

# Comprehensive molecular characterization of clear cell renal cell carcinoma

The Cancer Genome Atlas Research Network\*

**Genetic changes underlying clear cell renal cell carcinoma (ccRCC) include alterations in genes controlling cellular oxygen sensing (for example, *VHL*) and the maintenance of chromatin states (for example, *PBRM1*). We surveyed more than 400 tumours using different genomic platforms and identified 19 significantly mutated genes. The PI(3)K/AKT pathway was recurrently mutated, suggesting this pathway as a potential therapeutic target. Widespread DNA hypomethylation was associated with mutation of the H3K36 methyltransferase *SETD2*, and integrative analysis suggested that mutations involving the SWI/SNF chromatin remodelling complex (*PBRM1*, *ARID1A*, *SMARCA4*) could have far-reaching effects on other pathways. Aggressive cancers demonstrated evidence of a metabolic shift, involving downregulation of genes involved in the TCA cycle, decreased AMPK and PTEN protein levels, upregulation of the pentose phosphate pathway and the glutamine transporter genes, increased acetyl-CoA carboxylase protein, and altered promoter methylation of miR-21 (also known as *MIR21*) and *GRB10*. Remodelling cellular metabolism thus constitutes a recurrent pattern in ccRCC that correlates with tumour stage and severity and offers new views on the opportunities for disease treatment.**

Kidney cancers, or renal cell carcinomas (RCC), are a common group of chemotherapy-resistant diseases that can be distinguished by histopathological features and underlying gene mutations<sup>1</sup>. Inherited predisposition to RCC has been shown to arise from genes involved in regulating cellular metabolism, making RCC a model for the role of an oncologic-metabolic shift, commonly referred to as the ‘Warburg effect’, leading to malignancy<sup>2</sup>. The most common type of RCC, clear cell renal cell carcinoma (ccRCC), is closely associated with *VHL* gene mutations that lead to stabilization of hypoxia inducible factors (*HIF-1 $\alpha$*  and *HIF-2 $\alpha$* , also known as *HIF1A* and *EPAS1*) in both sporadic and familial forms. *PBRM1*, a subunit of the PBAF SWI/SNF chromatin remodelling complex, as well as histone deubiquitinase *BAP1* and histone methyltransferase *SETD2*, were recently found to be altered in ccRCC<sup>3–5</sup>, implicating major roles for epigenetic regulation of additional functional pathways participating in the development and progression of the disease. Oncogenic metabolism and epigenetic reprogramming have thus emerged as central features of ccRCC.

In the present study, clinical and pathological features, genomic alterations, DNA methylation profiles, and RNA and proteomic signatures were evaluated in ccRCC. We accrued more than 500 primary nephrectomy specimens from patients with histologically confirmed ccRCC that conformed to the requirements for genomic study defined by the Cancer Genome Atlas (TCGA), together with matching ‘normal’ genomic material. Samples were restricted to those that contained at least 60% tumour nuclei (median 85%) by pathological review (clinical data summary provided in Supplementary Table 1). A data freeze representing 446 samples was generated from at least one analytical platform (‘Extended’ data set) and data from all platforms were available for 372 samples for coordinated, integrative analyses (‘Core’ data set) (Supplementary Data 1, Supplementary Table 2). No substantial batch effects in the data that might confound analyses were detected (Supplementary Figs 1–20).

## Somatic alterations

The global pattern of somatic alterations, determined from analysis of 417 samples, is shown in Fig. 1a. DNA hybridizations showed that recurrent arm-level and focal somatic copy number alterations

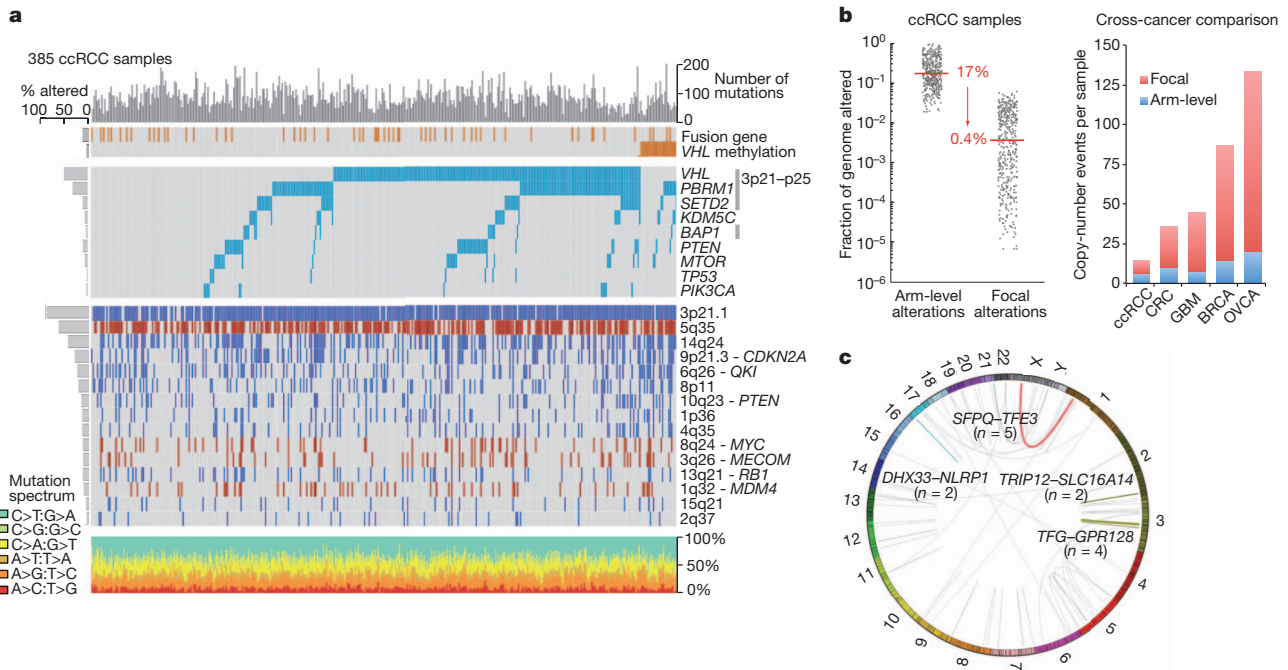
(SCNAs) occurred at a fewer sites than is generally observed in other cancers ( $P < 0.0004$ ; Supplementary Figs 21–22 and Supplementary Table 3). However, SCNAs that were observed more commonly involved entire chromosomes or chromosome arms, rather than focal events (17% vs 0.4%, Fig. 1b). Notably, the most frequent arm-level events involved loss of chromosome 3p (ref. 6; 91% of samples), encompassing all of the four most commonly mutated genes (*VHL*, *PBRM1*, *BAP1* and *SETD2*).

The data also suggested lower and more variable tumour cellularity<sup>7</sup> in the accrued samples, compared to conventional pathological review (median 54%  $\pm$  14%). This may reflect stromal or endothelial cell contributions, or tumour cell heterogeneity. A recent study of multiple samples from single tumours has demonstrated significant regional genomic heterogeneity, but with shared mutations in frequently mutated genes and convergent evolution of other common gene level events<sup>8</sup>. The mutation frequencies of key genes (*VHL*, *PBRM1* and so on), as well as copy number gains and losses found here, were, however, consistent with previous reports. Tumour purity was therefore not determined to be a limitation in the current study.

Arm level losses on chromosome 14q, associated with loss of *HIF1A*, which has been predicted to drive more aggressive disease<sup>9</sup>, were also frequent (45% of samples). Gains of 5q were observed (67% of samples) and additional focal amplifications refined the region of interest to 60 genes in 5q35, which was particularly informative as little has been known about the importance of this region in ccRCC since the 5q gain was initially described. Focal amplification also implicated the protein kinase C member *PRKCI* (ref. 10), and the MDS1 and EVI1 complex locus *MECOM* at 3p26, the p53 regulator *MDM4* at 1q32, *MYC* at 8q24 and *JAK2* on 9p24. Focally deleted regions included the tumour suppressor genes *CDKN2A* at 9p21 and *PTEN* at 10q23, putative tumour suppressor genes *NEGR1* at 1p31, *QKI* at 6q26, and *CADM2* at 3p12 and the genes that are frequently deleted in cancer, *PTPRD* at 9p23 and *NRXN3* at 14q24 (ref. 11).

Whole-exome sequencing (WES) of tumours from 417 patients identified 36,353 putative somatic mutations, including 16,821 missense mutations, 6,383 silent mutations and 2,999 indels, with an average of

\*Lists of participants and their affiliations appear at the end of the paper.



**Figure 1 | Somatic alterations in ccRCC.** **a**, Top histogram, mutation events per sample; left histogram, samples affected per alteration. Upper heat map, distribution of fusion transcripts and *VHL* methylation across samples ( $n = 385$  samples, with overlapping exome/SCNA/RNA-seq/methylation data); middle heat map, mutation events; bottom heat map, copy number gains (red) and losses (blue). Lower chart, mutation spectrum by indicated categories.

