ARTICLE

OPEN

doi:10.1038/nature12965

Comprehensive molecular characterization of urothelial bladder carcinoma

The Cancer Genome Atlas Research Network*

Urothelial carcinoma of the bladder is a common malignancy that causes approximately 150,000 deaths per year worldwide. So far, no molecularly targeted agents have been approved for treatment of the disease. As part of The Cancer Genome Atlas project, we report here an integrated analysis of 131 urothelial carcinomas to provide a comprehensive landscape of molecular alterations. There were statistically significant recurrent mutations in 32 genes, including multiple genes involved in cell–cycle regulation, chromatin regulation, and kinase signalling pathways, as well as 9 genes not previously reported as significantly mutated in any cancer. RNA sequencing revealed four expression subtypes, two of which (papillary–like and basal/squamous–like) were also evident in microRNA sequencing and protein data. Whole–genome and RNA sequencing identified recurrent in–frame activating *FGFR3–TACC3* fusions and expression or integration of several viruses (including HPV16) that are associated with gene inactivation. Our analyses identified potential therapeutic targets in 69% of the tumours, including 42% with targets in the phosphatidylinositol–3–OH kinase/AKT/mTOR pathway and 45% with targets (including ERBB2) in the RTK/MAPK pathway. Chromatin regulatory genes were more frequently mutated in urothelial carcinoma than in any other common cancer studied so far, indicating the future possibility of targeted therapy for chromatin abnormalities.

Urothelial carcinoma of the bladder is a major cause of morbidity and mortality worldwide, causing an estimated 150,000 deaths per year[1]. Previous studies have identified multiple regions of somatic copy number alteration, including amplification of *PPARG*, *E2F3*, *EGFR*, *CCND1* and *MDM2*, as well as loss of *CDKN2A* and *RB1* (refs 2, 3). Sequencing of candidate pathways has identified recurrent mutations in *TP53*, *FGFR3*, *PIK3CA*, *TSC1*, *RB1* and *HRAS* (refs 2, 3). Whole-exome sequencing of nine bladder cancers, followed by a replication analysis of 88 cancers, identified mutations at >10% frequency in several chromatin remodelling genes: *KDM6A*, *CREBBP*, *EP300* and *ARID1A* (ref. 4). Focused molecular analyses[5,6] have delineated tumour subtypes and identified kinase-activating *FGFR3* gene fusions[7,8].

We report here a comprehensive, integrated study of 131 high-grade muscle-invasive urothelial bladder carcinomas as part of The Cancer Genome Atlas (TCGA) project. Included are data on DNA copy number, somatic mutation, messenger RNA and microRNA (miRNA) expression, protein and phosphorylated protein expression, DNA methylation, transcript splice variation, gene fusion, viral integration, pathway perturbation, clinical correlates and histopathology to characterize the molecular landscape of urothelial carcinoma. This study identifies a number of mutations and regions of copy number variation that involve genes not previously reported as altered in a significant fraction of bladder cancers. It also identifies potential therapeutic targets in most of the samples analysed.

Demographic, clinical and pathological data

Samples (from 19 tissue source sites) consisted of 131 chemotherapy-naive, muscle-invasive, high-grade urothelial tumours (T2-T4a, Nx, Mx), as well as peripheral blood ($n = 118$) and/or tumour-adjacent, histologically normal-appearing bladder tissue ($n = 23$). Cases were retained only if they met the following criteria: tumour nuclei constituted ≥60% of all nuclei; tumour necrosis was ≤20% of the specimen; and variant histologies (squamous or small cell) were ≤50% (Supplementary Information, section 'Biospecimen collection and clinical data'). Clinical and demographic characteristics are described in Supplementary Data 1.1. Five expert genitourinary pathologists re-reviewed all of the cases for multiple parameters, including the extent of variant histology (Supplementary Fig. 1.1a and Supplementary Information, section 'Biospecimen collection and clinical data').

Somatic DNA alterations

The tumours displayed a large number of DNA alterations, slightly fewer than in lung cancer and melanoma, but more than in other adult malignancies studied by TCGA (Fig. 1)[9]. On average, there were 302 exonic mutations, 204 segmental alterations in genomic copy number and 22 genomic rearrangements per sample. We analysed somatic copy number alterations (CNAs) using both SNP 6.0 arrays and low-pass whole-genome sequencing; the two were strongly concordant (Supplementary Methods 6.1 and Supplementary Fig. 6.1). There were 22 significant arm-level copy number changes (Supplementary Data 6.1.1), and GISTIC (genomic identification of significant targets in cancer) (Supplementary Methods 6.2) identified 27 amplified and 30 deleted recurrent focal somatic CNAs (Supplementary Data 6.2.1 and 6.3.1). Focal amplifications involved genes previously reported to be altered in bladder cancer (Fig. 1c and Supplementary Fig. 6.2.1) and some not previously implicated. The latter included *PVRL4*, *BCL2L1* and *ZNF703*. The most common recurrent focal deletion, seen in 47% of samples, contained *CDKN2A* (9p21.3) and correlated with reduced expression (Fig. 1 and Supplementary Fig. 2.7). Other focal deletions containing <10 genes appeared to target *PDE4D*, *RB1*, *FHIT*, *CREBBP*, *IKZF2*, *FOXQ1*, *FAM190A* (also called *CCSER1*), *LRP1B* and *WWOX*.

Whole-exome sequencing of 130 tumours and matched normal samples targeted 186,260 exons in 18,091 genes (mean coverage 100-fold, with 82% of target bases covered >30×). MuTect[10] identified 39,312 somatic mutations (including 38,012 point mutations and 1,138 indels (insertions or deletions)), yielding mean and median somatic mutation rates of 7.7 and 5.5 per megabase (Mb), respectively (Fig. 1a and Supplementary Table 2.1.1). Thirty-two genes showed statistically significant levels of recurrent somatic mutation (Fig. 1b and Supplementary Table 2.1.2) by analysis using MutSig 1.5 (refs 9, 11) (Supplementary

*A list of authors and affiliations appears at the end of the paper.

