Comprehensive molecular profiling of lung adenocarcinoma

The Cancer Genome Atlas Research Network*

Adenocarcinoma of the lung is the leading cause of cancer death worldwide. Here we report molecular profiling of 230 resected lung adenocarcinomas using messenger RNA, microRNA and DNA sequencing integrated with copy number, methylation and proteomic analyses. High rates of somatic mutation were seen (mean 8.9 mutations per megabase). Eighteen genes were statistically significantly mutated, including RIT1 activating mutations and newly described loss-of-function MGA mutations which are mutually exclusive with focal MYC amplification. EGF mutations were more frequent in female patients, whereas mutations in RBM10 were more common in males. Aberrations in NFI, MET, ERBB2 and RIT1 occurred in 13% of cases and were enriched in samples otherwise lacking an activated oncogene, suggesting a driver role for these events in certain tumours. DNA and mRNA sequence from the same tumour highlighted splicing alterations driven by somatic genomic changes, including exon 14 skipping in MET mRNA in 4% of cases. MAPK and PI(3)K pathway activity, when measured at the protein level, was explained by known mutations in only a fraction of cases, suggesting additional, unexplained mechanisms of pathway activation. These data establish a foundation for classification and further investigations of lung adenocarcinoma molecular pathogenesis.

Lung cancer is the most common cause of global cancer-related mortality, leading to over a million deaths each year and adenocarcinoma is its most common histological type. Smoking is the major cause of lung adenocarcinoma but, as smoking rates decrease, proportionally more cases occur in never-smokers (defined as less than 100 cigarettes in a lifetime). Recently, molecularly targeted therapies have dramatically improved outcomes for patients whose tumours harbour somatically activated oncogenes such as mutant EGFR or translocated ALK, RET, or ROS1 (refs 2–4). Mutant BRAF and ERBB2 (ref. 5) are also investigational targets. However, most lung adenocarcinomas either lack an identifiable driver oncogene, or harbour mutations in KRAS and are therefore still treated with conventional chemotherapy. Tumour suppressor gene abnormalities, such as those in TP53 (ref. 6), STK11 (ref. 7), CDKN2A (ref. 8), KEAP1 (ref. 9), and SMARCA4 (ref. 10) are also common but are not currently clinically actionable. Finally, lung adenocarcinoma shows high rates of somatic mutation and genomic rearrangement, challenging identification of all but the most frequent driver gene alterations because of a large burden of passenger events per tumour genome11-13. Our efforts focused on comprehensive, multiplatform analysis of lung adenocarcinoma, with attention towards pathobiology and clinically actionable events.

Clinical samples and histopathologic data

We analysed tumour and matched normal material from 230 previously untreated lung adenocarcinoma patients who provided informed consent (Supplementary Table 1). All major histologic types of lung adenocarcinoma were represented: 5% lepidic, 33% acinar, 9% papillary, 14% micropapillary, 25% solid, 4% invasive mucinous, 0.4% colloid and 8% unclassifiable adenocarcinoma (Supplementary Fig. 1). Median follow-up was 19 months, and 163 patients were alive at the time of last follow-up. Eighty-one percent of patients reported past or present smoking. Supplementary Table 2 summarizes demographics. DNA, RNA and protein were extracted from specimens and quality-control assessments were performed as described previously15. Supplementary Table 3 summarizes molecular estimates of tumour cellularity16.

Figure 1 | Somatic mutations in lung adenocarcinoma. a, Co-mutation plot from whole exome sequencing of 230 lung adenocarcinomas. Data from TCGA samples were combined with previously published data12 for statistical analysis. Co-mutation plot for all samples used in the statistical analysis (n = 412) can be found in Supplementary Fig. 2. Significant genes with a corrected P value less than 0.025 were identified using the MutSig2CV algorithm and are ranked in order of decreasing prevalence. b, c. The differential patterns of mutation between samples classified as transversion high and transversion low samples (b) or male and female patients (c) are shown for all samples used in the statistical analysis (n = 412). Stars indicate statistical significance using the Fisher’s exact test (black stars: P < 0.05, grey stars: P < 0.05) and are adjacent to the sample set with the higher percentage of mutated samples.