$1.1 \pm 0.5$  non-silent mutations per megabase (Supplementary Figs 23–25). Mutations from 50 genes with high apparent somatic mutation frequencies (Supplementary Table 4) were independently validated using alternative sequencing instrumentation (Supplementary Fig. 26). In tumours from 22 patients, whole-genome sequencing was also used to validate and calibrate the WES data and confirmed 83% of the WES mutation-calls (Supplementary Tables 5 and 6). In line with results of previous studies (Supplementary Tables 7 and 8), the validated mutation data identified nineteen significantly mutated genes (SMGs) (false discovery rate (FDR)  $< 0.1$ ), with *VHL*, *PBRM1*, *SETD2*, *KDM5C*, *PTEN*, *BAP1*, *MTOR* and *TP53* representing the eight most extreme members ( $q < 0.00001$ ) (Fig. 1a). Eleven additional SMGs were of considerably lower significance ( $q < 0.1–0.5$ ) but included known cancer genes. Among all SMGs, only mutation of *BAP1* correlated with poor survival outcome (Supplementary Fig. 27)<sup>12</sup>. Approximately 20% of cases had none of the 19 recorded SMGs, although many contained rare mutations in other known oncogenes or tumour suppressors, involving survival associations, illustrating the genetic complexity of ccRCC<sup>8</sup> (Supplementary Figs 28–30 and Supplementary Table 9).

Eighty-four putative RNA fusions were identified in 416 ccRCC samples<sup>13</sup>. Eleven of thirteen predicted events (Fig. 1c) were validated using targeted methods, consistent with an 85% true-positive rate (Supplementary Table 10 and Supplementary Figs 31–35). A recurrent *SFPQ-TFE3* fusion (previously linked to non-clear cell translocation-associated RCC<sup>14</sup>) was found in five samples, all of which were *VHL* wild type, indicating either that these tumours are a clear cell variant or that translocation-associated renal tumours may be histologically indistinguishable from conventional ccRCC. Furthermore, the TFE3 protein as well as an X(p11) rearrangement was found in three of those samples, where there were available slides.

## DNA methylation profiles

We observed epigenetic silencing of *VHL* in about 7% of ccRCC tumours, which was mutually exclusive with mutation of *VHL* (Fig. 1a), reflecting the central role of this locus in ccRCC<sup>15</sup>. An additional 289 genes showed

**b**, Left panel, frequency of arm-level copy-number alterations versus focal copy number alterations. Right panel, comparison of the average numbers of arm-level and focal copy-number changes in ccRCC, colon cancer (CRC), glioblastoma (GBM), breast cancer (BRCA) and ovarian cancer (OVCA).

**c**, Circos plot of fusion transcripts identified in 416 samples of ccRCC, with recurrent fusions highlighted.

evidence of epigenetic silencing in at least 5% of tumours. The top-ranked gene by inverse correlation between gene expression and DNA methylation was *UQCRH*, hypermethylated in 36% of the tumours. *UQCRH* has been previously suggested to be a tumour suppressor<sup>16</sup>, but not linked to ccRCC. Interestingly, increasing promoter hypermethylation frequency correlated with higher stage and grade (Fig. 2a, b).

We also evaluated the global consequences of mutation in specific epigenetic modifiers. Mutations in *SETD2*, a non-redundant H3K36 methyltransferase, were associated with increased loss of DNA methylation at non-promoter regions (Fig. 2c, d). This discovery is consistent with the emerging view that H3K36 trimethylation may be involved in the maintenance of a heterochromatic state<sup>17</sup>, whereby DNA methyltransferase 3A (DNMT3A) binds H3K36me3 and methylates nearby DNA<sup>18</sup>. Thus, reductions of H3K36me3 through *SETD2* inactivation could lead indirectly to regional loss of DNA methylation.

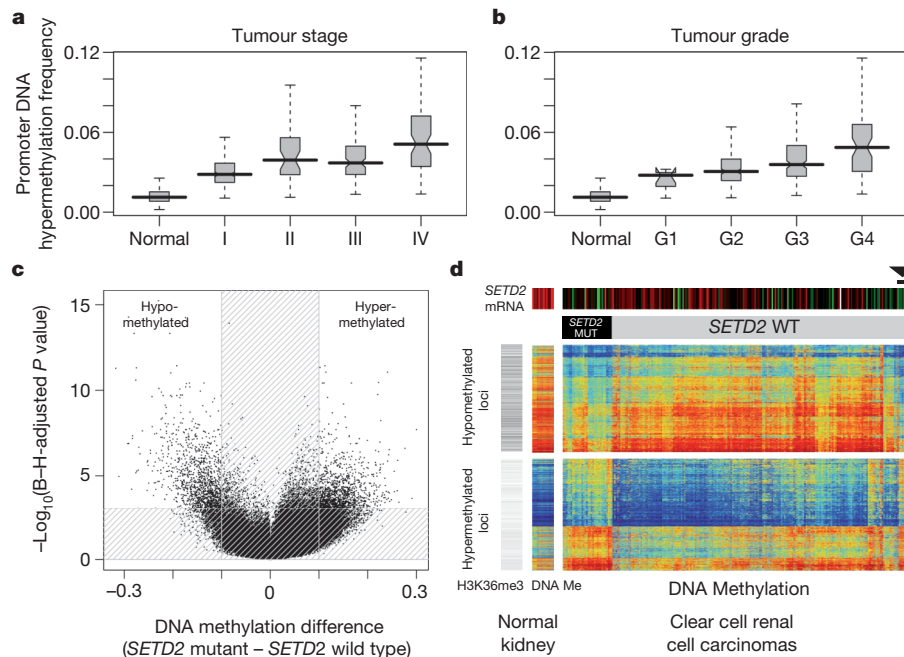
## RNA expression

Unsupervised clustering methods identified four stable subsets in both mRNA (m1–m4) and miRNA (mi1–mi4) expression data sets (Fig. 3a and Supplementary Figs 36–39). Supervised clustering revealed the similarity of these new mRNA classes to the previously reported ccA and ccB expression subtypes<sup>19</sup>, with cluster m1 corresponding to ccA and ccB divided between m2 and m3 (Supplementary Table 11). Cluster m4 probably accounts for the roughly 15% of tumours previously unclassified in the ccA/ccB classification scheme. Similarly, the survival advantage previously observed for ccA cases was again identified for m1 tumours (Fig. 3b).

The m1 subtype was characterized by gene sets associated with chromatin remodelling processes and a higher frequency of *PBRM1* mutations (39% in m1 vs 27% in others,  $P = 0.027$ ). Deletion of *CDKN2A* (53% vs 26%;  $P < 0.0001$ ) and mutations in *PTEN* (11% vs 1%;  $P < 0.0001$ ) were more frequent in m3 tumours (Supplementary Fig. 5). The m4 group showed higher frequencies of *BAP1* mutations (17% vs 7%;  $P = 0.002$ ) and base-excision repair; however, this group

**Figure 2 | DNA methylation and ccRCC.**

**a, b**, Overall promoter DNA hypermethylation frequency in the tumour increases with rising stage (a) and grade (b). The promoter DNA hypermethylation frequency is calculated as the percentage of CpG loci hypermethylated among 15,101 loci which are unmethylated in the normal kidney tissue and normal white blood cells (boxplots, median with 95% confidence interval). **c**, Volcano plots showing a comparison of DNA methylation for *SETD2* mutant versus non-mutant tumours ( $n = 224$ , HumanMethylation450 platform). Unshaded area: CpG loci with Benjamini–Hochberg (B–H) FDR = 0.001 and difference in mean beta value  $> 0.1$  ( $n = 2,557$ ). **d**, Heat map showing CpG loci with *SETD2* mutation-associated DNA methylation (from part c); blue to red indicates low to high DNA methylation. The loci are split into those hypomethylated (top panel;  $n = 1,251$ ) or hypermethylated (bottom panel;  $n = 1,306$ ) in *SETD2* mutants. Top colour bars indicate *SETD2* mRNA expression (red: high, green: low) and *SETD2* mutation status. Grey-scale row-side colour bar on left-hand side represents the relative number of overlapping reads, based on H3K36me3 ChIP-seq experiment in normal adult kidney (<http://nihroadmap.nih.gov/epigenomics/>); black, high read count. DNA methylation patterns include 14 normal kidney samples. Among the tumours without *SETD2* mutations, six (arrowhead) have both the signature pattern of *SETD2* mutation and low *SETD2* mRNA expression.



also harboured more *mTOR* mutations (12% vs 4%;  $P = 0.01$ ) and ribosomal gene sets.