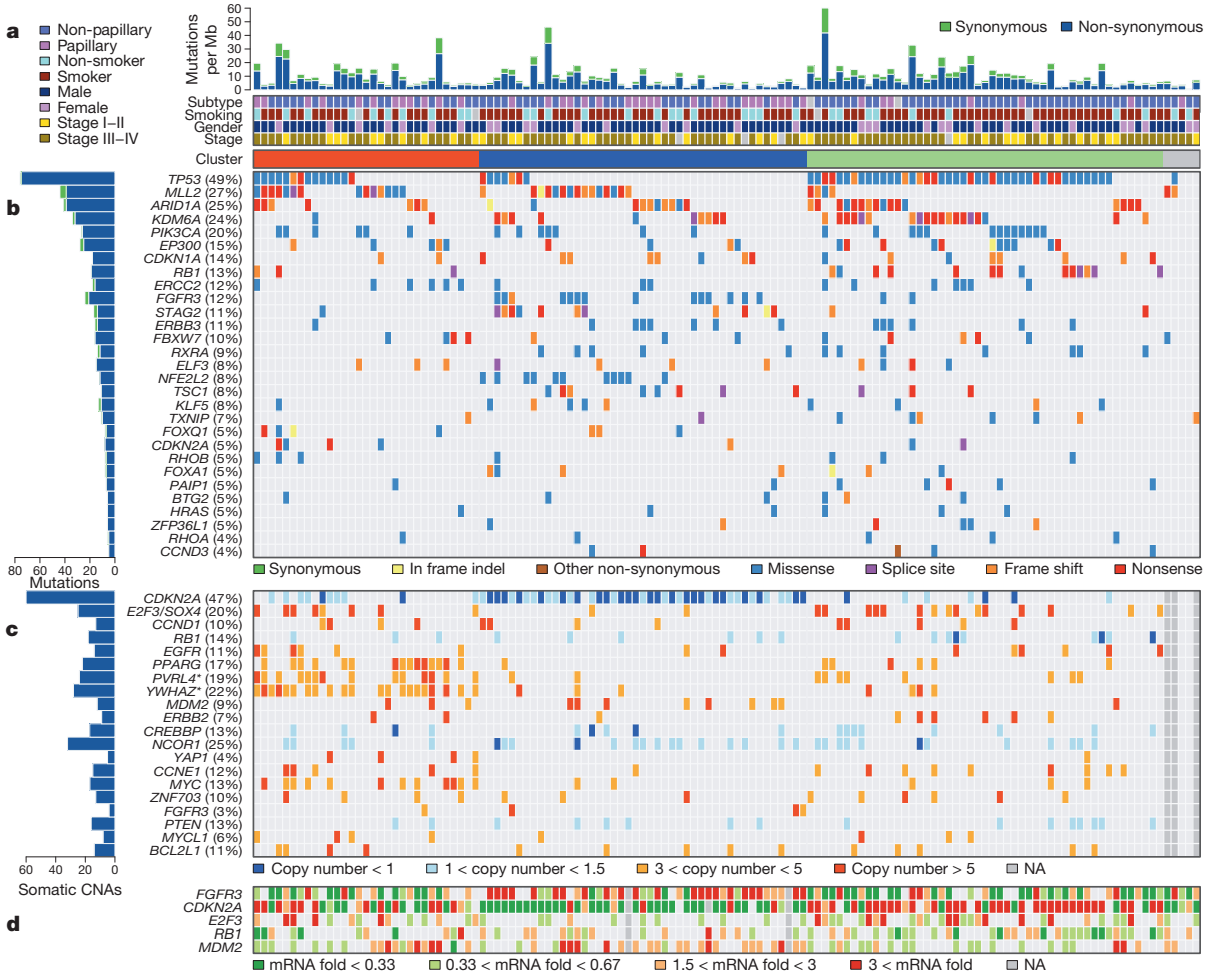

Figure 1 | The genomic landscape of bladder cancer. **a**, Mutation rate and type, histological subtype, smoking status, gender, tumour stage and cluster type. **b**, Genes with statistically significant levels of mutation (MutSig, false discovery rate <0.1) and mutation types. **c**, Deletions and amplifications for genomic regions with statistically significant focal copy number changes (GISTIC2.0). 'Copy number' refers to absolute copy number. Note that two amplification peaks (*) contain several genes, any of which could be the target, as opposed to the single gene listed here. **d**, RNA expression level for selected genes, expressed as fold change from the median value for all samples. Tumour samples were grouped into three clusters (red, blue and green) using consensus NMF clustering (see the main text and Supplementary Fig. 2.1.2). Three samples with no copy number data and two samples with no mutations in the genes were not used in the clustering and are shown in grey.

Methods 2.2). Three other genes identified by MutSig were not considered further because of low or undetectable expression (Supplementary Fig. 2.1.1). A similar analysis considering only mutations in the COSMIC database[2] identified three more significantly mutated genes: *ERBB2*, *ATM* and *CTNNB1* (Supplementary Table 2.1.3). We validated the mutation findings in three ways: targeted re-sequencing of all significantly mutated gene mutations, comparison with RNA-seq data for 123 samples and comparison with whole-genome sequence data for 18 samples. Overall, the validation rate was >99% in selected mutations by a combination of the methods (Supplementary Methods 2.4).

Nearly half (49%) of the samples had *TP53* mutations (Fig. 1b), which were mutually exclusive in their relationship with amplification (9%) and overexpression (29%) of *MDM2*; hence, TP53 function was inactivated in 76% of samples. Most *RB1* mutations were inactivating, were associated with significantly reduced mRNA level (Supplementary Fig. 2.7) and were mutually exclusive with *CDKN2A* deletions (Supplementary Fig. 2.8 and Supplementary Table 2.8.1). *FGFR3* mutations (12%) typically affected known kinase-activating sites. *PIK3CA* mutations were relatively common (20%), clustering in the helical domain near E545 (Supplementary Fig. 2.4). Most *TSC1* mutations (8%) were truncating, and six were homozygous (allele fraction >0.5).

Many of the 32 genes identified in Fig. 1b have not previously been reported as statistically significantly mutated in bladder cancer: *MLL2* (also called *KMT2D*; 27%), *CDKN1A** (14%), *ERCC2** (12%), *STAG2* (11%), *RXRA** (9%), *ELF3** (8%), *NFE2L2* (8%), *KLF5** (8%), *TXNIP* (7%), *FOXQ1** (5%), *RHOB** (5%), *FOXA1* (5%), *PAIP1** (5%), *BTG2** (5%), *ZFP36L1* (5%), *RHOA* (4%) and *CCND3* (4%). The nine genes marked with asterisks have not been reported as significantly mutated genes in any other TCGA cancer type or reported in another study as mutated at >3% frequency[2]. *CDKN1A* (*$p21^{CIP1}$*), a cyclin-dependent kinase inhibitor[12], had predominantly null or truncating mutations, indicating loss of function. Fifteen of sixteen mutations in *ERCC2*, a nucleotide excision repair gene[13], were deleterious missense mutations, suggesting dominant-negative effects. *ERCC2*-mutant tumours also had significantly fewer C>G mutations than did *ERCC2*-wild-type tumours (Supplementary Figs 2.3.1 and 2.3.2), and they trended towards higher overall mutation rate (Supplementary Fig. 2.12). Seven of twelve mutations in *RXRA* (retinoid X nuclear receptor alpha)[14] occurred at the same amino acid (five S427F; two S427Y) in the ligand-binding domain. Those seven tumours showed increased expression of genes involved in adipogenesis and lipid metabolism (Supplementary Fig. 2.6 and Supplementary Data 2.6.1–2.6.3), suggesting that the mutations cause constitutive activation.

Eleven tumours (8%) had deleterious missense mutations in the Neh2 domain of *NFE2L2*, a transcription factor that regulates the anti-oxidant program in response to oxidative stress[15]. Those tumours

showed markedly increased expression of genes involved in genotoxic metabolism and the reactive oxygen species (ROS) response (Supplementary Figs 2.5.1–2.5.3 and Supplementary Data 2.5.2). Furthermore, nine samples had mutations in redox regulator *TXNIP* (ref. 16) (five of them inactivating) and were mutually exclusive of samples with *NFE2L2* mutations, providing another mechanism for dysregulation of redox metabolism. Predominant inactivating mutations were seen in STAG2, an X-linked cohesin complex component required for separation of sister chromatids during cell division[17] (Supplementary Fig. 2.4).

Unsupervised clustering by non-negative matrix factorization of mutations and focal somatic CNAs in 125 samples identified three distinct groups (Fig. 1a and Supplementary Fig. 2.1.2). Group A (red), classified as 'focally amplified', is highly enriched in focal somatic CNAs in several genes, as well as mutations in *MLL2* (Fig. 1 and Supplementary Tables 2.1.4 and 2.1.5). Group B (blue), classified as 'papillary *CDKN2A*-deficient *FGFR3* mutant', is enriched in papillary histology. Nearly all group B samples show loss of *CDKN2A*, and most have one or more alterations in *FGFR3*. Group C (green), classified as '*TP53*/cell-cycle-mutant', shows *TP53* mutations in nearly all samples, as well as enrichment with *RB1* mutations and amplifications of *E2F3* and *CCNE1* (Fig. 1 and Supplementary Table 2.1.4). These differences in pattern of mutation suggest the possibility of different oncogenic mechanisms.

Seventy-two per cent of the cancers in this study were from current or past smokers, consistent with extensive epidemiological studies indicating an association between smoking and urothelial cancer risk. In contrast with lung cancer, however, there was no statistically significant association between smoking status and the mutational spectrum, frequency of mutation in any significantly mutated gene, occurrence of focal somatic CNAs or expression subtype (Supplementary Tables 2.9.1 and 2.9.2). Never-smokers did have a slightly higher fraction of C>G mutations than did current/former smokers (28.5% versus 23.8%, $P = 0.032$; Supplementary Figs 2.3.2 and 2.3.3). Unsupervised clustering of promoter CpG island DNA methylation data revealed a major subgroup (34%) of tumours (CIMP) characterized by cancer-specific DNA hypermethylation (Supplementary Fig. 7.1). Multivariate regression analysis with age, sex and tumour stage as covariates identified smoking pack-years as the only significant predictor of CIMP phenotype, as has also been reported for colorectal cancer[18].

Fifty-one per cent of mutations overall were Tp*C->(T/G) (Supplementary Table 2.1.1), a class of mutation recently reported to be mediated by one of the DNA cytosine deaminases, APOBEC (refs 19, 20). *APOBEC3B* was expressed at high levels in all of the tumours, suggesting a major role for APOBEC-mediated mutagenesis in bladder carcinogenesis (Supplementary Figs 12.1 and 12.2).

Four genes involved in epigenetic regulation were significantly mutated genes: *MLL2*, *ARID1A*, *KDM6A* and *EP300* (Fig. 1). Truncating mutations were significantly enriched in each of those genes (Supplementary Fig. 2.2 and Supplementary Data 2.2.1–2). Three of the genes had previously been identified as mutated in urothelial cancers[4], but mutation of *MLL2*, which encodes a histone H3 lysine 4 (H3K4) methyltransferase, is a novel finding. Several other chromatin-regulating genes had mutation rates ≥10% but were not statistically significant by MutSig analysis: *MLL3*, *MLL*, *CREBBP*, *CHD7* and *SRCAP*. Many other epigenetic regulators were mutated at lower frequency but were also enriched with truncating mutations, indicating functional significance (Supplementary Fig. 2.2 and Supplementary Data 2.2.1 and 2.2.2). Non-silent mutations in chromatin regulatory genes overall were significantly enriched in bladder cancer in comparison with the entire exome, in contrast with all other epithelial cancers studied so far in the TCGA project (Supplementary Table 2.10). Mutations in *MLL2* and *KDM6A* (the latter encoding a histone H3 lysine 27 (H3K27) demethylase) were mutually exclusive (Supplementary Fig. 2.8 and Supplementary Table 2.8.1), suggesting that mutations in the two genes have redundant downstream effects on carcinogenesis or that the combined loss is synthetically lethal.