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

©2014 Macmillan Publishers Limited. All rights reserved
Somatically acquired DNA alterations

We performed whole-exome sequencing (WES) on tumour and germline DNA, with a mean coverage of 97.6× and 95.8×, respectively, as performed previously. The mean somatic mutation rate across the TCGA cohort was 8.87 mutations per megabase (Mb) of DNA (range: 0.5–48, median: 5.78). The non-synonymous mutation rate was 6.86 per Mb. MutSigCV identified significantly mutated genes among our 230 cases along with 182 similarly-sequenced, previously reported lung adenocarcinomas. Analysis of these 412 tumour/normal pairs highlighted 18 statistically significant mutated genes (Fig. 1a) shows co-mutation plot of TCGA samples (n = 230), Supplementary Fig. 2 shows co-mutation plot of all tumour samples used in the statistical analysis (n = 412) and Supplementary Table 4 contains complete MutSig2CV results, which also appear on the TCGA Data Portal along with many associated data files (https://tcga-data.nci.nih.gov/docs/publications/luad_2014/). TP53 was commonly mutated (46%). Mutations in KRAS (33%) were mutually exclusive with those in EGFR (14%), BRAF was also commonly mutated (10%), as were PIK3CA (7%), MET (7%) and the small GTPase gene, RIT1 (2%). Mutations in tumour suppressor genes including STK11 (17%), KEAP1 (17%), NFI1 (11%), RB1 (4%) and CDKN2A (4%) were observed. Mutations in chromatin modifying genes SETD2 (9%), ARID1A (7%) and SMARCA4 (6%) and the RNA splicing genes RBM10 (8%) and U2AF1 (3%) were also common. Recurrent mutations in the MGA gene (which encodes a Max-interacting protein on the MYC pathway) occurred in 8% of samples. Loss-of-function (frameshift and nonsense) mutations in MGA were mutually exclusive with focal MYC amplification (Fisher’s exact test P = 0.04), suggesting a hitherto unappreciated potential mechanism of MYC pathway activation. Coding single nucleotide variants and indel variants were verified by resequencing at a rate of 99% and 100%, respectively (Supplementary Fig. 3a, Supplementary Table 5). Tumour purity was not associated with the presence of false negatives identified in the validation data (P = 0.31; Supplementary Fig. 3b).

Past or present smoking associated with cytosine to adenine (C to A) nucleotide transversions as previously described both in individual genes and genome-wide. C to A nucleotide transversion fraction showed two peaks; this fraction correlated with total mutation count (R² = 0.30) and inversely correlated with cytosine to thymine (C to T) transition frequency (R² = 0.75) (Supplementary Fig. 4). We classified each sample (Supplementary Methods) into one of two groups named transversion-high (TH, n = 269), and transversion-low (TL, n = 144). The transversion-high group was strongly associated with past or present smoking (P < 2.2 × 10⁻¹⁶), consistent with previous reports. The transversion-high and transversion-low patient cohorts harboured different gene mutations. Whereas KRAS mutations were significantly enriched in the transversion-high cohort (P = 2.1 × 10⁻¹⁵), EGFR mutations were significantly enriched in the transversion-low group (P = 3.3 × 10⁻⁶). PIK3CA and RB1 mutations were likewise enriched in transversion-low tumours (P < 0.05). Additionally, the transversion-low tumours were specifically enriched for in-frame insertions in EGFR and ERBB2 (ref. 5) and for frameshift indels in RB1 (Fig. 1b). RB1 is commonly mutated in small-cell lung carcinoma (SCLC). We found RB1 mutations in transversion-low adenocarcinomas were enriched for frameshift indels versus single nucleotide substitutions compared to SCLC (P < 0.05), suggesting a mutational mechanism in transversion-low adenocarcinoma that is probably distinct from smoking in SCLC.

Gender is correlated with mutation patterns in lung adenocarcinoma. Only a fraction of significantly mutated genes from the complete set reported in this study (Fig. 1a) were enriched in men or women (Fig. 1c). EGFR mutations were enriched in tumours from the female cohort (P = 0.03) whereas loss-of-function mutations within RBM10, an RNA-binding protein located on the X chromosome were enriched in tumours from men (P = 0.002). When examining the transversion-high group, 16 out of 21 RBM10 mutations were observed in males (P = 0.003, Fisher’s exact test).

