonal selection drives genetic divergence of metastatic medulloblastoma. *Nature* **482**, 529–533 (2012).
49. Zhang, J. *et al.* Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. *Science* **346**, 256–259 (2014).
50. Sakr, W.A. *et al.* High grade prostatic intraepithelial neoplasia (HGPIN) and prostatic adenocarcinoma between the ages of 20–69: an autopsy study of 249 cases. *In Vivo* **8**, 439–443 (1994).
51. Garraway, L.A. & Sellers, W.R. Lineage dependency and lineage-survival oncogenes in human cancer. *Nat. Rev. Cancer* **6**, 593–602 (2006).
52. Jones, M.J. & Huang, T.T. The Fanconi anemia pathway in replication stress and DNA crosslink repair. *Cell. Mol. Life Sci.* **69**, 3963–3974 (2012).

© 2016 Nature America, Inc. All rights reserved.

ONLINE METHODS

Patient enrollment and tissue acquisition. All procedures involving human subjects were approved by the Institutional Review Board (IRB) of the University of Washington and of the Fred Hutchinson Cancer Research Center. Samples were obtained from patients who died of mCRPC and who signed written informed consent for a rapid autopsy performed within 6 h of death, under the aegis of the Prostate Cancer Donor Program at the University of Washington[53]. The inclusion criteria for the Prostate Cancer Donor Program included a diagnosis of prostate cancer, informed consent to perform an autopsy and informed consent to use biospecimens for biomedical research. Visceral metastases were identified at a gross level. Biopsies of all bone sites, except ribs (a Dyke cutter was used for rib biopsies) were obtained using a drill with an attached trephine (11 mm diameter coring drill bit). All specimens were either flash frozen and embedded in Optimal Cutting Temperature compound (OCT; Tissue-Tek) or formalin-fixed and embedded in paraffin, respectively (bone metastases were decalcified in 10% formic acid before paraffin embedding). This study included 176 localized and metastatic tumors from 63 men with CRPC and a matched benign tissue sample from each individual to facilitate the identification of somatic mutations. For the molecular studies, the investigators were blinded to sample identities, and samples were randomly assigned to whole-exome sequencing runs and microarray CGH analyses. Samples were included for the molecular studies if histological analysis estimated the presence of >50% neoplastic cells, DNA quantity to be ≥3 μg and the presence of high molecular weight DNA, as determined by gel electrophoresis.

Transcript profiling. A total of 176 tumors from 63 patients were collected for profiling by expression microarray. For five tumors, RNA was of insufficient RNA quality or quantity, leaving 171 tumors from 63 patients that were successfully profiled. A Leica CM3050S cryostat was used to cut 8 μm sections, which were collected on PEN Membrane Frame Slides (Life Technologies) and immediately fixed in 95% ethanol. Sections were briefly stained (5–10 s) with hematoxylin solution according to Mayer (Sigma-Aldrich), and then dehydrated in 100% ethanol. Laser-capture microdissection (LCM) was performed with an Arcturus Veritas instrument, using both UV cutting and infrared capture lasers, with CapSure Macro LCM Caps (Life Technologies), collecting 5,000–10,000 tumor cells per sample. Digital photographs were taken of tissue sections before, during and after LCM and assessed by a pathologist to confirm the tumor content. After LCM, captured cells were lysed in Arcturus RNA Extraction Buffer. RNA was isolated using the Arcturus PicoPure RNA Isolation Kit, and DNase was treated using the Qiagen RNase-Free DNase Set. The bone metastases were sampled using a 1 mm diameter tissue punch of the frozen OCT-embedded tumors in a −20 °C cryostat. The sample was obtained from the region of the block in which there was tumor, based on a section of an adjacent decalcified formalin-fixed, paraffin-embedded (FFPE) block. RNA was isolated from the tissue cores using the RNeasy Plus Micro Kit (Qiagen, Inc.). Subsequently, total RNA from all tumors was amplified for two rounds by using the Ambion MessageAmp aRNA Kit. Probe labeling and hybridization to Agilent 44K whole human genome expression oligonucleotide microarray slides (Agilent Technologies, Inc., Santa Clara, California) was performed in accordance with the manufacturer's suggested protocols. Fluorescent array images were collected using the Agilent DNA microarray scanner G2565BA. Agilent Feature Extraction software (Agilent Technologies) was used to grid and extract the data. Data was loess-normalized within arrays (normexp background correction with offset 50) and quantile-normalized between arrays in R, using the Limma Bioconductor package by batch. Control probes were removed, duplicate probes were averaged and spots were flagged by Agilent Feature Extraction software as being foreground feature nonuniformit, and population outliers were assigned a value of 'NA'. An additional normalization step was applied to remove systematic batch effects resulting from the date of RNA-amplification by application of the ComBat function within the sva Bioconductor package to the $\log_2$ Cy3 signal intensities. Expression microarray data are deposited in the Gene Expression Omnibus database under the accession number GSE77930.