Chromosomal rearrangements and viral integration

To identify structural variations and pathogen sequences, we used low-pass, paired-end, whole-genome sequencing (WGS; 6–8× coverage) of 114 tumours and RNA sequencing of all tumours. We detected 2,529 structural aberrations, including 1,153 that involve gene–gene fusions. Among the translocations, 379 were inter-chromosomal, 237 were intra-chromosomal, 274 were the result of inversions and 263 resulted from deletions (Supplementary Table 3.1). We found several recurrent translocations of probable pathogenic significance, including an intra-chromosomal translocation on chromosome 4 involving *FGFR3* and *TACC3* ($n = 3$). The breakpoints were in intron 16 (two cases) or exon 17 (one case) of *FGFR3* and intron 10 of *TACC3* (confirmed by DNA sequencing and RNA-seq). All three lead to fusion mRNA products for which the predicted proteins include the amino-terminal 758 amino acids of FGFR3 fused with the carboxy-terminal 191 amino acids of TACC3 (Fig. 2a). On the basis of the structure of the FGFR3–TACC3 fusion protein, we predict that it can auto-dimerize, leading to constitutive activation of the kinase domain of FGFR3. FGFR3–TACC3 fusion, which was recently described in both glioblastoma[21] and bladder cancer[7,8], represents a promising therapeutic target. The *ERBB2* gene was also involved in translocations in four tumours, all with different fusion partners and all confirmed by DNA sequencing, RNA-seq or both. In one case, exons 4 to 29 of *ERBB2* were fused to the promoter plus exon 1 of *DIP2B*, and the fusion product was amplified (Fig. 2b). Two other fusion products resulted in novel mRNA products, the biological significance of which is not known.

We identified viral DNAs in 7 of 122 tumours (6%), and viral transcripts in 5 of 122 (4%). Three tumours expressed cytomegalovirus (CMV) transcripts (encoding RL5A, RNA2.7, RL9A, RNA1.2, UL5 and UL22A), one expressed BK polyoma virus and one expressed human papilloma virus 16 (HPV16). HPV16 and human herpesvirus 6B DNA were each identified in one other sample but without expression. None of the tumours expressing CMV showed evidence of CMV integration into the host genome, suggesting the presence of a stable episome. In the BK-positive tumour, two BK genes were integrated into *GRB14*, a signalling adaptor protein for receptor tyrosine kinases. In the HPV-16-expressing case, the virus integrated into *BCL2L1*, an apoptosis-regulating gene (Fig. 2c). In that tumour, *BCL2L1* was amplified (~6×) and

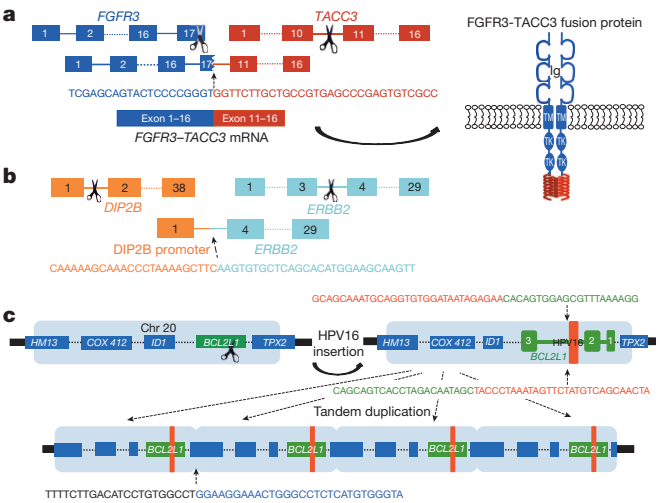

Figure 2 | Structural rearrangements and viral integration. **a**, FGFR3–TACC3 fusion in sample TCGA-CF-A3MH showing the breakpoints in the two genes, the breakpoint junction sequences and the predicted fusion protein. **b**, Rearrangement involving *DIP2B* and *ERBB2* in TCGA-DK-A2I6. The *ERBB2* gene has swapped its promoter with that of *DIP2B*, resulting in overexpression of *ERBB2*. **c**, Insertion of human papilloma virus 16 (HPV16) into the *BCL2L1* gene on chromosome 20 in TCGA-GC-A3I6. The region of *BCL2L1* into which the virus has integrated and the integration junction sequence are shown.

overexpressed (~10× median; >2× any of the other samples). Overall, these findings indicate that viral infection may have a role in the development of a small percentage of urothelial carcinomas.

mRNA, miRNA and protein expression

Analysis of RNA-seq data from 129 tumours identified four clusters (clusters I–IV) (Fig. 3 and Supplementary Fig. 4.1). Cluster I ('papillary-like') is enriched in tumours with papillary morphology ($P = 0.0002$), *FGFR3* mutations ($P = 0.0007$, $q = 0.02$), *FGFR3* copy number gain ($P = 0.04$, $q = 0.1$) and elevated FGFR3 expression ($P < 0.0001$) (Fig. 3a). It includes all three samples with *FGFR3–TACC3* fusions. Cluster I samples also show significantly lower expression of miR-99a and miR-100, miRNAs that downregulate *FGFR3* expression ($P = 0.0002$, Figs 3a and Supplementary Fig. 5.3)[22]. Cluster I samples also show lower expression of miR-145 and miR-125b, which have been reported as frequently downregulated in bladder cancer[23]. Tumours with *FGFR3* alterations, and perhaps other tumours that share the cluster I expression profile, may respond to inhibitors of FGFR or its downstream targets.

Reverse-phase protein array (RPPA) data indicate that clusters I and II express high HER2 (ERBB2) levels and an elevated oestrogen receptor beta (ESR2) signalling signature, indicating potential targets for hormone therapies such as tamoxifen or raloxifene (Fig. 3d). In fact, HER2 protein levels in a subset of the tumours are comparable to those found in TCGA HER2-positive breast cancers[23].

For comparison, we asked whether any of the four clusters show gene signatures similar to those identified in any other tumour type(s) among the first 11 analysed by TCGA. We found that the signature of bladder cancer cluster III ('basal/squamous-like') is similar to that of basal-like breast cancers, as well as squamous cell cancers of the head and neck and lung (Supplementary Fig. 4.2)[24,25]. All four of those cancer types express characteristic epithelial lineage genes, including *KRT14*, *KRT5*, *KRT6A* and *EGFR*. Basal-like subtype[26] and squamous cell subtype[27] of urothelial carcinoma have been independently reported. Many of the samples in bladder cluster III express cytokeratins (that is, *KRT14* and *KRT5*) that were recently reported to mark stem/progenitor cells[26]. Some of those samples also show a level of variant squamous histology (Fig. 3b). Bladder clusters I and II show features similar to those of luminal A breast cancer, with high mRNA and protein expression of luminal breast differentiation markers, including GATA3 and FOXA1 (Fig. 3c). Markers of urothelial differentiation such as the uroplakins (for example, UPK3A) are also highly expressed in clusters I and II, as are the epithelial marker E-cadherin and members of the miR-200 family of miRNAs (which target multiple regulators of epithelial–mesenchymal transition)[28] (Fig. 3c). Taken together, these observations indicate that, despite their diverse tissue origins, some bladder, breast, head and neck and lung cancers share common pathways of tumour development.

To determine whether the expression-based clusters could be seen in other data sets, we used the muscle-invasive bladder cancer samples from ref. 27, hierarchically clustering them with the genes used in our analysis. From the sample dendrogram, we identified four groups (Supplementary Fig. 4.3a). The four groups identified in the data set of ref. 27 correlated well with the four clusters identified in our TCGA data (Supplementary Fig. 4.3b).

When we analysed the RNA-seq data for transcript splice variation using SpliceSeq[29] (Supplementary Information, section 11), one finding of interest was an average of 3% *PKM1* and 97% *PKM2* transcripts in the tumour samples. The *PKM2* isoform of pyruvate kinase is the principal driver of a shift to aerobic glycolysis in tumours (the Warburg effect)[30]. Therefore, urothelial bladder cancers (and other cancer types) may prove sensitive to inhibition of glycolysis or related metabolic pathways.

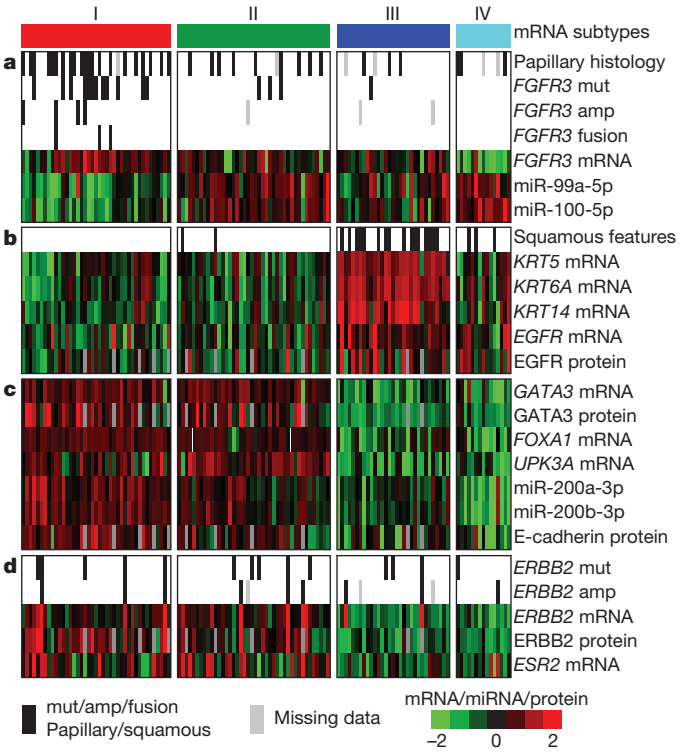

Figure 3 | Expression characteristics of bladder cancer. Integrated analysis of mRNA, miRNA and protein data led to identification of distinct subsets of urothelial carcinoma. Data for mRNA, miRNA and protein were *z*-normalized, and samples were organized in the horizontal direction by mRNA clustering. **a**, Papillary histology, *FGFR3* alterations, *FGFR3* expression and reduced *FGFR3*-related miRNA expression are enriched in cluster I. **b**, Expression of epithelial lineage genes and stem/progenitor cytokeratins are generally high in cluster III, some of which show variant squamous histology. **c**, Luminal breast and urothelial differentiation factors are enriched in clusters I and II. **d**, *ERBB2* mutation and oestrogen receptor beta (*ESR2*) expression are enriched in clusters I and II.