Survival differences evident in miRNA-based subtypes (Supplementary Figs 40–44) correlated with the mRNA data (Fig. 3b–d). For example, miR-21, previously shown to demonstrate strong regulatory interactions in ccRCC<sup>20</sup> and with established roles in metabolism<sup>17,21,22</sup> correlated strongly with worse outcome, and DNA promoter methylation levels inversely correlated with expression of miR-21, miR-10b and miR-30a (Supplementary Tables 12–14). miRNA interactions thus represent a significant component of the epigenetic regulation observed in ccRCC.

### Integrative data analyses

We used a combination of approaches for integrative pathway analysis. The HotNet<sup>23</sup> algorithm uses a heat diffusion model, to find sub-networks distinguished by both the frequency of mutation in genes (nodes in the network) and the topology of interactions between genes (edges in the network). In ccRCC, HotNet identified twenty-five sub-networks of genes within a genome-scale protein–protein interaction network (Supplementary Table 15 and Supplementary Fig. 45). The largest and most frequently mutated network contained *VHL* and interacting partners. The second most frequently mutated sub-network included *PBRM1*, *ARID1A* and *SMARCA4*, key genes in the PBAF *SWI/SNF* chromatin remodelling complex.

We also inferred activities for known pathways, by using the PARADIGM algorithm to incorporate mutation, copy and mRNA expression data, with pathway information catalogued in public databases. This method identified a highly significant sub-network of 2,398 known regulatory interactions, connecting 1,218 molecular features (645 distinct proteins) (Supplementary Figs 46–49 and Supplementary Tables 16 and 17). Several ‘active’ transcriptional ‘hubs’ were identified, by searching for transcription factors with targets that were inferred to be active in the PARADIGM network. The active hubs found included HIF1A/ARNT, the transcription factor program activated by *VHL* mutation, as well as MYC/MAX, SP1, FOXM1, JUN and FOS. These hubs, together with

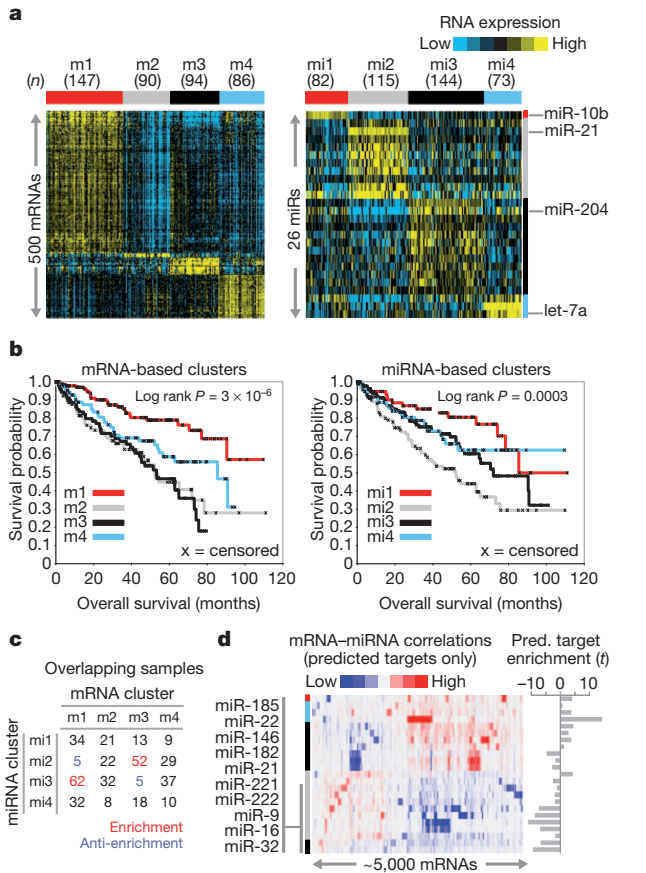
several other less well-studied transcription factors, interlink much of the transcriptional program promoting glycolytic shift, de-differentiation and growth promotion in ccRCC.

We next searched for causal regulatory interactions connecting ccRCC somatic mutations to these transcriptional hubs, using a bi-directional extension to HotNet (‘TieDIE’) and identified a chromatin-specific sub-network (Fig. 4a and Supplementary Figs 50–52). TieDIE defines a set of transcriptional targets, whose state in the tumour cells is proposed to be influenced by one or more of the significantly mutated genes. The chromatin modification pathway intersects a wide variety of processes, including the regulation of hormone receptors (for example, *ESR1*), RAS signalling via the *SRC* homologue (*SHC1*), immune-related signalling (for example, *NFKB1* and *IL6*)<sup>24</sup>, transcriptional output (for example, *HIF1A*, *JUN*, *FOS* and *SP1*), DNA repair (via *BAP1*) and beta-catenin (*CTNBN1*) and transforming growth factor (TGF)- $\beta$  (*TGFBR2*) signalling via interactions with a *SMARC-PBRM1-ARID1A* complex. The complexity of these interactions reflects the potential for highly pleiotropic effects following primary events in chromatin modification genes.

The mutations in the chromatin regulators *PBRM1*, *BAP1* and *SETD2* were differentially associated with altered expression patterns of large numbers of genes when compared to samples bearing a background of *VHL* mutation (Supplementary Tables 18–21 and Supplementary Fig. 53). Each chromatin regulator had a distinct set of downstream effects, reflecting diverse roles for chromatin remodelling in the transcriptome.

Additionally, an unsupervised pathway analysis using the MEMO algorithm<sup>25</sup> identified mutually exclusive patterns of alterations targeting multiple components of the PI(3)K/AKT/MTOR pathway in 28% of the tumours (Fig. 4b and Supplementary Table 22). Interestingly, the altered gene module included two genes from the broad amplicon on 5q35.3: *GNB2L1* and *SQSTM1*. Both these genes have previously been associated with activation of PI(3)K signalling<sup>26,27</sup>. Furthermore, mRNA expression levels of these two genes were correlated with both DNA copy number increases and alteration status of the PI(3)K pathway (Supplementary Figs 54–55). The mutual exclusivity module also includes





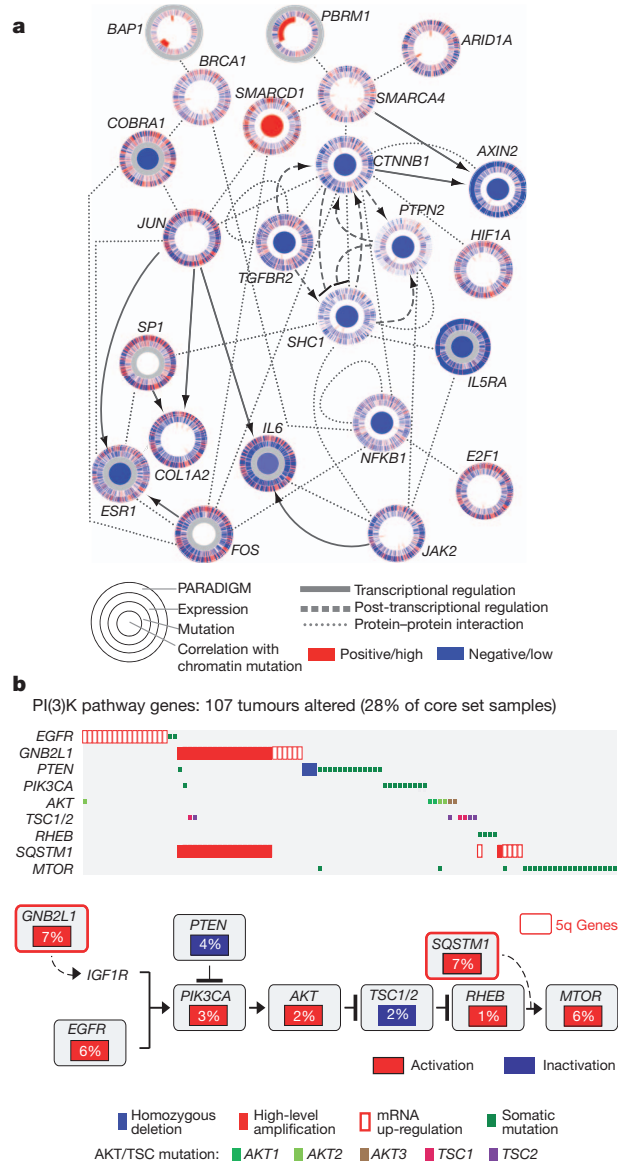
**Figure 3 | mRNA and miRNA patterns reflect molecular subtypes of ccRCC.** **a**, Tumours were separated into four sample groups (that is, ‘clusters’) by unsupervised analyses, on the basis of either differentially expressed mRNA patterns (left panel, showing 500 representative genes: m1–m4) or differentially expressed miRNA patterns (right panel, showing 26 representative miRNAs: mi1–mi4). **b**, Significant differences in patient survival were identified for both the mRNA-based clusters (left panel) and the miRNA-based clusters (right panel). **c**, Numbers of samples overlapping between the two sets of clusters, with significant concordance observed between m1 and m3 and between m3 and mi2; red, significant overlap ( $P < 10^{-5}$ , chi-squared test). **d**, mRNA–miRNA correlations, for predicted targeting interactions. Rows indicate miRNAs from **a** (indicated by cluster-specific colour bar); columns, mRNAs (5,000 differentially regulated genes selected for average RPKM  $> 10$  and at least one predicted miRNA interaction); mRNA–miRNA entries with no predicted targeting are white. To the right of the correlation matrix,  $t$  statistics (Spearman’s rank) indicate group target enrichment.

frequent overexpression of EGFR, which correlates with increased phosphorylation of the receptor (Supplementary Fig. 56), and which has been previously associated with lapatinib response in ccRCC<sup>28</sup>.