Somatic copy number alterations were very similar to those previously reported for lung adenocarcinoma (Supplementary Fig. 5, Supplementary Table 6). Significant amplifications included NKK2-2, TERT, MDM2, KRAS, EGFR, MET, CCNE1, CCND1, TERc and MECOM (Supplementary Table 6), as previously described, 8q24 near MYC, and a novel peak containing CCND3 (Supplementary Table 6). The CDKN2A locus was the most significant deletion (Supplementary Table 6). Supplementary Table 7 summarizes molecular and clinical characteristics by sample. Low-pass whole-genome sequencing on a subset (n = 93) of the samples revealed an average of 36 gene–gene and gene–inter-gene
rearrangements per tumour. Chromothripsis occurred in six of the 93 samples (6%) (Supplementary Fig. 6, Supplementary Table 8). Low-pass whole genome sequencing-detected rearrangements appear in Supplementary Table 9.

**Description of aberrant RNA transcripts**

Gene fusions, splice site mutations or mutations in genes encoding splicing factors promote or sustain the malignant phenotype by generating aberrant RNA transcripts. Combining DNA with mRNA sequencing enabled us to catalogue aberrant RNA transcripts and, in many cases, to identify the DNA-encoded mechanism for the aberration. Seventy-five per cent of somatic mutations identified by WES were present in the RNA transcriptome when the locus in question was expressed (minimum 5×) (Supplementary Fig. 7a) similar to prior analyses. Previously identified fusions involving ALK (3/230 cases), ROS1 (4/230) and RET (2/230) (Fig. 2a, Supplementary Table 10), all occurred in transversion-low tumours (P = 1.85 × 10^{-4}, Fisher’s exact test).

**MET** activation can occur by exon 14 skipping, which results in a stabilized protein. Ten tumours had somatic MET DNA alterations with MET exon 14 skipping in RNA. In nine of these samples, a 5’ or 3’ splice site mutation or deletion was identified. MET exon 14 skipping was also found in the setting of a MET Y1003* stop codon mutation (Fig. 2b, Supplementary Fig. 8a). The codon affected by the Y1003* mutation is predicted to disrupt multiple splicing enhancer sequences, but the mechanism of skipping remains unknown in this case.

S34F mutations in U2AF1 have recently been reported in lung adenocarcinoma but their contribution to oncogenesis remains unknown. Eight samples harboured U2AF1S34F. We identified 129 splicing events strongly associated with U2AF1S34F mutation, consistent with the role of U2AF1 in 3’-splice site selection. Cassette exons and alternative 3’ splice sites were most commonly affected (Fig. 2c, Supplementary Table 11). Among these events, alternative splicing of the CTNNB1 proto-oncogene was strongly associated with U2AF1 mutations (Supplementary Fig. 8b). Thus, concurrent analysis of DNA and RNA enabled delineation of both cis and trans mechanisms governing RNA processing in lung adenocarcinoma.

**Candidate driver genes**

The receptor tyrosine kinase (RTK)/RAS/RAF pathway is frequently mutated in lung adenocarcinoma. Striking therapeutic responses are often achieved when mutant pathway components are successfully inhibited. Sixty-two per cent (143/230) of tumours harboured known activating mutations in known driver oncogenes, as defined by others. Cancer-associated mutations in KRAS (32%, n = 74), EGFR (11%, n = 26) and BRAF (7%, n = 16) were common. Additional, previously uncharacterized KRAS, EGFR and BRAF mutations were observed, but were not classified as driver oncogenes for the purposes of our analyses (see Supplementary Fig. 9a for depiction of all mutations of known and unknown significance); explaining the differing mutation frequencies in each gene between this analysis and the overall mutational analysis described above.

We also identified known activating ERBB2 in-frame insertion and point mutations (n = 5), as well as mutations in MAP2K1 (n = 2), NRAS and HRAS (n = 1 each). RNA sequencing revealed the aforementioned MET exon 14 skipping (n = 10) and fusions involving ROSI (n = 4), ALK (n = 3) and RET (n = 2). We considered these tumours collectively as oncogene-positive, as they harboured a known activating RTK/RAS/RAF pathway somatic event. DNA amplification events were not considered to be driver events before the comparisons described below.