Pathway analysis and therapeutic targeting

Integrated analysis of the mutation and copy-number data revealed three main pathways as frequently dysregulated in bladder cancer: cell cycle regulation (altered in 93% of cases); kinase and phosphatidylinositol-3-OH kinase (PI(3)K) signalling (72%); and chromatin remodelling, including mutations/somatic CNAs in histone-modifying genes (89%) and components of the SWI/SNF nucleosome remodelling complex (64%) (Fig. 4a). To complement these results for well-defined pathways, we applied network analysis methods to examine other possible interactions between genes and pathways (Fig. 4b). In particular, we used the TieDIE algorithm to search for causal regulatory interactions within the PARADIGM network, which connects mutated genes to active transcriptional hubs[31,32]. The analysis identified a sub-network linking mutated histone-modifying genes to a large array of activated transcription factors, indicating potential far-reaching effects of histone modification on other pathways (Supplementary Fig. 8.2.1) converging on MYC/MAX regulation. Both MYC and MAX showed similar levels of pathway activity, independent of mutations in chromatin genes, suggesting that mutations in histone-modifying genes provide just one mechanism for disruption of the MYC/MAX hub. By contrast, tumours with chromatin-related mutations showed differential activity of transcription factors *FOXA2* and *SP1*, implicating de-differentiation processes as a result of the mutations. Our network analysis also identified *HSP90AA1* as a critical signalling hub, indicating that inhibitors of HSP90 may have therapeutic value in urothelial carcinoma. Although the linkages between mutations and transcriptional changes were statistically significant in terms of their proximity in the network (as determined by permutation tests; see Supplementary Fig. 8.2), further studies will be needed to assess the biological relevance of the findings.

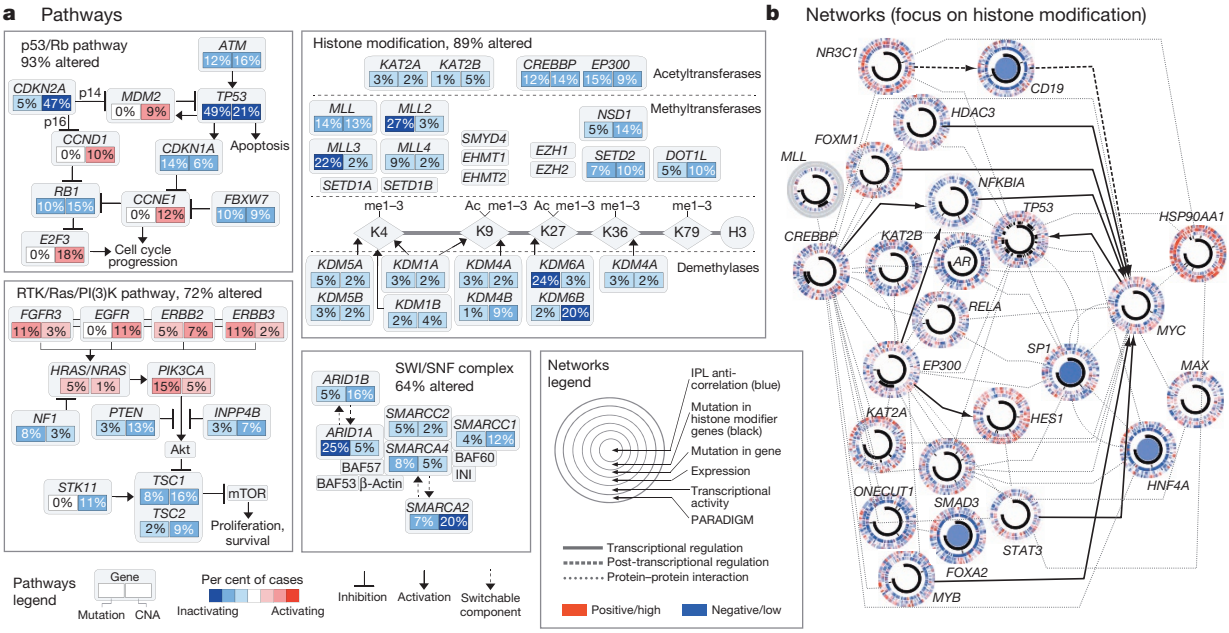

Figure 4 | Altered pathways and networks in bladder cancer. **a**, Somatic mutations and copy number alterations (CNA) in components of the p53/Rb pathway, RTK/RAS/PI(3)K pathway, histone modification system and SWI/SNF complex. Red, activating genetic alterations; blue, inactivating genetic alterations. Percentages shown denote activation or inactivation of at least one allele. **b**, The network connecting mutated histone-modifying genes to transcription factors with differential activity (methodology and larger implicated network in Supplementary Fig. 8.2.1). Each gene is depicted as a multi-ring circle with various levels of data, plotted such that each 'spoke' in the ring represents a single patient sample (same sample ordering for all genes). 'PARADIGM' ring, bioinformatically inferred levels of gene activity (red, higher activity); 'Transcriptional activity', mean mRNA levels of all of the targets of each transcription factor; 'expression', mRNA levels relative to normal (red, high); 'Mutation in gene', somatic mutation; 'Mutation in histone modifier genes', somatic mutation in at least one such gene; 'IPL anti-correlation', genes with PARADIGM integrated pathway levels (IPLs) inversely correlated with histone-gene mutation status. Gene–gene relationships are inferred using public resources.

Integrated analysis also identified mutations, copy number alterations or RNA expression changes affecting the PI(3)K/AKT/mTOR pathway in 42% of the tumours (Fig. 5a). Included were activating point mutations in *PIK3CA* (17%; potentially responsive to PI(3)K inhibitors), mutation or deletion of *TSC1* or *TSC2* (9%; potentially responsive to mTOR inhibitors) and overexpression of *AKT3* (10%; potentially responsive to AKT inhibitors). We also observed mutations, genomic amplifications or gene fusions that affect the RTK/RAS pathway in 44% of the tumours (Fig. 5b, c). Included were events that can activate FGFR3 (17%; potentially responsive to FGFR inhibitors or antibodies), amplification of EGFR (9%; potentially responsive to EGFR antibodies or inhibitors), mutations of *ERBB3* (6%; potentially