Copy number profiling. A total of 176 tumors from 63 patients were collected for profiling by microarray CGH. For 27 tumors, DNA was of insufficient quality or quantity, which left 149 tumors from 60 patients successfully profiled. Genomic DNA was isolated from OCT sections of localized and metastatic tissues with ≥50% tumor content, 1 mm punches from frozen bone metastases, as well as patient-matched normal tissues, using the Qiagen QiaAmp DNA Micro kit that incorporated RNase treatment. DNA quality was assessed by agarose gel electrophoresis. 1 μg of high molecular weight genomic DNA from each sample was labeled by random priming using the Agilent Genomic DNA Enzymatic Labeling Kit (Cy3-dUTP). A pool of reference normal DNA (Promega) was labeled with Cy5-dUTP. Cy3 and Cy5 probes were combined and hybridized to Agilent 2 × 400K SurePrint G3 CGH microarrays and washed following the manufacturer's specifications. Fluorescent array images were collected using the Agilent DNA microarray scanner G2505C and Agilent Feature Extraction software. Control probes were removed, duplicate probes were averaged and $\log_2$ ratios were defined as gProcessedSignal/rProcessedSignal. Recurrent germline CNVs were identified across benign tissues uninvolved with cancer obtained from 56 patients from whom DNA of sufficient quality and quantity was obtained for this study, as well as the Grasso *et al.*[19] copy number data set of 28 benign samples, and a list of CNV regions defined by the Broad Institute. CGH probes mapping to sequences within these regions were removed from further analysis. The $\log_2$ ratios were corrected for the GC content wave effect using 1 Mb windows for genome GC content. Outliers were smoothed, and the normalized $\log_2$ ratios were segmented using the Circular Binary Segmentation (CBS) algorithm implemented with the DNAcopy Bioconductor package using default parameters and undo.splits = 'sdundo' (undo.sd = 1). The mode of the histogram of segment means was set to zero. To determine significantly recurrent regions of SCNA, the GISTIC 2.0 version 6.2 GenePattern module was applied to the segmented data using amplifications and deletions thresholds of 0.2, join.segment.size of four probes, a copy-ratio cap of 2.5, arm-level peel-off enabled, broad length cutoff of 0.7 chromosome arms and a confidence level of 95%. CGH array data are deposited in the Gene Expression Omnibus database under the accession number GSE77930.