We sought to nominate previously unrecognized genomic events that might activate this critical pathway in the 38% of samples without a RTK/RAS/RAF oncogene mutation. Tumour cellularity did not differ between oncogene-negative and oncogene-positive samples (Supplementary Fig. 9b). Analysis of copy number alterations using GISTIC identified unique focal ERBB2 and MET amplifications in the oncogene-negative subset (Fig. 3a, Supplementary Table 6); amplifications in other wild-type proto-oncogenes, including KRAS and EGFR, were not significantly different between the two groups.

We next analysed WES data independently in the oncogene-negative and oncogene-positive subsets. We found that TP53, KEAP1, NF1 and RIT1 mutations were significantly enriched in oncogene-negative tumours (P < 0.01; Fig. 3b, Supplementary Table 12). NF1 mutations have previously been reported in lung adenocarcinoma, but this is the first study, to our knowledge, capable of identifying all classes of loss-of-function mutations in known driver oncogenes, as defined by others. Candidate driver oncogenes, as defined by others.

**Figure 3** | Identification of novel candidate driver genes. a, GISTIC analysis of focal amplifications in oncogene-negative (n = 87) and oncogene-positive (n = 143) TCGA samples identifies focal gains of MET and ERBB2 that are specific to the oncogene-negative set (purple). b, TP53, KEAP1, NF1 and RIT1 mutations are significantly enriched in samples otherwise lacking oncogene mutations (adjusted P < 0.05 by Fisher’s exact test). c, Co-mutation plot of variants of known significance within the RTK/RAS/RAF pathway in lung adenocarcinoma. Not shown are the 63 tumours lacking an identifiable driver lesion. Only canonical driver events, as defined in Supplementary Fig. 9, and proposed driver events, are shown; hence not every alteration found is displayed. d, New candidate driver oncogenes (blue: 13% of cases) and known somatically activated drivers (red: 63%) that activate the RTK/RAS/RAF pathway can be found in the majority of the 230 lung adenocarcinomas.
**Recurrent alterations in key pathways**

Recurrent aberrations in multiple key pathways and processes characterize lung adenocarcinoma (Fig. 4a). Among these were RTK/RAS/RAF pathway activation (76% of cases), PI3K-mTOR pathway activation (25%), p53 pathway alteration (65%), alteration of cell cycle regulators (64%, Supplementary Fig. 10), alteration of oxidative stress pathways (22%, Supplementary Fig. 11), and mutation of various chromatin and RNA splicing factors (49%).

We then examined the phenotypic sequelae of some key genomic events in the tumours in which they occurred. Reverse-phase protein arrays provided proteomic and phosphoproteomic phenotypic evidence of pathway activity. Antibodies on this platform are listed in Supplementary Table 13. This analysis suggested that DNA sequencing did not identify all samples with phosphoprotein evidence of activation of a given signalling pathway. For example, whereas KRAS-mutant lung adenocarcinomas had higher levels of phosphorylated MAPK than KRAS wild-type tumours had on average, many KRAS wild-type tumours displayed significant MAPK pathway activation (Fig. 4b, Supplementary Fig. 10). The multiple mechanisms by which lung adenocarcinomas achieve MAPK activation suggest additional, still undetected RTK/RAS/RAF pathway alterations. Similarly, we found significant activation of mTOR and its effectors (p70S6kinase, S6, 4E-BP1) in a substantial fraction of the tumours (Fig. 4c). Analysis of mutations in PIK3CA and STK11, STK11 protein levels, and AMPK and AKT phosphorylation led to the identification of three major mTOR pathways in lung adenocarcinoma: (1) tumours with minimal or basal mTOR pathway activation, (2) tumours showing higher mTOR activity accompanied by either STK11-inactivating mutation or combined low STK11 expression and low AMPK activation, and (3) tumours showing high mTOR activity accompanied by either phosphorylated AKT activation, PIK3CA mutation, or both. As with MAPK, many tumours lack an obvious underlying genomic alteration to explain their apparent mTOR activation.