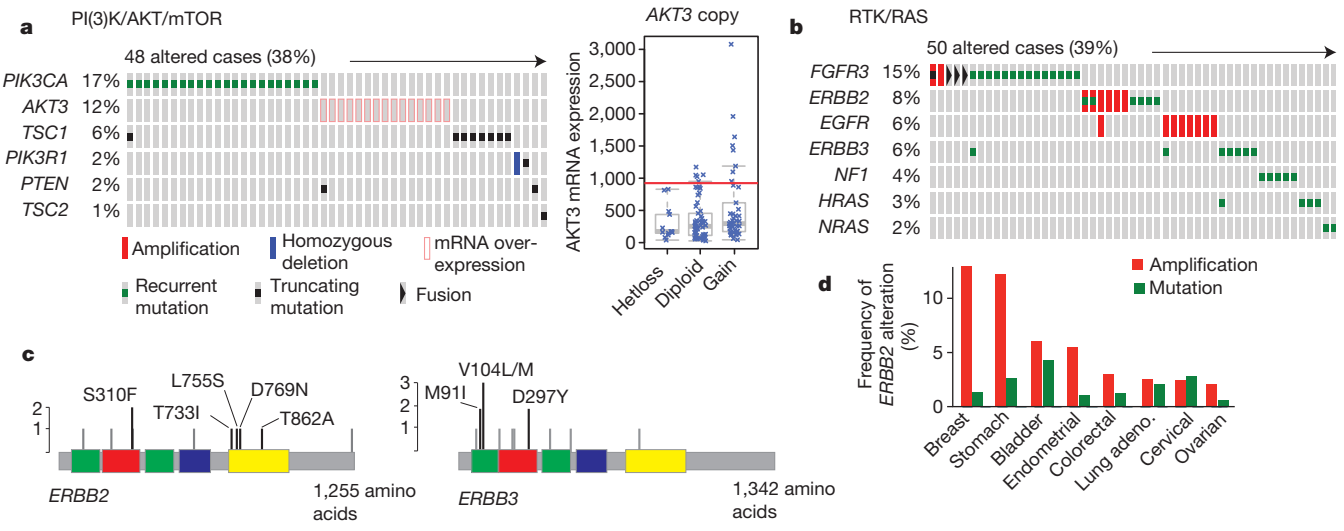

Figure 5 | Potential targets in bladder cancer. **a**, Alterations in the PI(3)K/AKT/mTOR pathway are mutually exclusive. Tumour samples are shown in columns; genes in rows. Only samples with at least one alteration are shown. AKT3 shows elevated expression in 10% of samples, independent of copy number (right panel). Hetloss, heterozygous loss. **b**, Receptor tyrosine kinases are altered, by any of several different mechanisms (amplification, mutation or fusion), in 45% of samples. Only mutations that are recurrent in this data set or previously reported in COSMIC are shown. **c**, Recurrent mutations in ERBB2 and ERBB3. The mutations shown in black are either recurrent in the TCGA data set or reported in COSMIC. Green, receptor L domain; red, furin-like cysteine-rich region; blue, growth factor receptor domain IV; yellow, tyrosine kinase domain. **d**, *ERBB2* amplifications and recurrent mutations in other cancers profiled by TCGA. Missense mutations were counted in the following positions: G309, S310, L313, R678, T733, L755, V777, D769, V842, T862, R896 and M916I. In-frame insertions were counted between amino acids 774 and 776. Only tumour types with an alteration frequency ≥2% are shown.

sensitive to ERBB kinase inhibitors) and mutation or amplification of *ERBB2* (9%; potentially sensitive to ERBB2 kinase inhibitors or antibodies). *ERBB3* mutations in bladder cancer have been noted previously[4], but statistically significant mutation of *ERBB2* in bladder cancer has not been reported. Both genes are potential therapeutic targets in other diseases[33–35]. Notably, *ERBB2* alterations were approximately as frequent in this study as in TCGA breast cancers, but with fewer amplifications and more mutations (Fig. 5d)[24].

Discussion

This integrated study of 131 invasive urothelial bladder carcinomas provides numerous novel insights into disease biology and delineates multiple potential opportunities for therapeutic intervention. Treatment for muscle-invasive bladder cancer has not advanced beyond cisplatin-based combination chemotherapy and surgery in the past 30 years[36], and no new drugs for the disease have been approved in that time. Median survival for patients with recurrent or metastatic bladder cancer remains 14–15 months with cisplatin-based chemotherapy, and there is no widely recognized second-line therapy[37]. With the exception of a single case report, there is also no known benefit from treatment with newer, targeted agents[38]. Several of the genomic alterations identified in this study, particularly those involving the PI(3)K/AKT/mTOR, CDKN2A/CDK4/CCND1 and RTK/RAS pathways, including ERBB2 (Her-2), ERBB3 and FGFR3, are amenable in principle to therapeutic targeting. Clinical trials based on patients with relevant druggable genomic alterations are warranted.

FGFR3 mutation is a common feature of low-grade non-invasive papillary urothelial bladder cancer, but it occurs at a much lower frequency in high-grade invasive bladder cancer. The cluster analysis in Fig. 3 highlights multiple mechanisms of *FGFR3* activation, and its strong association with papillary morphology. The data presented here suggest a subset of muscle-invasive cancers that can potentially be targeted through *FGFR3*. Similarly, ERBB2 amplification may be targetable by strategies used in breast cancer, by small-molecule tyrosine kinase inhibitors or by novel immunotherapeutic approaches (NCT01353222)[34]. The data here provide further support for several on-going ERBB2-targeted trials in bladder cancer and further define the subpopulation of cancers suited to that approach. Finally, cluster III of the integrated expression profiling analysis reveals the existence of a urothelial carcinoma subtype with cancer stem-cell expression features (including KRT14 and KRT5), perhaps providing another avenue for therapeutic targeting.

The alterations identified in epigenetic pathways also suggest new possibilities for bladder cancer treatment. Ninety-nine (76%) of the tumours analysed here had an inactivating mutation in one or more of the chromatin regulatory genes, and 53 (41%) had at least two such mutations. Overall, the bladder cancers showed a mutational spectrum highly enriched with mutations in chromatin regulatory genes (Supplementary Table 2.10). Furthermore, integrated network analyses revealed a profound impact of those mutations on the activity levels of various transcription factors and pathways implicated in cancer. Drugs that target chromatin modifications—for example, recently developed agents that bind acetyl-lysine binding motifs (bromodomains)—might prove useful for treatment of the subset of bladder tumours that exhibit abnormalities in chromatin-modifying enzymes[39]. Our findings overall indicate bladder cancer as a prime candidate for exploration of that approach to therapy.

METHODS SUMMARY

Tumour and normal samples were obtained with institutional-review-board-approved consent and processed using a modified AllPrep kit (Qiagen) to obtain purified DNA and RNA. Quality-control analyses revealed only modest batch effects (Supplementary Information, section 'Batch effects'). The tumours were profiled using Affymetrix SNP 6.0 microarrays for somatic CNAs, low-pass WGS (HiSeq) for somatic CNAs and translocations, RNA-seq (HiSeq) for mRNA and miRNA expression, Illumina Infinium (HumanMethylation450) arrays for DNA methylation, HiSeq for exome sequencing and RPPA for protein expression and phosphorylation. Statistical analysis and biological interpretation of the data were spearheaded by the TCGA genome data analysis centres. Sequence files are in CGHub (https://cghub.ucsc.edu/). All other molecular, clinical and pathological data are available through the TCGA Data Portal (https://tcga-data.nci.nih.gov/tcga/). The data can be explored through a compendium of next-generation clustered heat maps (http://bioinformatics.mdanderson.org/TCGA/NGCHMPortal/), the cBio Cancer Genomics Portal (http://cbioportal.org), TieDIE (http://sysbiowiki.soe.ucsc.edu/tiedie), SpliceSeq (http://bioinformatics.mdanderson.org/main/SpliceSeq:Overview), MBatch batch effects assessor (http://bioinformatics.mdanderson.org/tcgambatch/) and Regulome Explorer (http://explorer.cancerregulome.org/). Also see Supplementary Information.

Received 17 June; accepted 19 December 2013.
Published online 29 January; corrected online 19 March 2014 (see full-text HTML version for details).

1. Jemal, A. *et al.* Global cancer statistics. *CA Cancer J. Clin.* **61,** 69–90 (2011).
2. Forbes, S. A. *et al.* COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* **39,** D945–D950 (2011).
3. Goebell, P. J. & Knowles, M. A. Bladder cancer or bladder cancers? Genetically distinct malignant conditions of the urothelium. *Urol. Oncol.* **28,** 409–428 (2010).
4. Gui, Y. *et al.* Frequent mutations of chromatin remodeling genes in transitional cell carcinoma of the bladder. *Nature Genet.* **43,** 875–878 (2011).
5. Hurst, C. D., Platt, F. M., Taylor, C. F. & Knowles, M. A. Novel tumor subgroups of urothelial carcinoma of the bladder defined by integrated genomic analysis. *Clin. Cancer Res.* **18,** 5865–5877 (2012).
6. Lindgren, D. *et al.* Integrated genomic and gene expression profiling identifies two major genomic circuits in urothelial carcinoma. *PLoS ONE* **7,** e38863 (2012).
7. Williams, S. V., Hurst, C. D. & Knowles, M. A. Oncogenic FGFR3 gene fusions in bladder cancer. *Hum. Mol. Genet.* **22,** 795–803 (2013).
8. Wu, Y. M. *et al.* Identification of targetable FGFR gene fusions in diverse cancers. *Cancer Discov.* **3,** 636–647 (2013).
9. Lawrence, M. S. *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499,** 214–218 (2013).
10. Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nature Biotechnol.* **31,** 213–219 (2013).
11. Lawrence, M. S. *et al.* Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* **505,** 495–501 (2014).
12. Warfel, N. A. & El-Deiry, W. S. $p21^{WAF1}$ and tumourigenesis: 20 years after. *Curr. Opin. Oncol.* **25,** 52–58 (2013).
13. Lehmann, A. R. The xeroderma pigmentosum group D (XPD) gene: one gene, two functions, three diseases. *Genes Dev.* **15,** 15–23 (2001).
14. Tontonoz, P. *et al.* Adipocyte-specific transcription factor ARF6 is a heterodimeric complex of two nuclear hormone receptors, PPARγ and RXRα. *Nucleic Acids Res.* **22,** 5628–5634 (1994).
15. Shibata, T. *et al.* Cancer related mutations in NRF2 impair its recognition by Keap1-Cul3 E3 ligase and promote malignancy. *Proc. Natl Acad. Sci. USA* **105,** 13568–13573 (2008).
16. Zhou, J., Yu, Q. & Chng, W. J. TXNIP (VDUP-1, TBP-2): a major redox regulator commonly suppressed in cancer by epigenetic mechanisms. *Int. J. Biochem. Cell Biol.* **43,** 1668–1673 (2011).
17. Solomon, D. A. *et al.* Mutational inactivation of STAG2 causes aneuploidy in human cancer. *Science* **333,** 1039–1043 (2011).
18. Samowitz, W. S. *et al.* Association of smoking, CpG island methylator phenotype, and V600E BRAF mutations in colon cancer. *J. Natl. Cancer Inst.* **98,** 1731–1738 (2006).
19. Nik-Zainal, S. *et al.* Mutational processes molding the genomes of 21 breast cancers. *Cell* **149,** 979–993 (2012).
20. Roberts, S. A. *et al.* Clustered mutations in yeast and in human cancers can arise from damaged long single-strand DNA regions. *Mol. Cell* **46,** 424–435 (2012).
21. Singh, D. *et al.* Transforming fusions of FGFR and TACC genes in human glioblastoma. *Science* **337,** 1231–1235 (2012).
22. Oneyama, C. *et al.* MicroRNA-mediated downregulation of mTOR/FGFR3 controls tumor growth induced by Src-related oncogenic pathways. *Oncogene* **30,** 3489–3501 (2011).
23. Yoshino, H. *et al.* Aberrant expression of microRNAs in bladder cancer. *Nature Rev. Urol.* **10,** 396–404 (2013).
24. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490,** 61–70 (2012).
25. Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **489,** 519–525 (2012).
26. Ho, P. L., Kurtova, A. & Chan, K. S. Normal and neoplastic urothelial stem cells: getting to the root of the problem. *Nature Rev. Urol.* **9,** 583–594 (2012).
27. Sjodahl, G. *et al.* A molecular taxonomy for urothelial carcinoma. *Clin. Cancer Res.* **18,** 3377–3386 (2012).
28. Korpal, M., Lee, E. S., Hu, G. & Kang, Y. The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2. *J. Biol. Chem.* **283,** 14910–14914 (2008).
29. Ryan, M. C., Cleland, J., Kim, R., Wong, W. C. & Weinstein, J. N. SpliceSeq: a resource for analysis and visualization of RNA-Seq data on alternative splicing and its functional impacts. *Bioinformatics* **28,** 2385–2387 (2012).
30. Christofk, H. R. *et al.* The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. *Nature* **452,** 230–233 (2008).