### Correlations with survival

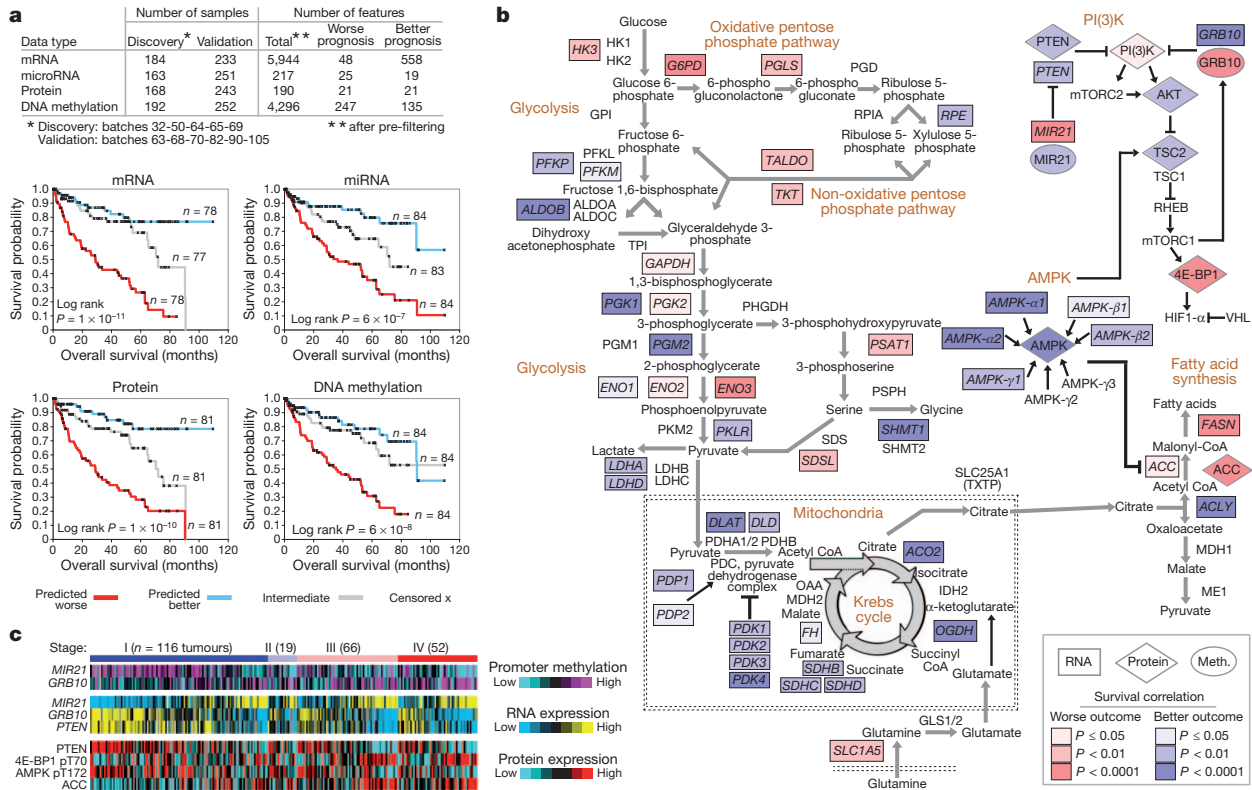
Where unsupervised analyses had indicated that common molecular patterns were associated with patient survival, we sought to further define molecular prognostic signatures at the levels of mRNA, miRNA, DNA methylation and protein. Data were divided into ‘discovery’ ( $n = 193$ ) and ‘validation’ ( $n = 253$ ) sets and platform-specific signatures were defined using Cox analyses<sup>24</sup>. Kaplan–Meier analysis for each signature showed statistically significant associations with survival in the validation subset (Fig. 5a and Supplementary Fig. 57). Multivariate Cox analyses, incorporating established clinical variables, showed that the mRNA, miRNA and protein signatures provided additional prognostic power (Supplementary Table 23). In addition, these signatures could provide molecular clues as to the drivers of aggressive cancers.

Top protein correlates of worse survival included reduced AMP-activated kinase (AMPK) and increased acetyl-CoA carboxylase (ACC)



**Figure 4 | Genomically-altered pathways in ccRCC.** **a**, Alterations in chromatin remodelling genes were predicted to affect a large network of genes and pathways (larger implicated network in Supplementary Information). 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); ‘Expression’, mRNA levels relative to normal (red, high); ‘Mutation’, somatic event; centre, correlation of gene expression or activity to mutation events in chromatin-related genes (red, positive). Protein–protein relationships inferred using public resources. **b**, For the PI(3)K/AKT/MTOR pathway (altered in ~28% of tumours), the MEMo algorithm identified a pattern of mutually exclusive gene alterations (somatic mutations, copy alterations and aberrant mRNA expression) targeting multiple components, including two genes from the recurrent amplicon on 5q35.3. The alteration frequency and inferred alteration type (blue for inactivation and red for activation) is shown for each gene in the pathway diagram.

(Supplementary Fig. 58). Together, downregulation of AMPK and upregulation of ACC activity contribute to a metabolic shift towards increased fatty acid synthesis<sup>29</sup>. A metabolic shift to an altered use of key metabolites and pathways was also apparent when considering the full set of genes involved in the core metabolic processes, including a shift towards a ‘Warburg effect’-like state (Fig. 5b). Poor prognosis correlated with downregulation of AMPK complex and the Krebs cycle genes, and with upregulation of genes involved in the pentose phosphate pathway



**Figure 5 | Molecular correlates of patient survival involve metabolic pathways.** **a**, Sample profiles were separated into discovery and validation subsets, with the top survival correlates within the discovery subset being defined for each of the four platforms examined (mRNA, microRNA, protein, DNA methylation). Kaplan–Meier plots show results of applying the four prognostic signatures to the validation subset, comparing survival for patients with predicted higher risk (red, top third of signature scores), lower risk (blue, bottom third) or intermediate risk (grey, middle third); successful predictions were observed in each case. **b**, When viewed in the context of metabolism, the molecular survival correlates highlight a widespread metabolic shift, with tumours altering their usage of key pathways and metabolites (red and blue

shading representing the correlation of increased gene expression with worse or better survival respectively, univariate Cox based on extended cohort). Worse survival correlates with upregulation of pentose phosphate pathway genes (*G6PH*, *PGLS*, *TALDO* (also known as *TALDO1P1*), *TKT*) and fatty acid synthesis (*FASN*, *ACC* (also known as *ACACA*)). Additionally, specific promoter methylation events, including hypermethylation of PI(3)K pathway repressor *GRB10*, associate with outcome. **c**, Heat map of selected key features from the metabolic shift schematic (**b**) demonstrating coordinate expression by stage at DNA methylation, RNA, and protein levels (data from validation subset).

(*G6PD*, *PGLS*, *TALDO* (also known as *TALDO1P1*), *TKT*) and fatty acid synthesis (*FASN*, *ACC* (also known as *ACACA*)).

Examination of potential genetic or epigenetic drivers of a glycolytic shift led us to identify methylation events involving *MIR21* and *GRB10*, with decreased promoter methylation of each gene (thereby higher expression) being associated with worse or better outcome, respectively (Fig. 5b, Supplementary Fig. 59 and Supplementary Table 24). Both genes regulate the PI(3)K pathway: miR-21 is inducible by high glucose levels and downregulates *PTEN*<sup>22</sup>; whereas the tumour suppressor *GRB10* negatively regulates PI(3)K and insulin signalling<sup>30</sup>. Promoter methylation of *MIR21* and *GRB10* were coordinated with their mRNA expression patterns, as well as with the mRNA expression of other key genes and protein expression in the metabolic pathways (Fig. 5c and Supplementary Fig. 60). In addition to the PI(3)K pathway (Fig. 5b and Supplementary Fig. 61), molecular survival correlations involved several pro-metastatic matrix metalloproteinases (Supplementary Fig. 62).