Clustering. Unsupervised clustering of expression and CN profiles was performed in R on the 5000 most variable genes for each data set by using Pearson's correlation coefficient as the distance metric and average-linkage hierarchical clustering. Clusters were visualized using the ape R package (CRAN, The Comprehensive R Archive Network, https://cran.r-project.org/) Circular plots of expression and CN profiles by chromosome location were created with the OmicCircos R package version 1.8.0 (Bioconductor, http://bioconductor.org/)

Statistical analyses. Pearson's correlation coefficient was used to study the relationships between variables shown in scatterplots by using the cor.test function in R. *RB1* inactivation status was determined to be homozygous if there was a two-copy loss or one-copy loss accompanied by mutation. Single-copy loss or mutation only was categorized as heterozygous loss. The CCP activity in each of these groups was compared to the *RB1* WT group using both pairwise *t* tests and pairwise Wilcoxon rank-sum tests, using the pairwise.t.test and pairwise.wilcox.test functions in R. The pairwise *t* test and pairwise Wilcoxon rank-sum test were also used to compare the number of nonsynonymous mutations and the percentage of genome altered in different metastatic tissue sites, and all copy number versus expression comparisons. Assessments of the proportions of bone or soft-tissue metastases with either homozygous phosphatase and tensin homolog (*PTEN*) CN loss or high gain of *AR* were performed by Fisher's exact test using GraphPad Prism version 6.02 (GraphPad Software, La Jolla, California). Duration on carboplatin treatment was compared for percentages of patients harboring a DNA-repair defect with those who did not by Kaplan-Meier plot with log-rank test using GraphPad Prism version 6.02. Patients with a DNA-repair defect were defined as those with any combination of the following criteria: homozygous or heterozygous loss in *ATM* by copy-number loss and/or mutation; homozygous loss of one of the following 15 FA-complex–associated genes (*FANCA*, *FANCB*, *FANCC*, *FANCD2*, *FANCE*, *FANCF*, *FANCG*, *FANCI*, *BRIP1*, *FANCL*, *FANCM*, *PALB2*, *SLX4*, *BRCA1*); or germline mutation in *BRCA2* and/or *ATM*.

Exome sequencing. A total of 176 tumors from 63 patients were collected for whole-exome sequencing. For 35 tumors, DNA was of insufficient quality or

© 2016 Nature America, Inc. All rights reserved.

quantity, leaving 141 tumors from 56 patients that were successfully sequenced by using the Nimblegen V2 or V3 platforms (Nimblegen, Inc.) as previously described[18], with the following modification: in a subset of tumors, individually barcoded libraries were pooled in pairs before capture (**Supplementary Table 2**). Genomic DNA was isolated as described above. Sequencing was performed using the Illumina Hiseq 2000 with either 50 bp or 100 bp paired-end sequences. Reads were mapped to the human reference genome sequence (hg19) with bwa v0.7.1 (ref. 54). After the removal of PCR duplicate pairs, we performed local realignment around indels using the Genome Analysis Toolkit (GATK)[47]. We subsequently called mutations using the MuTect software package (Broad Institute, https://www.broadinstitute.org) with the following parameters: '–minimum_normal_allele_fraction 0.02–max_alt_alleles_in_normal_count 12–intervals poscont.list–fraction_contamination 0.02' and called small insertions and deletions using GATK-2.3.5 SomaticIndelDetector (Broad Institute, https://www.broadinstitute.org). The minimum phred-scaled base quality at a position to count a read was 15. A subset of paired libraries displayed a low level of cross-contamination because of barcode cross-talk. To remove common polymorphisms and to enrich for probable somatic mutations, we imposed a number of additional requirements across all samples, including requiring variants to be observed in at least 10% of reads at a position in tumor samples, requiring mutant reads to be greater than 10% of the number of mutant reads in the paired library, and removing variants present within a modified database of single-nucleotide polymorphisms (SNPs) (dbSNP v137) and matched controls that had first been stripped of all COSMIC variants. We investigated mutations for significance with MutSigCV (Broad Institute, https://www.broadinstitute.org) and using standard parameters, and we manually inspected mutations in the top 50 significant genes using the Integrative Genomics Viewer (IGV) (Broad Institute, https://www.broadinstitute.org) to remove sequence artifacts. A file of the mutations identified through whole-exome sequencing is provided in **Supplementary Table 4**.