31. Vaske, C. J. *et al.* Inference of patient-specific pathway activities from multi-dimensional cancer genomics data using PARADIGM. *Bioinformatics* **26,** i237–i245 (2010).
32. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* **499,** 43–49 (2013).
33. Bose, R. *et al.* Activating HER2 mutations in HER2 gene amplification negative breast cancer. *Cancer Discov.* **3,** 224–237 (2013).
34. Greulich, H. *et al.* Functional analysis of receptor tyrosine kinase mutations in lung cancer identifies oncogenic extracellular domain mutations of ERBB2. *Proc. Natl Acad. Sci. USA* **109,** 14476–14481 (2012).
35. Jaiswal, B. S. *et al.* Oncogenic ERBB3 mutations in human cancers. *Cancer Cell* **23,** 603–617 (2013).
36. National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in Oncology for Bladder Cancer. Vol. **1.2012,** http://www.nccn.org/professionals/physician_gls/f_guidelines.asp#site (2012).
37. von der Maase, H. *et al.* Long-term survival results of a randomized trial comparing gemcitabine plus cisplatin, with methotrexate, vinblastine, doxorubicin, plus cisplatin in patients with bladder cancer. *J. Clin. Oncol.* **23,** 4602–4608 (2005).
38. Iyer, G. *et al.* Genome sequencing identifies a basis for everolimus sensitivity. *Science* **338,** 221 (2012).
39. Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468,** 1067–1073 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements We are grateful to all of the patients and families who contributed to this study, as well as C. Gunter and L. Chastain for scientific editing and M. Sheth, J. Zhang and C. Ron Bouchard for administrative support. This work was supported by the following grants from the United States National Institutes of Health: U54 HG003273, U54 HG003067, U54 HG003079, U24 CA143799, U24 CA143835, U24 CA143840, U24 CA143843, U24 CA143845, U24 CA143848, U24 CA143858, U24 CA143866, U24 CA143867, U24 CA143882, U24 CA143883, U24 CA144025 and P01 CA120964. Additional personnel and funding sources are acknowledged in the Supplementary Information.

Author Information The primary and processed data used to generate the analyses presented here can be downloaded by registered users from The Cancer Genome Atlas at https://tcga-data.nci.nih.gov/tcga/tcgaDownload.jsp. All of the primary sequence files are deposited in CGHub and all other data are deposited at the Data Coordinating Center (DCC) for public access (http://cancergenome.nih.gov/, https://cghub.ucsc.edu/ and https://tcga-data.nci.nih.gov/docs/publications/blca_2013/). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.N.W. (jweinste@mdanderson.org), S.P.L. (slerner@bcm.edu) or D.J.K. (dk@rics.bwh.harvard.edu).

Author Contributions The Cancer Genome Atlas research network contributed collectively to this study. Biospecimens were provided by the tissue source sites and processed by the Biospecimen Core Resource. Data generation and analyses were performed by the genome-sequencing centres, cancer genome-characterization centres and genome data analysis centres. All data were released through the Data Coordinating Center. Project activities were coordinated by the NCI and NHGRI project teams. We also acknowledge the following TCGA investigators of the Bladder Analysis Working Group who contributed substantially to the project. Project leaders: J. N. Weinstein and S. P. Lerner. Data coordinator: C. J. Creighton. Analysis coordinators: R. Akbani and J. Kim. Manuscript coordinator: M. B. Morgan. Project coordinator: M. Sheth. Writing team: J. N. Weinstein, D. J. Kwiatkowski, S. P. Lerner, C. J. Creighton, P. W. Laird, R. Kucherlapati, R. Akbani, X. Su, K. A. Hoadley and M. C. Ryan. Clinical expertise: S. Lerner, D. J. Kwiatkowski, J. E. Rosenberg and D. Bajorin. Pathology review: H. Al-Ahmadie, B. A. Czerniak, D. Hansel, V. Reuter and B. Robinson. DNA sequence and copy number analysis: J. Kim, D. J. Kwiatkowski, A. D. Cherniack and J. E. Rosenberg. DNA methylation analysis: P. W. Laird and T. Hinoue. mRNA analysis: K. A. Hoadley, W. Y. Kim, J. S. Damrauer, W. Zhang, Y. Liu and R. Akbani. miRNA analysis: G. Robertson and A. J. Mungall. Transcript splicing analysis: M. Ryan and J. N. Weinstein. Protein analysis: R. Akbani and G. B. Mills. APOBEC: D. A. Gordenin. Pathway/integrated analysis: C. J. Creighton, N. Schultz, Evan O. Paull and J. Stuart. Chromosomal rearrangements and viral integration: X. Su, R. Kucherlapati, N. Santoso, S. Lee and M. Parfenov. Batch effects: R. Akbani and J. N. Weinstein. Manuscript review: R. Gibbs, C. Gunter and M. Meyerson. Contact PIs: J. N. Weinstein, S. P. Lerner and D. J. Kwiatkowski.

The Cancer Genome Atlas Research Network

Analysis working group: The University of Texas MD Anderson Cancer Center John N. Weinstein[1,2], Rehan Akbani[1], Bradley M. Broom[1], Wenyi Wang[1], Roeland G. W. Verhaak[1], David McConkey[3]; **Baylor College of Medicine** Seth Lerner[4,5], Margaret Morgan[5,6], Chad J. Creighton[7], Carolyn Smith[8]; **Broad Institute** David J. Kwiatkowski[9,10,11], Andrew D. Cherniack[9], Jaegil Kim[9], Chandra Sekhar Pedamallu[9,12], Michael S. Noble[9]; **Memorial Sloan-Kettering Cancer Center** Hikmat A. Al-Ahmadie[13], Victor E. Reuter[13], Jonathan E. Rosenberg[13], Dean F. Bajorin[13], Bernard H. Bochner[13], David B. Solit[13]; **Oregon Health and Science University, Department of Urology** Theresa Koppie[14]; **Weill Medical College of Cornell University** Brian Robinson[15]; **National Institute of Environmental Health Sciences** Dmitry A. Gordenin[16], David Fargo[16], Leszek J. Klimczak[16], Steven A. Roberts[16]; **Optimum Therapeutics LLC** Jessie Au[17]; **University of Southern California Epigenome Center** Peter W. Laird[18], Toshinori Hinoue[18]; **Computational Biology Center, Memorial Sloan-Kettering Cancer Center** Nikolaus Schultz[19], Ricardo Ramirez[19]; **UCSD Department of Pathology** Donna Hansel[20]; **Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill** Katherine A. Hoadley[21], William Y. Kim[21,22,23]; **Department of Genetics, University of North Carolina at Chapel Hill** Jeffrey S. Damrauer[21,22]; **The Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins University** Stephen B. Baylin[24]; **Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency** Andrew J. Mungall[30], A. Gordon Robertson[30], Andy Chu[30].

Genome Sequencing Center: Broad Institute David J. Kwiatkowski[9,10,11], Carrie Sougnez[9], Kristian Cibulskis[9], Lee Lichtenstein[9], Andrey Sivachenko[9], Chip Stewart[9], Michael S. Lawrence[9], Gad Getz[9,25], Eric Lander[9], Stacey B. Gabriel[9].