**Discussion**

Our study sampled a single site of the primary tumour, in a disease with a potentially high level of tumour heterogeneity<sup>8</sup>. The extent to which convergent evolutionary events are a common theme in ccRCC remains to be determined, but may indicate that critical genes will be represented across the tumour landscape for an individual mass. In general, the large sample size seemed to overcome the intrinsic challenges of studying a genetically complex disease, revealing rare variants

at rates similar to what has been described previously<sup>3</sup>. The samples, taken from primary tumour specimens, were reflective of patients fit for either definitive or cytoreductive nephrectomy, whereas future work could explore the genomic landscape of metastatic lesions.

Pathway and integrated analyses highlighted the importance of the well-known VHL/HIF pathway, the newly emerging chromatin remodelling/histone methylation pathway, and the PI(3)K/AKT pathway. The observation of chromatin modifier genes being frequently mutated in ccRCC strongly supports the model of nucleosome dynamics, providing a key function in renal tumorigenesis. Although the mechanistic details remain to be defined as to how such modulation promotes tumour formation, the data presented here revealed alterations in DNA methylation associated with *SETD2* mutations. As an epigenetic process that can potentially modify many transcriptional outputs, these mutational events have the potential to change the landscape of the tumour genome through altered expression of global sets of genes and genetic elements. Molecular correlates of patient survival further implicated PI(3)K/AKT as having a role in tumour progression, involving specific DNA methylation events. The PI(3)K/AKT pathway presents a strong therapeutic target in ccRCC, supporting the potential value of MTOR and/or related pathway inhibitor drugs for this cancer<sup>31,32</sup>.

Cross-platform molecular analyses indicated a correlation between worsened prognosis in patients with ccRCC and a metabolic shift involving increased dependence on the pentose phosphate shunt, decreased AMPK, decreased Krebs cycle activity, increased glutamine transport and fatty acid production. These findings are consistent



with the isotopomer spectral analysis of a pair of *VHL*<sup>-/-</sup> clear cell kidney cancer cell lines, both of which were notably derived from patients with aggressive, metastatic disease, which revealed a dependence on reductive glutamine metabolism for lipid biosynthesis<sup>33</sup>. The metabolic shift identified in poor prognosis ccRCC remarkably mirrors the Warburg metabolic phenotype (increased glycolysis, decreased AMPK, glutamine-dependent lipogenesis) identified in type 2 papillary kidney cancer characterized by mutation of the Krebs cycle enzyme, fumarate hydratase<sup>33</sup>. Further studies to dissect out the role of the commonly mutated chromosome 3 chromatin remodelling genes, *PBRM1*, *SETD2* and *BAP1*, in ccRCC tumorigenesis and their potential role in the metabolic remodelling associated with progression of this disease will hopefully provide the foundation for the development of effective forms of therapy for this disease.

## METHODS SUMMARY

Specimens were obtained from patients, with appropriate consent from institutional review boards. Using a co-isolation protocol, DNA and RNA were purified. In total, 446 patients were assayed on at least one molecular profiling platform, which platforms included: (1) RNA sequencing, (2) DNA methylation arrays, (3) miRNA sequencing, (4) Affymetrix single nucleotide polymorphism (SNP) arrays, (5) exome sequencing, and (6) reverse phase protein arrays. As described above and in the Supplementary Methods, both single platform analyses and integrated cross-platform analyses were performed.

Received 25 October 2012; accepted 24 April 2013.

Published online 23 June 2013.

- Linehan, W. M., Walther, M. M. & Zbar, B. The genetic basis of cancer of the kidney. *J. Urol.* **170**, 2163–2172 (2003).
- Linehan, W. M., Srinivasan, R. & Schmidt, L. S. The genetic basis of kidney cancer: a metabolic disease. *Nature Rev. Urol.* **7**, 277–285 (2010).
- Dalglish, G. L. *et al.* Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature* **463**, 360–363 (2010).
- Guo, G. *et al.* Frequent mutations of genes encoding ubiquitin-mediated proteolysis pathway components in clear cell renal cell carcinoma. *Nature Genet.* **44**, 17–19 (2012).
- Varela, I. *et al.* Exome sequencing identifies frequent mutation of the SWI/SNF complex gene *PBRM1* in renal carcinoma. *Nature* **469**, 529–542 (2011).
- Zbar, B., Brauch, H., Talmadge, C. & Linehan, M. Loss of alleles of loci on the short arm of chromosome 3 in renal cell carcinoma. *Nature* **327**, 721–724 (1987).
- Carter, S. L. *et al.* Absolute quantification of somatic DNA alterations in human cancer. *Nature Biotechnol.* **30**, 413–421 (2012).
- Gerlinger, M. *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366**, 883–892 (2012).
- Shen, C. *et al.* Genetic and functional studies implicate HIF1 $\alpha$  as a 14q kidney cancer suppressor gene. *Cancer Discov.* **1**, 222–235 (2011).
- Eder, A. M. *et al.* Atypical PKC $\alpha$  contributes to poor prognosis through loss of apical basal polarity and cyclin E overexpression in ovarian cancer. *Proc. Natl Acad. Sci. USA* **102**, 12519–12524 (2005).
- Herbers, J. *et al.* Significance of chromosome arm 14q loss in nonpapillary renal cell carcinomas. *Genes Chromosom. Cancer* **19**, 29–35 (1997).
- Hakimi, A. A. *et al.* Adverse outcomes in clear cell renal cell carcinoma with mutations of 3p21 epigenetic regulators BAP1 and SETD2: a report by MSKCC and the KIRC TCGA research network. *Clin. Cancer Res.* <http://dx.doi.org/10.1158/1078-0432.CCR-12-3886> (2013).
- Lewis, B. P., Shih, I., Jones-Rhoades, M., Bartel, D. & Burge, C. Prediction of mammalian microRNA targets. *Cell* **115**, 787–798 (2003).
- Clark, J. *et al.* Fusion of splicing factor genes *PSF* and *NonO* (*p54<sup>nrb</sup>*) to the *TFE3* gene in papillary renal cell carcinoma. *Oncogene* **15**, 2233–2239 (1997).
- Herman, J. G. *et al.* Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma. *Proc. Natl Acad. Sci. USA* **91**, 9700–9704 (1994).
- Modena, P. *et al.* *UQCRH* gene encoding mitochondrial hinge protein is interrupted by a translocation in a soft-tissue sarcoma and epigenetically inactivated in some cancer cell lines. *Oncogene* **22**, 4586–4593 (2003).
- Wagner, E. J. & Carpenter, P. B. Understanding the language of Lys36 methylation at histone H3. *Nature Rev. Mol. Cell Biol.* **13**, 115–126 (2012).
- Dhayanal, A. *et al.* The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J. Biol. Chem.* **285**, 26114–26120 (2010).
- Brannon, A. R. *et al.* Molecular stratification of clear cell renal cell carcinoma by consensus clustering reveals distinct subtypes and survival patterns. *Genes Cancer* **1**, 152–163 (2010).
- Liu, H. *et al.* Identifying mRNA targets of microRNA dysregulated in cancer: with application to clear cell renal cell carcinoma. *BMC Syst. Biol.* **4**, 51 (2010).
- Creighton, C. J. *et al.* Integrated analyses of microRNAs demonstrate their widespread influence on gene expression in high-grade serous ovarian carcinoma. *PLoS ONE* **7**, e34546 (2012).

- Dey, N. *et al.* MicroRNA-21 orchestrates high glucose-induced signals to TOR complex 1, resulting in renal cell pathology in diabetes. *J. Biol. Chem.* **286**, 25586–25603 (2011).
- Vandin, F., Upfal, E. & Raphael, B. J. Algorithms for detecting significantly mutated pathways in cancer. *J. Comput. Biol.* **18**, 507–522 (2011).
- The Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609–615 (2011).
- Ciriello, G., Cerami, E., Sander, C. & Schultz, N. Mutual exclusivity analysis identifies oncogenic network modules. *Genome Res.* **22**, 398–406 (2012).
- He, X., Wang, J., Messing, E. M. & Wu, G. Regulation of receptor for activated C kinase 1 protein by the von Hippel–Lindau tumor suppressor in IGF-I-induced renal carcinoma cell invasiveness. *Oncogene* **30**, 535–547 (2011).
- Duran, A. *et al.* p62 is a key regulator of nutrient sensing in the mTORC1 pathway. *Mol. Cell* **44**, 134–146 (2011).
- Ravaud, A. *et al.* Lapatinib versus hormone therapy in patients with advanced renal cell carcinoma: a randomized phase III clinical trial. *J. Clin. Oncol.* **26**, 2285–2291 (2008).
- Tong, W. H. *et al.* The glycolytic shift in fumarate-hydratase-deficient kidney cancer lowers AMPK levels, increases anabolic propensities and lowers cellular iron levels. *Cancer Cell* **20**, 315–327 (2011).
- Yu, Y. *et al.* Phosphoproteomic analysis identifies Grb10 as an mTORC1 substrate that negatively regulates insulin signaling. *Science* **332**, 1322–1326 (2011).
- Motzer, R. J. *et al.* Efficacy of everolimus in advanced renal cell carcinoma: a double-blind, randomised, placebo-controlled phase III trial. *Lancet* **372**, 449–456 (2008).
- Hudes, G. *et al.* Temsirolimus, interferon alfa, or both for advanced renal-cell carcinoma. *N. Engl. J. Med.* **356**, 2271–2281 (2007).
- Metallo, C. M. *et al.* Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia. *Nature* **481**, 380–384 (2012).