**Molecular subtypes of lung adenocarcinoma**

Broad transcriptional and epigenetic profiling can reveal downstream consequences of driver mutations, provide clinically relevant classification and offer insight into tumours lacking clear drivers. Prior unsupervised analyses of lung adenocarcinoma gene expression have used varying nomenclature for transcriptional subtypes of the disease. To coordinate naming of the transcriptional subtypes with the histopathological, anatomic and mutational classifications of lung adenocarcinoma, we propose an updated terminology: the terminal respiratory unit (TRU), formerly bronchioid, the proximal-inflammatory (PI, formerly squamoid), and the proximal-proliferative (PP, formerly magnoïd) transcriptional subtypes (Fig. 5a). Previously reported associations of expression signatures with pathways and clinical outcomes were observed (Supplementary Fig. 7b) and integration with multi-analyte data revealed statistically significant genomic alterations associated with these transcriptional subtypes. The PP subtype was enriched for mutation of KRAS, along with inactivation of the STK11 tumour suppressor gene by chromosomal loss, inactivating mutation, and reduced gene expression. In contrast, the PI subtype was characterized by solid histopathology and...
co-mutation of *NF1* and *TP53*. Finally, the TRU subtype harboured the majority of the *EGFR*-mutated tumours as well as the kinase fusion expressing tumours. TRU subtype membership was prognostically favourable, as seen previously\(^4\) (Supplementary Fig. 7c). WNT pathway overexpression was significantly associated with tumours in the proximal-inflammatory CIMP-H group (cluster 3) and a low ploidy, low mutation rate (Fig. 5c) and were significantly enriched for *SETD2* mutation, suggesting an important role for this chromatin-modifying gene in the development of certain tumours.

Integrative clustering\(^4\) of copy number, DNA methylation and mRNA expression data found six clusters (Fig. 5c). Tumour ploidy and mutation rate are higher in clusters 1–3 than in clusters 4–6. Clusters 1–3 frequently harbour *TP53* mutations and are enriched for the two proximal transcriptional subtypes. Fisher’s combined probability tests revealed significant copy number associated gene expression changes on 3q in cluster one, 8q in cluster two, and chromosomes 7 and 15q in cluster three (Supplementary Fig. 15). The low ploidy and low mutation rate clusters four and five contain many TRU samples, whereas tumours in cluster 6 have comparatively lower tumour cellularity, and few other distinguishing molecular features. Significant copy number-associated gene expression changes are observed on 6q in cluster four and 19p in cluster five. The CIMP-H tumours divided into a high ploidy, high mutation rate, proximal-inflammatory CIMP-H group (cluster 3) and a low ploidy, low mutation rate, TRU-associated CIMP-H group (cluster 4), suggesting that the CIMP phenotype in lung adenocarcinoma can occur in markedly different genomic and transcriptional contexts. Furthermore, cluster four is enriched for *CDKN2A* methylation and *SETD2* mutations, suggesting an interaction between somatic mutation of *SETD2* and deregulated chromatin maintenance in this subtype. Finally, cluster membership was significantly associated with mutations in *TP53*, *EGFR* and *STK11* (Supplementary Fig. 15, Supplementary Table 6).

## Conclusions

We assessed the mutation profiles, structural rearrangements, copy number alterations, DNA methylation, mRNA, miRNA and protein expression...
of 230 lung adenocarcinomas. In recent years, the treatment of lung adenocarcinoma has been advanced by the development of multiple therapies targeted against alterations in the RTK/RAS/RAF pathway. We nominate amplifications in MET and ERBB2 as well as mutations of NF1 and RIT1 as driver events specifically in otherwise oncogene-negative lung adenocarcinomas. This analysis increases the fraction of lung adenocarcinoma cases with somatic evidence of RTK/RAS/RAF activation from 62% to 76%. While all lung adenocarcinomas may activate this pathway by some mechanism, only a subset show tonic pathway activation at the protein level, suggesting both diversity between tumours with seemingly similar activating events and as yet undescribed mechanisms of pathway activation. Therefore, the current study expands the range of possible targetable alterations within the RTK/RAS/RAF pathway in general and suggests increased implementation of MET and ERBB2/HER2 inhibitors in particular. Our discovery of inactivating mutations of MGA further underscores the importance of the MYC pathway in lung adenocarcinoma.