Genome characterization centres: Dan L. Duncan Cancer Center, Human Genome Sequencing Center, Baylor College of Medicine Chad J. Creighton[7], Lawrence Donehower[7,26]; **Broad Institute** Andrew D. Cherniack[9], Jaegil Kim[9], Scott L. Carter[9], Gordon Saksena[9], Steven E. Schumacher[9,27], Carrie Sougnez[9], Samuel S. Freeman[9], Joonil Jung[9], Chandra Sekhar Pedamallu[9,12], Ami S. Bhatt[9,12], Trevor Pugh[9,12], Gad Getz[9,25], Rameen Beroukhim[9,12,28], Stacey B. Gabriel[9], Matthew Meyerson[9,12,29]; **Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency** Andrew J. Mungall[30], A. Gordon Robertson[30], Andy Chu[30], Adrian Ally[30], Miruna Balasundaram[30], Yaron S. N. Butterfield[30], Noreen Dhalla[30], Carrie Hirst[30], Robert A. Holt[30], Steven J. M. Jones[30], Darlene Lee[30], Haiyan I. Li[30], Marco A. Marra[30], Michael Mayo[30], Richard A. Moore[30], Jacqueline E. Schein[30], Payal Sipahimalani[30], Angela Tam[30], Nina Thiessen[30], Tina Wong[30], Natasja Wye[30], Reanne Bowlby[30], Eric Chuah[30], Ranabir Guin[30], Steven J. M. Jones[30], Marco A. Marra[30]; **University of Southern California Epigenome Center** Toshinori Hinoue[18], Hui Shen[18], Moiz S. Bootwalla[18], Timothy Triche Jr[18], Phillip H. Lai[18], David J. Van Den Berg[18], Daniel J. Weisenberger[18], Peter W. Laird[18]; **UCSD Department of Pathology** Donna Hansel[20]; **Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill** Katherine A. Hoadley[21], Saianand Balu[21], Tom Bodenheimer[21], Jeffrey S. Damrauer[21,22] Alan P. Hoyle[21], Stuart R. Jefferys[21], Shaowu Meng[21], Lisle E. Mose[21], Janae V. Simons[21], Mathew G. Soloway[21], Junyuan Wu[21], William Y. Kim[21,22,23], Joel S. Parker[21,22], D. Neil Hayes[21,31]; **Research Computing Center, University of North Carolina at Chapel Hill** Jeffrey Roach[32]; **Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill** Elizabeth Buda[33]; **Department of Biology, University of North Carolina at Chapel Hill** Corbin D. Jones[33,34], Piotr A. Mieczkowski[34], Donghui Tan[34], Umadevi Veluvolu[34], Scot Waring[34]; **Eshelman School of Pharmacy, University of North Carolina at Chapel Hill** J. Todd Auman[35]; **Department of Genetics, University of North Carolina at Chapel Hill** Charles M. Perou[22], Matthew D. Wilkerson[22]; **Department of Genetics, Harvard Medical School** Netty Santoso[36], Michael Parfenov[36], Xiaojia Ren[36], Angeliki Pantazi[36], Angela Hadjipanayis[36,37], Jonathan Seidman[36], Raju Kucherlapati[36,37]; **The Center for Biomedical Informatics, Harvard Medical School** Semin Lee[38], Lixing Yang[38], Peter J. Park[37,38,39]; **Cancer Biology Division, The Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins University** Stephen B. Baylin[24]; **Division of Genetics, Brigham and Women's Hospital** Andrew Wei Xu[37]; **Institute for Applied Cancer Science, Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center** Alexei Protopopov[40], Jianhua Zhang[40], Christopher Bristow[40], Harshad S. Mahadeshwar[40], Sahil Seth[40], Xingzhi Song[40], Jiabin Tang[40], Dong Zeng[40], Lynda Chin[9,40]; **The University of Texas MD Anderson Cancer Center, Department of Pathology** Charles Guo[41]

Genome data analysis centres: The University of Texas M.D. Anderson Cancer Center John N. Weinstein[1,2], Rehan Akbani[1], Bradley M. Broom[1], David McConkey[3], Tod D. Casasent[10], Wenbin Liu[1,2], Zhenlin Ju[1,2], Thomas Motter[1], Bo Peng[1], Michael Ryan[1], Wenyi Wang[1], Roeland G. W. Verhaak[1], Xiaoping Su[1], Ji-Yeon Yang[1,2], Philip L. Lorenzi[1], Hui Yao[1], Nianxiang Zhang[1], Jiexin Zhang[1], Gordon B. Mills[2]; **Broad Institute** Jaegil Kim[9], Michael S. Noble[9], Juok Cho[9], Daniel DiCara[9], Scott Frazer[9], Nils Gehlenborg[9], David I. Heiman[9], Pei Lin[9], Yingchun Liu[9], Petar Stojanov[9,12], Doug Voet[9], Hailei Zhang[9], Lihua Zou[9], Lynda Chin[9,40], Gad Getz[9,25]; **Institute for Systems Biology** Brady Bernard[42], Dick Kreisberg[42], Sheila Reynolds[42], Hector Rovira[42], Ilya Shmulevich[42]; **Computational Biology Center, Memorial Sloan-Kettering Cancer Center** Ricardo Ramirez[19], Nikolaus Schultz[19], Jianjiong Gao[19], Anders Jacobsen[19], B. Arman Aksoy[19], Yevgeniy Antipin[19], Giovanni Ciriello[19], Gideon Dresdner[19], Benjamin Gross[19], William Lee[19], Boris Reva[19], Ronglai Shen[19], Rileen Sinha[19], S. Onur Sumer[19], Nils Weinhold[19], Marc Ladanyi[19], Chris Sander[19]; **Buck Institute for Research on Aging** Christopher Benz[43]; **University of California Santa Cruz** Daniel Carlin[44], David Haussler[44], Sam Ng[44], Evan O. Paull[44], Joshua Stuart[44], Jing Zhu[44]; **Department of Pathology, MD Anderson Cancer Center** Yuexin Liu[45], Wei Zhang[45]; **Helen Diller Family Comprehensive Cancer Center, University of California** Barry S. Taylor[46]

Biospecimen core resource: The Research Institute at Nationwide Children's Hospital Tara M. Lichtenberg[47], Erik Zmuda[47], Thomas Barr[47], Aaron D. Black[47], Myra George[47], Benjamin Hanf[47], Carmen Helsel[47], Cynthia McAllister[47], Nilsa C. Ramirez[47,48], Teresa R. Tabler[47], Stephanie Weaver[47], Lisa Wise[47], Jay Bowen[47], Julie M. Gastier-Foster[47,48]

Tissue source sites: The University of Texas MD Anderson Cancer Center John N. Weinstein[1,2]; **Scott Department of Urology, Baylor College of Medicine** Seth Lerner[4,5], Weiguo Jian[4,5], Sebrina Tello[4,5]; **Texas Cancer Research Biobank (TCRB), Baylor College of Medicine** Michael Ittman[5,49], Patricia Castro[5,49], Whitney D. McClenden[5], Margaret Morgan[5,6], Richard Gibbs[5,6]; **Broad Institute** Yingchun Liu[9]; **Analytical Biological Services, Inc.** Charles Saller[50], Katherine Tarvin[50]; **Cleveland Clinic Foundation** Jennifer M. DiPiero[51], Jennifer Owens[51]; **Georgia Regents University Cancer Center** Roni Bollag[52], Qiang Li[52], Paul Weinberger[52]; **Helen F. Graham Cancer Center at Christiana Care** Christine Czerwinski[53], Lori Huelsenbeck-Dill[53], Mary Iacocca[53], Nicholas Petrelli[53], Brenda Rabeno[53], Pat Swanson[53]; **International Genomics Consortium** Troy Shelton[54], Erin Curley[54], Johanna Gardner[54], David Mallery[54], Robert Penny[54]; **ILSbio, LLC** Nguyen Van Bang[55,56], Phan Thi Hanh[55,56], Bernard Kohl[55], Xuan Van Le[55], Bui Duc Phu[55,56], Richard Thorp[55], Nguyen Viet Tien[55,56], Le Quang Vinh[55,56]; **IU School of Medicine** George Sandusky[57]; **Lahey Hospital and Medical Center** Eric Burks[58], Kimberly Christ[58], Jason Gee[58], Antonia Holway[58], Alireza Moinzadeh[58], Andrea Sorcini[58], Travis Sullivan[58]; **Memorial Sloan-Kettering Cancer Center** Hikmat A. Al-Ahmadie[13], Dean F. Bajorin[13], Bernard H. Bochner[13], Ilana R. Garcia-Grossman[13], Ashley M. Regazzi[13], David B. Solit[13], Jonathan E. Rosenberg[13], Victor E. Reuter[13]; **Oregon Health and Science University, Department of Urology** Theresa Koppie[14]; **University of North Carolina, Lineberger Cancer Center** Lori Boice[59], Wendy Kimryn Rathmell[59], Leigh Thorne[59]; **University of Pittsburgh** Sheldon Bastacky[60], Benjamin Davies[60], Rajiv Dhir[60], Jeffrey Gingrich[60], Ronald Hrebinko[60], Jodi Maranchie[60], Joel Nelson[60], Anil Parwani[60]; **Roswell Park Cancer Institute** Wiam Bshara[61], Carmelo Gaudioso[61], Carl Morrison[61]; **Ontario Tumour Bank—Hamilton site, St Joseph's Healthcare Hamilton** Vina Alexopoulou[62], John Bartlett[62], Jay Engel[62], Sugy Kodeeswaran[62]; **The University of Chicago** Tatjana Antic[63], Peter H. O'Donnell[63], Norm D. Smith[63], Gary D. Steinberg[63]; **University of Miami, Sylvester Comprehensive Cancer Center** Sophie Egea[64], Carmen Gomez-Fernandez[64], Lynn Herbert[64], Merce Jorda[64], Mark Soloway[64]; **UT Southwestern Medical Center** Allison Beaver[65], Suzie Carter[65], Payal Kapur[65], Cheryl Lewis[65], Yair Lotan[65]; **Weill Medical College of Cornell University** Brian Robinson[15]; **UCSD Department of Pathology** Donna Hansel[20]; **The University of Texas MD Anderson Cancer Center, Department of Pathology** Charles Guo[41], Jolanta Bondaruk[41], Bogdan Czerniak[41]