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

**Acknowledgements** We wish to thank all patients and families who contributed to this study. A full list of grant support and acknowledgments is included in the supplement.

**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 centers, cancer genome-characterization centers and genome data analysis centers. All data were released through the Data Coordinating Center. Project activities were coordinated by the NCI and NHGRI project teams. The following TCGA investigators of the Kidney Analysis Working Group contributed substantially to the analysis and writing of this manuscript: Project leaders: Richard A. Gibbs, W. Marston Linehan. Data Coordinator: Margaret Morgan. Analysis Coordinators: Chad J. Creighton, Roel G. W. Verhaak. Manuscript Coordinators: Richard A. Gibbs, Chad J. Creighton. Writing Team: W. Marston Linehan, Chad J. Creighton, W. Kimryn Rathmell, Roel G. W. Verhaak, Richard A. Gibbs. DNA Sequence analysis: David A. Wheeler, Kristian Cibulskis. mRNA analysis: Roel G. W. Verhaak, A. Rose Brannon, W. Kimryn Rathmell, Wandaliz Torres-Garcia. microRNA analysis: A. Gordon Robertson, Andy Chu, Preethi H. Gunaratne. DNA methylation analysis: Hui Shen, Peter W. Laird. Copy number analysis: Rameen Beroukhi, Sabina Signoretti. Protein analysis: Dimitra Tsavachidou, Yiling Lu, Gordon B. Mills. Pathway/Integrated Analysis: Rehan Akbani, Giovanni Ciriello, Chad J. Creighton, Suzanne S. Fei, Anders Jacobsen, Evan O. Paull, Ben Raphael, Sheila M. Reynolds, Christopher J. Ricketts, Nikolaus Schultz, Joshua M. Stuart, Fabio Vandin. Clinical Data: W. Kimryn Rathmell, A. Ari Hakimi, Johanna Gardener, Candace Shelton. Pathology and Clinical Expertise: James Hsieh, Marston W. Linehan, Pheroze Tamboli, W. Kimryn Rathmell, Victor Reuter.

**Author Information** All of the primary sequence files are deposited in CGHub (file IDs in Supplementary Information) and all other data are deposited at the Data Coordinating Center (DCC) for public access (<http://cancergenome.nih.gov/>). Reprints and permissions information is available at [www.nature.com/reprints](http://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 R.A.G. (agibbs@bcm.edu) or W.M.L. (linehan@mail.nih.gov).



This work is licensed under a Creative Commons Attribution-NonCommercial-Share Alike 3.0 Unported licence. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-sa/3.0>

## The Cancer Genome Atlas Research Network

**Analysis working group: Baylor College of Medicine** Chad J. Creighton<sup>1,2</sup>, Margaret Morgan<sup>1</sup>, Preethi H. Gunaratne<sup>1,3</sup>, David A. Wheeler<sup>1</sup>, Richard A. Gibbs<sup>1</sup>; **BC Cancer Agency** A. Gordon Robertson<sup>4</sup>, Andy Chu<sup>4</sup>; **Broad Institute** Rameen Beroukhi<sup>5,6</sup>, Kristian Cibulskis<sup>6</sup>; **Brigham & Women's Hospital** Sabina Signoretti<sup>7,54,59</sup>; **Brown University** Fabio Vandin<sup>8</sup>; Hsin-Ta Wu<sup>8</sup>, Benjamin J. Raphael<sup>8</sup>; **The University of Texas MD Anderson Cancer Center** Roel G. W. Verhaak<sup>9</sup>, Pheroze Tamboli<sup>10</sup>, Wandaliz Torres-Garcia<sup>9</sup>, Rehan Akbani<sup>9</sup>, John N. Weinstein<sup>9</sup>; **Memorial Sloan-Kettering Cancer Center** Victor Reuter<sup>11</sup>, James J. Hsieh<sup>12</sup>, A. Rose Brannon<sup>11</sup>, A. Ari Hakimi<sup>12</sup>, Anders Jacobsen<sup>13</sup>, Giovanni Ciriello<sup>13</sup>, Boris Reva<sup>13</sup>; **National Cancer Institute** Christopher J. Ricketts<sup>14</sup>, W. Marston Linehan<sup>14</sup>; **University of California Santa Cruz** Joshua M. Stuart<sup>15</sup>; **University of North Carolina, Chapel Hill** W. Kimryn Rathmell<sup>16</sup>; **University of Southern California** Hui Shen<sup>17</sup>, Peter W. Laird<sup>17</sup> **Genome sequencing centres: Baylor College of Medicine** Donna Muzny<sup>1</sup>, Caleb Davis<sup>1</sup>, Margaret Morgan<sup>1</sup>, Liu Xi<sup>1</sup>,