This study further implicates both chromatin modifications and splicing alterations in lung adenocarcinoma through the integration of DNA, transcriptome and methylyme analysis. We identified alternative splicing due to both splicing factor mutations in trans and mutation of splice sites in cis, the latter leading to activation of the MET gene by exon 14 skipping. Cluster analysis separated tumours based on single-gene driver events as well as large-scale aberrations, emphasizing lung adenocarcinoma's molecular heterogeneity and combinatorial alterations, including the identification of coincident SETD2 mutations and CDKN2A methylation in a subset of CIMP-H tumours, providing evidence of a somatic event associated with a genome-wide methylation phenotype. These studies provide new knowledge by illuminating modes of genomic alteration, highlighting previously unappreciated altered genes, and enabling further refinement in sub-classification for the improved personalization of treatment for this deadly disease.

METHODS SUMMARY

All specimens were obtained from patients with appropriate consent from the relevant institutional review board. DNA and RNA were collected from samples using the AllPrep kit (Qiagen). We used standard approaches for capture and sequencing of exomes from tumour DNA and normal DNA and whole-genome shotgun sequencing. Significantly mutated genes were identified by comparing them with expectation models based on the exact measured rates of specific sequence lesions. GISTIC analysis of the circular-binary-segmented Affymetrix SNP 6.0 copy number data was used to identify recurrent amplification and deletion peaks. Consensus clustering approaches were used to analyse mRNA, miRNA and methylation subtypes using previous approaches. The publication web page is (https://tcga-data.nci.nih.gov/docs/publications/luad_2014/). Sequence files are in CGHub (https://cghub.ucsc.edu/).

Received 11 June 2013; accepted 22 April 2014.

Published online 9 July 2014.
The Cancer Genome Atlas Research Network

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 coördinating centre. The National Cancer Institute and National Human Genome Research Institute project teams coordinated project activities. We also acknowledge the following TCGA investigators who made substantial contributions: E. A. Collisson (manuscript coordinator); J. D. Campbell, J. Chmielecki, (analysis coordinators); C. Sougnez (data coordinator); J. D. Campbell, M. Rosenberg, W. Lee, J. Chmielecki, M. Ladanyi, and G. Getz (DNA sequence analysis); M. D. Wilkerson, A. N. Brooks, and D. N. Hayes (miRNA sequence analysis); L. Danila and L. Cope (DNA methylation analysis); A. D. Cherniack (copy number analysis); M. D. Wilkerson and A. Hajdujajiany (translocations); N. Schultz, W. Lee, E. A. Collisson, A. H. Berger, J. Chmielecki, C. J. Creighton, L. A. Byers and M. Ladanyi (pathway analyses); A. Chiu and A. G. Romanov (sequence analysis); W. Lee and A. D. Wiggle (pathology and clinical expert); L. A. Byers and G. B. Mills (reverse phase protein arrays); S. B. Baylin, R. Govindan and M. Meyerson (project chairs).

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 data files are deposited in pub and all other data are released at the Data Coordinating Center (DCC) for public access (http://cancergenome.nih.gov/). (https://cgbhu.ucsc.edu/) and (https://tcga-data.nci.nih.gov/docs/publications/luad_2014/). 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 author manuscript. Correspondence requests for materials should be addressed to M.M. (mattbay@dfci.harvard.edu).

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported licence. The images or other material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons licence, users will need to obtain permission from the licence holder to reproduce the material. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-sa/3.0/
Corrigendum: Comprehensive molecular profiling of lung adenocarcinoma
The Cancer Genome Atlas Research Network

Nature 511, 543–550 (2014); doi:10.1038/nature13385

In this Article, the surname of author Kristen Rodgers was incorrectly spelled Rogers. This error has been corrected in the HTML and PDF of the original paper.
Author Correction: Comprehensive molecular profiling of lung adenocarcinoma

The Cancer Genome Atlas Research Network

Correction to: Nature https://doi.org/10.1038/nature13385, published online 9 July 2014; corrected online 8 October 2014.

In this Article, the Supplementary Table 7 iCLUSTER output column included incorrect cluster labels for the integrated subtypes presented in Fig. 5c. These changes affect only the iCLUSTER output column and do not affect the analysis or the conclusions of the work. The authors apologise for the error. Supplementary Table 7 has been corrected online, and the original incorrect table is provided as Supplementary Information to this Amendment for transparency.

Supplementary Information is available in the online version of this Amendment.