Disease working group: The University of Texas MD Anderson Cancer Center Rehan Akbani[1], Bradley M. Broom[1], Yuexin Liu[45], Wei Zhang[45], John N. Weinstein[1,2]; **Scott Department of Urology, Baylor College of Medicine** Seth Lerner[4,5]; **Baylor College of Medicine** Margaret Morgan[5,6]; **Broad Institute** Jaegil Kim[9], Andrew D. Cherniack[9], Samuel S. Freeman[9], Chandra Sekhar Pedamallu[9,12], Michael S. Noble[9], David J. Kwiatkowski[9,10,11]; **Memorial Sloan-Kettering Cancer Center** Hikmat A. Al-Ahmadie[13], Dean F. Bajorin[13], Bernard H. Bochner[13], David B. Solit[13], Jonathan E. Rosenberg[13], Victor E. Reuter[13]; **Oregon Health and Science University, Department of Urology** Theresa Koppie[14]; **Weill Medical College of Cornell University** Brian Robinson[15]; **Stanford University, Department of Urology** Eila Skinner[66]; **Computational Biology Center, Memorial Sloan-Kettering Cancer Center** Ricardo Ramirez[19], Nikolaus Schultz[19]; **UCSD Department of Pathology** Donna Hansel[20]; **Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill** William Y. Kim[21,22,23]; **The University of Texas MD Anderson Cancer Center, Department of Pathology** Charles Guo[41], Jolanta Bondaruk[41], Kenneth Aldape[41], Bogdan Czerniak[41]

Data coordination centre: SRA International Mark A. Jensen[67], Ari B. Kahn[67], Todd D. Pihl[67], David A. Pot[67], Deepak Srinivasan[67], Yunhu Wan[67]

Project team: MLF Consulting Martin L. Ferguson[68]; **National Cancer Institute** Jean Claude Zenklusen[69], Tanja Davidsen[69], John A. Demchok[69], Kenna R. Mills Shaw[3,69], Margi Sheth[69], Roy Tarnuzzer[69], Zhining Wang[69], Liming Yang[69]; **National Human Genome Research Institute** Carolyn Hutter[70], Bradley A. Ozenberger[70], Heidi J. Sofia[70]; **Scimentis, LLC** Greg Eley[71]

[1]Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. [2]Department of Systems Biology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. [3]The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. [4]Scott Department of Urology, Baylor College of Medicine, Houston, Texas 77030, USA. [5]Texas Cancer Research Biobank (TCRB), Baylor College of Medicine, Houston, Texas 77030, USA. [6]Human Genome Sequencing Center at Baylor College of Medicine, Houston, Texas 77030, USA. [7]Dan L. Duncan Cancer Center, Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas 77030, USA. [8]Baylor College of Medicine, Houston, Texas 77030, USA. [9]The Eli and Edythe L. Broad Institute of Massachusetts Institute of Technology and Harvard University Cambridge, Massachusetts 02142, USA. [10]Brigham and Women's Hospital, 75 Francis St, Boston, Massachusetts 02115, USA. [11]Harvard Medical School, Boston, Massachusetts 02115, USA. [12]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts 02215, USA. [13]Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. [14]Oregon Health and Science University, Department of Urology, 3303 SW Bond Avenue, CHH10U, Portland, Oregon 97239, USA. [15]Weill Medical College of Cornell University, New York, New York 10065, USA. [16]National Institute of Environmental Health Sciences, 111 T.W. Alexander Drive, Research Triangle Park, North Carolina 27709, USA. [17]Optimum Therapeutics LLC, 9363 Towne Centre Drive, San Diego, California 92121, USA. [18]University of Southern California Epigenome Center, University of Southern California, Los Angeles, California 90033, USA. [19]Computational Biology Center, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10065, USA. [20]UCSD Department of Pathology 9500 Gilman Drive, La Jolla, California 92093, USA. [21]Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [22]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [23]Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [24]Cancer Biology Division, The Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins University, Baltimore, Maryland 21231, USA. [25]Massachusetts General Hospital, Cancer Center and Department of Pathology, 55 Fruit Street, Boston, Massachusetts 02114, USA. [26]Department of Molecular Virology and Microbiology, Baylor College of Medicine, 1 Baylor Plaza, Houston, Texas 77030, USA. [27]Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, Massachusetts 02215, USA. [28]Department of Medicine, Harvard Medical School, Boston, Massachusetts 02215, USA. [29]Department of Pathology, Harvard Medical School, Boston, Massachusetts 02215, USA. [30]Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency, Vancouver, British Columbia V5Z 4S6, Canada. [31]Department of Internal Medicine, Division of Medical Oncology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [32]Research Computing Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [33]Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [34]Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [35]Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [36]Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA. [37]Division of Genetics, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. [38]The Center for Biomedical Informatics, Harvard Medical School, Boston, Massachusetts 02115, USA. [39]Informatics Program, Children's Hospital, Boston, Massachusetts 02115, USA. [40]Institute for Applied Cancer Science, Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. [41]The University of Texas MD Anderson Cancer Center, Department of Pathology, Unit 085; 1515 Holcombe Boulevard, Houston, Texas 77030, USA. [42]Institute for Systems Biology, 401 Terry Ave N, Seattle, Washington 98109, USA. [43]Buck Institute for Research on Aging; 8001 Redwood Blvd, Novato, California 94945, USA. [44]University California Santa Cruz, 1156 High Street, Santa Cruz, California 95064, USA. [45]Department of Pathology, MD Anderson Cancer Center, Houston, Texas 77030, USA. [46]Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, California 94158, USA. [47]The Research Institute at Nationwide Children's Hospital, Columbus, Ohio 43205, USA. [48]The Ohio State University, Columbus, Ohio 43210, USA. [49]Department of Pathology, Baylor College of Medicine, Houston, Texas 77030, USA. [50]Analytical Biological Services, Inc., 701 Cornell Drive, Wilmington, Delaware 19801, USA. [51]Cleveland Clinic Foundation, 9500 Euclid Avenue, Cleveland, Ohio 44195, USA. [52]Georgia Regents University Cancer Center. Augusta, Georgia 30912, USA. [53]Helen F. Graham Cancer Center at Christiana Care, 4701 Ogletown Stanton Road, Newark, Delaware 19713, USA. [54]International Genomics Consortium, 445 N. Fifth Street, Phoenix, Arizona 85004, USA. [55]ILSbio, LLC 100 Radcliffe Drive, Chestertown, Maryland 21620, USA. [56]Hue Central Hospital, Hue City, Vietnam. [57]IU School of Medicine, Med Science Bldg, Room 128A, 635 Barnhill Drive, Indianapolis, Indiana 46202, USA. [58]Lahey Hospital and Medical Center, Burlington, Massachusetts 01805, USA. [59]University of North Carolina, Lineberger Cancer Center, 450 West Drive, Chapel Hill, North Carolina 27599, USA. [60]University of Pittsburgh, Pittsburgh, Pennsylvania 15213, USA. [61]Roswell Park Cancer Institute, Elm and Carlton Streets, Buffalo, New York 14063, USA. [62]Ontario Tumour Bank—Hamilton site, St Joseph's Healthcare Hamilton, Hamilton, Ontario L8N 3Z5, Canada. [63]The University of Chicago, Chicago, Illinois 60637, USA. [64]University of Miami, Sylvester Comprehensive Cancer Center, 1550 NW 10th Avenue, Miami, Florida 33136, USA. [65]UT Southwestern Medical Center 5323 Harry Hines Blvd, Dallas, Texas 75390-9110, USA. [66]Stanford University, Department of Urology, 300 Pasteur Drive, Suite S287, Stanford, California 94305, USA. [67]SRA International, Fairfax, Virginia 22033, USA. [68]MLF Consulting, Arlington, Massachusetts 02474, USA. [69]National Cancer Institute, 31 Center Drive, 3A20, Bethesda, Maryland 20892, USA. [70]National Human Genome Research Institute, 5635 Fishers Lane, Rockville, Maryland 20852, USA. [71]Scimentis, LLC, Atlanta, Georgia 30666, USA.