Immunohistochemistry 155 metastatic PCa tissues (including 73 soft-tissue metastases and 82 bone metastases) from a total of 50 autopsy patients (up to four sites per patient) were embedded in paraffin. A human tissue microarray (TMA) was constructed with duplicate 1 mm cores using the aforementioned tissues. Because of insufficient tissue, not all tumors subjected to molecular profiling were included on the TMA. FFPE TMA sections (5 μm) were deparaffinized and rehydrated with three changes of xylene and graded ethanol. Antigen retrieval was performed with 10 mM citrate buffer (pH 6.0) in a high-pressure cooker for 30 min. Endogenous peroxidase and avidin/biotin were blocked respectively (Vector Laboratories, Inc.). Sections were then blocked with 5% normal goat-horse-chicken serum at room temperature for 1 h, and incubated with primary antibody at 4 °C overnight. After washing three times with 1 × phosphate buffered saline (PBS), slides were incubated with biotinylated secondary antibody (Vector Laboratories, Inc.), followed by ABC reagent (Vector Laboratories, Inc.) and stable DAB (Invitrogen Corp.). All sections were lightly counterstained with hematoxylin and mounted with Cytoseal XYL (Richard Allan Scientific). Mouse or rabbit IgG were used as appropriate for negative controls. Antibodies sources and dilutions: ER-β was from Biogenex cat#PU385-UP (1:100); ER-α was from Abcam, cat#ab173467 (1:500); NR3C1 was from Santa Cruz, cat#sc-8992 (1:75); AR was from Biogenex, cat#F39.4.1 (1:60); chromogranin A was from Dako, cat# DAK-A3 (1:100); and K_i-67 was from Dako, cat# M7240 (1:100).

Assessments of inter- and intraindividual tumor molecular heterogeneity Mutation and CNAs for 141 tumor samples from 56 patients whose samples had undergone successful sequencing and CGH analyses were initially examined separately. We identified 5,389 mutations across tumor samples. On the basis of visualization of individual profiles, we flagged five patients' tumors as hypermutated; these patients together contributed 2,843 (53%) unique mutations. To summarize heterogeneity across samples from nonhypermutated tumors, we excluded patients who provided these, and all mutations that were present only for these patients. This left 2,546 mutations from 129 tumor samples from 51 patients. Next, we identified 984 regions with defined chromosome copy gain or loss across tumor samples from all 56 patients. We collapsed copy numbers with one or two gains into a single category (and similarly for one or two losses). The resulting mutation and CNAs were then combined to create aggregate tumor samples. Finally, to compare heterogeneity across samples from the same individual to those between individuals, we restricted attention to individuals with at least two samples. This left a total of 3,530 mutations or copy number aberrations from 133 tumors from 39 patients.

We quantified heterogeneity between two tumor samples using a "diversity index," defined as the proportion of corresponding aberrations that differed between samples. For statistical inference, we randomly drew 10,000 pairs of tumor samples from the same individual, and separately, we drew another 10,000 random pairs of tumor samples from different patients. We summarized distributions of the resulting diversity index using kernel density estimates and tested for equality of distributions using a two-sided two-sample Wilcoxon rank-sum test.

As an alternative visualization of heterogeneity across tumor samples, we created heat maps of the same diversity index, which was calculated for all possible pairs of tumor samples. The heat maps were sorted on the basis of the results of a complete linkage cluster analysis, using the diversity index as the Euclidean distance between samples.

For the 3D plot of tumor molecular relationships (for example, **Fig. 2c**), we calculated a sample similarity score between the primary tumor of each patient and all other tumors using (a) single nucleotide variants, (b) CNAs and (c) gene expression. Similarity in single nucleotide alterations was measured by the number of affected genes that are shared between two tumors. Similarity in CNAs was calculated as the inner (dot) product of the vectors of GISTIC-score per-gene for each tumor. Similarity in expression was measured by Pearson's correlation coefficient across all genes.