Kyle Chang<sup>1</sup>, Nipun Kakkar<sup>1</sup>, Lisa R. Treviño<sup>1</sup>, Susan Benton<sup>1</sup>, Jeffrey G. Reid<sup>1</sup>, Donna Morton<sup>1</sup>, Harsha Doddapaneni<sup>1</sup>, Yi Han<sup>1</sup>, Lora Lewis<sup>1</sup>, Huyen Dinh<sup>1</sup>, Christie Kovar<sup>1</sup>, Yiming Zhu<sup>1</sup>, Jireh Santibanez<sup>1</sup>, Min Wang<sup>1</sup>, Walker Hale<sup>1</sup>, Divya Kalra<sup>1</sup>, Chad J. Creighton<sup>1,2</sup>, David A. Wheeler<sup>1</sup>, Richard A. Gibbs<sup>1</sup>; **Broad Institute** Gad Getz<sup>6,57</sup>, Kristian Cibulskis<sup>6</sup>, Michael S. Lawrence<sup>6</sup>, Carrie Sougnez<sup>6</sup>, Scott L. Carter<sup>6</sup>, Andrey Sivachenko<sup>6</sup>, Lee Lichtenstein<sup>6</sup>, Chip Stewart<sup>6</sup>, Doug Voet<sup>6</sup>, Sheila Fisher<sup>6</sup>, Stacey B. Gabriel<sup>6</sup>, Eric Lander<sup>6</sup> **Genome characterization centres: Broad Institute** Rameen Beroukhi<sup>5,6,54,56</sup>, Steve E. Schumacher<sup>6,56</sup>, Barbara Tabak<sup>6,56</sup>, Gordon Saxena<sup>6</sup>, Robert C. Onofrio<sup>6</sup>, Scott L. Carter<sup>6</sup>, Andrew D. Cherniack<sup>6</sup>, Jeff Gentry<sup>6</sup>, Kristin Ardlie<sup>6</sup>, Carrie Sougnez<sup>6</sup>, Gad Getz<sup>6,57</sup>, Stacey B. Gabriel<sup>6</sup>, Matthew Meyerson<sup>6,7,54</sup>; **BC Cancer Agency** A. Gordon Robertson<sup>4</sup>, Andy Chu<sup>4</sup>, Hye-Jung E. Chun<sup>4</sup>, Andrew J. Mungall<sup>4</sup>, Payal Sipahimalani<sup>4</sup>, Dominik Stolf<sup>4</sup>, Adrian Ally<sup>4</sup>, Miruna Balasundaram<sup>4</sup>, Yaron S. N. Butterfield<sup>4</sup>, Rebecca Carlsen<sup>4</sup>, Candace Carter<sup>4</sup>, Eric Chuah<sup>4</sup>, Robin J. N. Coope<sup>4</sup>, Noreen Dhalla<sup>4</sup>, Sharon Gorski<sup>4</sup>, Ranabir Guin<sup>4</sup>, Carrie Hirst<sup>4</sup>, Martin Hirst<sup>4</sup>, Robert A. Holt<sup>4</sup>, Chandra Lebovitz<sup>4</sup>, Darlene Lee<sup>4</sup>, Haiyan I. Li<sup>4</sup>, Michael Mayo<sup>4</sup>, Richard A. Moore<sup>4</sup>, Erin Pleasance<sup>4</sup>, Patrick Plettnar<sup>4</sup>, Jacqueline E. Schein<sup>4</sup>, Arash Shafiei<sup>4</sup>, Jared R. Slobodan<sup>4</sup>, Angela Tam<sup>4</sup>, Nina Thiessen<sup>4</sup>, Richard J. Varhol<sup>4</sup>, Natasia Wye<sup>4</sup>, Yongjun Zhao<sup>4</sup>, Inanc Biro<sup>4</sup>, Steven J. M. Jones<sup>4</sup>, Marco A. Marra<sup>4</sup>; **University of North Carolina, Chapel Hill** J. Todd Auman<sup>18</sup>, Donghui Tan<sup>19</sup>, Corbin D. Jones<sup>20</sup>, Katherine A. Hoadley<sup>16,21,22</sup>, Piotr A. Mieczkowski<sup>22</sup>, Lisle E. Mose<sup>21</sup>, Stuart R. Jefferys<sup>22</sup>, Michael D. Topal<sup>21,22</sup>, Christina Liquori<sup>16</sup>, Yidi J. Turman<sup>16</sup>, Yan Shi<sup>16</sup>, Scot Waring<sup>16</sup>, Elizabeth Buda<sup>16</sup>, Jesse Walsh<sup>16</sup>, Junyuan Wu<sup>16</sup>, Tom Bodenheimer<sup>16</sup>, Alan P. Hoyle<sup>16</sup>, Janae V. Simons<sup>16</sup>, Mathew G. Soloway<sup>16</sup>, Saianand Balu<sup>16</sup>, Joel S. Parker<sup>16</sup>, D. Neil Hayes<sup>16,23</sup>, Charles M. Perou<sup>16,21,22</sup>; **Harvard Medical School** Raju Kucherlapati<sup>24,25</sup>, Peter Park<sup>25,26,27</sup>; **University of Southern California & Johns Hopkins University** Hui Shen<sup>17</sup>, Timothy Triche Jr<sup>17</sup>, Daniel J. Weisenberger<sup>17</sup>, Phillip H. Lai<sup>17</sup>, Moiz S. Bootwalla<sup>17</sup>, Dennis T. Magiinte<sup>17</sup>, Swapna Mahurkar<sup>17</sup>, Benjamin P. Berman<sup>17</sup>, David J. Van Den Berg<sup>17</sup>, Leslie Cope<sup>28</sup>, Stephen B. Bayliss<sup>28</sup>, Peter W. Laird<sup>17</sup>; **Genome data analysis: Baylor College of Medicine** Chad J. Creighton<sup>12</sup>, David A. Wheeler<sup>17</sup>; **Broad Institute** Gad Getz<sup>6,57</sup>, Michael S. Noble<sup>6</sup>, Daniel DiCara<sup>6</sup>, Hailei Zhang<sup>6</sup>, Juok Cho<sup>6</sup>, David I. Heiman<sup>6</sup>, Nils Gehlenborg<sup>6,26</sup>, Doug Voet<sup>6</sup>, William Mallard<sup>6</sup>, Pei Lin<sup>6</sup>, Scott Frazier<sup>6</sup>, Petar Stojanov<sup>6,54</sup>, Yingchun Liu<sup>6</sup>, Lihua Zhou<sup>6</sup>, Jaegil Kim<sup>6</sup>, Michael S. Lawrence<sup>6</sup>, Lynda Chin<sup>6,31</sup>; **Brown University** Fabio Vandin<sup>8</sup>, Hsin-Ta Wu<sup>8</sup>, Benjamin J. Raphael<sup>8</sup>; **Buck Institute for Research on Aging** Christopher Benz<sup>55</sup>, Christina Yau<sup>55</sup>; **Institute for Systems Biology** Sheila M. Reynolds<sup>29</sup>, Ilya Shmulevich<sup>29</sup>; **The University of Texas MD Anderson Cancer Center** Roel G.W. Verhaak<sup>9</sup>, Wandaliz Torres-Garcia<sup>9</sup>, Rahul Vegesna<sup>9</sup>, Hoon Kim<sup>9</sup>, Wei Zhang<sup>10</sup>, David Cogdell<sup>10</sup>, Eric Jonasz<sup>9</sup>, Zhiyong Ding<sup>9</sup>, Yiling Lu<sup>30</sup>, Rehan Akbani<sup>9</sup>, Nianxiang Zhang<sup>9</sup>, Anna K. Unruh<sup>9</sup>, Tod D. Casasent<sup>9</sup>, Chris Wakefield<sup>9</sup>, Dimitra Tsavachidou<sup>30</sup>, Lynda Chin<sup>6,31</sup>, Gordon B. Mills<sup>30</sup>, John N. Weinstein<sup>9</sup>; **Memorial Sloan-Kettering Cancer Center** Anders Jacobsen<sup>13</sup>, A. Rose Brannon<sup>11</sup>, Giovanni Ciriello<sup>13</sup>, Nikolaus Schultz<sup>13</sup>, A. Ari Hakimi<sup>12</sup>, Boris Reva<sup>13</sup>, Yevgeniy Antipin<sup>13</sup>, Jianjiong Gao<sup>13</sup>, Ethan Cerami<sup>13</sup>, Benjamin Gross<sup>13</sup>, B. Arman Aksoy<sup>13</sup>, Rileen Sinha<sup>13</sup>, Nils Weinhold<sup>13</sup>, S. Onur Sumer<sup>13</sup>, Barry S. Taylor<sup>13</sup>, Ronglai Shen<sup>13</sup>, Irina Ostrovnya<sup>32</sup>, James J. Hsieh<sup>12</sup>, Michael F. Berger<sup>11</sup>, Marc Ladanyi<sup>12</sup>, Chris Sander<sup>13</sup>; **Oregon Health & Science University** Suzanne S. Fei<sup>33</sup>, Andrew Stout<sup>33</sup>, Paul T. Spellman<sup>33</sup>; **Stanford University** Daniel L. Rubin<sup>34</sup>, Tiffany T. Liu<sup>34</sup>; **University of California Santa Cruz** Joshua M. Stuart<sup>15</sup>, Sam Ng<sup>15</sup>, Evan O. Paull<sup>15</sup>, Daniel Carlin<sup>15</sup>, Theodore Goldstein<sup>15</sup>, Peter Waltman<sup>15</sup>, Kyle Elliott<sup>15</sup>, Jing Zhu<sup>15</sup>, David Haussler<sup>15,35</sup>; **University of Houston** Preethi H. Gunaratne<sup>13</sup>, Weimin Xiao<sup>3</sup>; **Biospecimen core resource: International Genomics Consortium** Candace Shelton<sup>36</sup>, Johanna Gardner<sup>36</sup>, Robert Penny<sup>36</sup>, Mark Sherman<sup>36</sup>, David Mallery<sup>36</sup>, Scott Morris<sup>36</sup>, Joseph Paulauskis<sup>36</sup>, Ken Burnett<sup>36</sup>, Troy Shelton<sup>36</sup>; **Tissue source sites: Brigham & Women's Hospital** Sabina Signoretti<sup>7,54,59</sup>, **Dana-Farber Cancer Institute** William G. Kaelin<sup>54,60</sup>, Toni Choueiri<sup>54</sup>; **Georgetown University** Michael B. Atkins<sup>37</sup>; **International Genomics Consortium** Robert Penny<sup>36</sup>, Ken Burnett<sup>36</sup>, David Mallery<sup>36</sup>, Erin Curley<sup>36</sup>; **Memorial Sloan-Kettering Cancer Center** Satish Tickoo<sup>11</sup>, Victor Reuter<sup>11</sup>; **University of North Carolina at Chapel Hill** W. Kimryn Rathmell<sup>16</sup>, Leigh Thorne<sup>16</sup>, Lori Boice<sup>16</sup>, Mei Huang<sup>16</sup>, Jennifer C. Fisher<sup>16</sup>; **National Cancer Institute** W. Marston Linehan<sup>14</sup>, Cathy D. Vocke<sup>14</sup>, James Peterson<sup>14</sup>, Robert Worrell<sup>14</sup>, Maria J. Merino<sup>14</sup>, Laura S. Schmidt<sup>14,38</sup>; **The University of Texas MD Anderson Cancer Center** Pheroze Tamboli<sup>10</sup>, Bogdan A. Czerniak<sup>10</sup>, Kenneth D. Aldape<sup>10</sup>, Christopher G. Wood<sup>39</sup>; **Fox Chase Cancer Center** Jeff Boyd<sup>40</sup>, JoEllen Weaver<sup>40</sup>; **Helen F. Graham Cancer Center at Christiana Care** Mary W. Iacocca<sup>41</sup>, Nicholas Petrelli<sup>41</sup>, Gary Witkin<sup>41</sup>, Jennifer Brown<sup>41</sup>, Christine Czerwinski<sup>41</sup>, Lori Huelsenbeck-Dill<sup>41</sup>, Brenda Rabeno<sup>41</sup>; **Penrose-St. Francis Health Services** Jerome Myers<sup>42</sup>, Carl Morrison<sup>42</sup>, Julie Bergsten<sup>42</sup>, John Eckman<sup>42</sup>, Jodi Harr<sup>42</sup>, Christine Smith<sup>42</sup>, Kelinda Tucker<sup>42</sup>, Leigh Anne Zach<sup>42</sup>; **Roswell Park Cancer Institute** Wiam Bshara<sup>43</sup>, Carmelo Gaudioso<sup>43</sup>, Carl Morrison<sup>43</sup>; **University of Pittsburgh** Rajiv Dhir<sup>44</sup>, Jodi Maranchie<sup>44</sup>, Joel Nelson<sup>44</sup>, Anil Parwani<sup>44</sup>; **Cureline** Olga Potapova<sup>45</sup>; **St. Petersburg City Clinical Oncology Dispensary** Konstantin Fedosenko<sup>46</sup>; **Mayo Clinic** John C. Cheville<sup>58</sup>, R. Houston Thompson<sup>58</sup>; **Disease working group: Brigham & Women's Hospital** Sabina Signoretti<sup>7,54,59</sup>; **Dana-Farber Cancer Institute** William G. Kaelin<sup>54,60</sup>; **Georgetown University** Michael B. Atkins<sup>37</sup>; **Memorial Sloan-Kettering Cancer Center** Satish Tickoo<sup>11</sup>, Victor Reuter<sup>11</sup>; **National Cancer Institute** W. Marston Linehan<sup>14</sup>, Cathy D. Vocke<sup>14</sup>, James Peterson<sup>14</sup>, Maria J. Merino<sup>14</sup>, Laura S. Schmidt<sup>14,38</sup>; **The University of Texas MD Anderson Cancer Center** Pheroze Tamboli<sup>10</sup>; **Weill Cornell Medical College** Juan M. Mosquera<sup>47</sup>, Mark A. Rubin<sup>47</sup>; **Massachusetts General Hospital** Michael L. Blute<sup>48</sup>; **University of North Carolina, Chapel Hill** W. Kimryn Rathmell<sup>16</sup>; **Data coordination centre: Todd Pihl**<sup>49</sup>, Mark Jensen<sup>49</sup>, Robert Sfeir<sup>49</sup>, Ari Kahn<sup>49</sup>, Anna Chu<sup>49</sup>, Prachi Kothiyal<sup>49</sup>, Eric Snyder<sup>49</sup>, Joan Pontius<sup>49</sup>, Brenda Ayala<sup>49</sup>, Mark Backus<sup>49</sup>, Jessica Walton<sup>49</sup>, Julien Baboud<sup>49</sup>, Dominique Berton<sup>49</sup>, Matthew Nicholls<sup>49</sup>, Deepak Srinivasan<sup>49</sup>, Rohini Raman<sup>49</sup>, Stanley Girshik<sup>49</sup>, Peter Kigonya<sup>49</sup>, Shelley Alonso<sup>49</sup>, Rashmi Sanbhatti<sup>49</sup>, Sean Barletta<sup>49</sup>, David Pot<sup>49</sup>; **Project team: National Cancer Institute** Margi Sheth<sup>50</sup>, John A. Demchok<sup>50</sup>, Tanja Davidsen<sup>50</sup>, Zhining Wang<sup>50</sup>, Liming Yang<sup>50</sup>, Roy W. Tarnuzzer<sup>50</sup>, Jiashan Zhang<sup>50</sup>, Greg Eley<sup>51</sup>, Martin L. Ferguson<sup>52</sup>, Kenna R. Mills Shaw<sup>50</sup>; **National Human Genome Research Institute** Mark S. Guyer<sup>53</sup>, Bradley A. Ozenberger<sup>53</sup> & Heidi J. Sofia<sup>53</sup>.