We generated one plot for each patient. Each coordinate axis in a plot represents one of the above three similarity measures. Plots were interactively rotated in 3D to obtain the clearest viewing angle. In the example plot of **Figure 2c**, the vertex at the top marks the location of an index tumor. The red plot points mark samples from the same patient as the index tumor. The blue plot points mark all other tumor samples.

Gene manipulation and cell growth assessments. LNCaP, PC3, VCaP and 22Rv1 cells were obtained from ATCC and used within six passages of growth after receipt, or validated using STR (DNA Diagnostics) and confirmed to be mycoplasma-negative. Cell lines were grown in RPMI 1640 supplemented with 10% fetal bovine serum (FBS). siRNAs were obtained from either Sigma-Aldrich (FANC F, L, M and BRIP1) or Qiagen (all others and the Allstars negative control). Target sequences for siRNAs are provided in **Supplementary Table 5**. Cells were transfected with the siRNAs by using Liopofectamine RNAiMAX (Invitrogen), by following the manufacturer's protocol. Briefly, for gene-expression analyses, 250,000 cells were plated in a six-well dish with 50 nM RNAi and 5 μl of RNAiMAX in Optimem. After 8 h, the media was replaced with RPMI 1640 with 10% FBS. After 5 d, RNA was extracted using the RNeasy kit (Qiagen). cDNA was made using SuperScript II (Invitrogen) and oligo dT primers. Gene expression was determined using qRT-PCR. Each assay comprised three biological replicates. Reactions were performed using Power SybrGreen on an Applied Biosystems (Foster City, California) 7900HT Fast Real-Time PCR System. Cycling conditions included a 10 min incubation at 95 °C, followed by 40 cycles of 95 °C for 15 s with a 1 min extension at 60 °C. Disassociation curves were generated using the machine's default conditions. Expression was normalized to RPL13A expression. Gene-specific primer sequences are provided in **Supplementary Table 6**.

Growth assays were performed using CellTiter 96 Aqueous One Solution (Promega). For each assay and/or condition, 10,000 cells per well were transfected as above, using 10 nM RNAi and 0.2 μl of RNAiMAX in 100 μl of Optimem. After 8 h, 100 μl of RPMI 1640 with 20% FBS was added to each well. Each day, 10 μl of CellTiter 96 Aqueous One Solution was added to designated wells, incubated at 37 °C for 2 h and the absorbance was measured at 490 nM.

PC3-AR were created by cloning the androgen receptor from a normal prostate cDNA library obtained from Clontech (Mountain View, California) into the pBABE retroviral expression vector. After virus production and infection, cells were selected by using puromycin, and clones were selected and screened for AR expression. AR expression was determined by immunoblotting using an antibody against AR (GTX62599, Genetex, Irvine, California).

© 2016 Nature America, Inc. All rights reserved.

doi:10.1038/nm.4053

AR-expressing cells were assayed for their response to varying concentrations of the synthetic androgen R1881. Five thousand cells were plated in charcoal-stripped growth media for 1 d before exposure to androgen. Cell growth was determined as above, using CellTiter 96 Aqueous One Solution. Percentage of cell growth was normalized to day 0 or to vector controls. Percentage of growth in each group was compared to the control using two-sample *t* tests. Variances in each group were compared and not found to be significantly different by F test.

DNA damage and γ-H2AX assessment. Cells were transfected in 96-well plates with siRNAs, as described above. Three days after transfection, half of the plates were dosed with 50 μM of carboplatin (Sigma) or water vehicle. After 16 h of exposure to either carboplatin or vehicle, the level of phosphorylated γ-H2AX was determined by using the H2A.X (pSer139) Human In-Cell ELISA Kit (Abcam). The manufacturer's procedures were followed precisely. PCR primer sequences and siRNA sequences are in **Supplementary Tables 5** and **6**.

53. Roudier, M.P. *et al.* Phenotypic heterogeneity of end-stage prostate carcinoma metastatic to bone. *Hum. Pathol.* **34**, 646–653 (2003).
54. Li, H. *et al.* & 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

© 2016 Nature America, Inc. All rights reserved.