<sup>1</sup>Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas 77030, USA. <sup>2</sup>Dan L Duncan Cancer Center, Baylor College of Medicine, Houston, Texas 77030, USA. <sup>3</sup>Department of Biology & Biochemistry, University of Houston, Houston, Texas 77204, USA. <sup>4</sup>Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency, Vancouver, British Columbia V5Z, Canada. <sup>5</sup>Department of Medicine, Harvard Medical School, Boston, Massachusetts 02215, USA. <sup>6</sup>The Eli and Edythe L. Broad Institute of Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts 02142, USA. <sup>7</sup>Department of Pathology, Harvard Medical School, Boston, Massachusetts 02215, USA. <sup>8</sup>Department of Computer Science, Brown University, Providence, Rhode Island 02912, USA. <sup>9</sup>Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. <sup>10</sup>Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. <sup>11</sup>Department of Pathology, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. <sup>12</sup>Human Oncology and Pathogenesis Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. <sup>13</sup>Computational Biology Center, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. <sup>14</sup>Urologic Oncology Branch, Center for Cancer Research, National Cancer Institute, Bethesda, Maryland 20892, USA. <sup>15</sup>Department of Biomolecular Engineering and Center for Biomolecular Science and Engineering, University of California Santa Cruz, Santa Cruz, California 95064, USA. <sup>16</sup>Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. <sup>17</sup>USC Epigenome Center, University of Southern California, Los Angeles, California 90033, USA. <sup>18</sup>Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. <sup>19</sup>Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. <sup>20</sup>Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. <sup>21</sup>Department of Pathology and Laboratory Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. <sup>22</sup>Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. <sup>23</sup>Department of Internal Medicine, Division of Medical Oncology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. <sup>24</sup>Department of Genetics, Harvard Medical School, Boston, Massachusetts 02215, USA. <sup>25</sup>Division of Genetics, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. <sup>26</sup>The Center for Biomedical Informatics, Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>27</sup>Informatics Program, Children's Hospital, Boston, Massachusetts 02115, USA. <sup>28</sup>Cancer Biology Division, The Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins University, Baltimore, Maryland 21231, USA. <sup>29</sup>Institute for Systems Biology, Seattle, Washington 98109, USA. <sup>30</sup>Department of Systems Biology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. <sup>31</sup>Department of Genomic Medicine, University of Texas MD Anderson Cancer Center, Houston, Texas 77054, USA. <sup>32</sup>Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. <sup>33</sup>Department of Molecular & Medical Genetics, Oregon Health & Science University, Portland, Oregon 97239, USA. <sup>34</sup>Department of Radiology, Stanford University Medical Center, Stanford, California 94305, USA. <sup>35</sup>Howard Hughes Medical Institute, University of California Santa Cruz, Santa Cruz, California 95064, USA. <sup>36</sup>International Genomics Consortium, Phoenix, Arizona 85004, USA. <sup>37</sup>Georgetown-Lombardi Comprehensive Cancer Center, Georgetown University, Washington DC 20057, USA. <sup>38</sup>Basic Science Program, SAIC-Frederick, Inc., Frederick National Lab, Frederick, Maryland 21702, USA. <sup>39</sup>Department of Urology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. <sup>40</sup>Cancer Biology Program, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111, USA. <sup>41</sup>Helen F. Graham Cancer Center, Christiana Care, Newark, Delaware 19713, USA. <sup>42</sup>Penrose-St. Francis Health Services, Colorado Springs, Colorado 80907, USA. <sup>43</sup>Department of Pathology, Roswell Park Cancer Institute, Buffalo, New York 14263, USA. <sup>44</sup>Department of Pathology, University of Pittsburgh, Pittsburgh, Pennsylvania 15213, USA. <sup>45</sup>Cureline, Inc. South San Francisco, California 94080, USA. <sup>46</sup>St Petersburg City Clinical Oncology Dispensary, St Petersburg 198255, Russia. <sup>47</sup>Department of Pathology and Laboratory Medicine, Weill Cornell College of Medicine, New York, New York 10065, USA. <sup>48</sup>Department of Urology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. <sup>49</sup>SRA International, 4300 Fair Lakes Court, Fairfax, Virginia 22033, USA. <sup>50</sup>The Cancer Genome Atlas Program Office, Center for Cancer Genomics, National Cancer Institute, Bethesda, Maryland 20892, USA. <sup>51</sup>TCGA Consultant, Scicentis, LLC, Atlanta, Georgia 30666, USA. <sup>52</sup>MLF Consulting, Arlington, Massachusetts 02474, USA. <sup>53</sup>National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892, USA. <sup>54</sup>Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts 02215, USA. <sup>55</sup>Buck Institute for Research on Aging, Novato, California 94945, USA. <sup>56</sup>Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, Massachusetts 02215, USA. <sup>57</sup>Cancer Center and Department of Pathology, Massachusetts General Hospital, 55 Fruit St, Boston, Massachusetts 02114. <sup>58</sup>Mayo Clinic, Rochester, Minnesota 55905, USA. <sup>59</sup>Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts 02215, USA. <sup>60</sup>Howard Hughes Medical Institute, Chevy Chase, Maryland 20815, USA.