ARTICLE

doi:10.1038/nature11404

Comprehensive genomic characterization of squamous cell lung cancers

The Cancer Genome Atlas Research Network*

Lung squamous cell carcinoma is a common type of lung cancer, causing approximately 400,000 deaths per year worldwide. Genomic alterations in squamous cell lung cancers have not been comprehensively characterized, and no molecularly targeted agents have been specifically developed for its treatment. As part of The Cancer Genome Atlas, here we profile 178 lung squamous cell carcinomas to provide a comprehensive landscape of genomic and epigenomic alterations. We show that the tumour type is characterized by complex genomic alterations, with a mean of 360 exonic mutations, 165 genomic rearrangements, and 323 segments of copy number alteration per tumour. We find statistically recurrent mutations in 11 genes, including mutation of *TP53* in nearly all specimens. Previously unreported loss-of-function mutations are seen in the *HLA-A* class I major histocompatibility gene. Significantly altered pathways included *NFE2L2* and *KEAP1* in 34%, squamous differentiation genes in 44%, phosphatidylinositol-3-OH kinase pathway genes in 47%, and *CDKN2A* and *RB1* in 72% of tumours. We identified a potential therapeutic target in most tumours, offering new avenues of investigation for the treatment of squamous cell lung cancers.

Lung cancer is the leading cause of cancer-related mortality worldwide, leading to an estimated 1.4 million deaths in 2010 (ref. 1). The discovery of recurrent mutations in the epidermal growth factor receptor (*EGFR*) kinase, as well as fusions involving anaplastic lymphoma kinase (*ALK*), has led to a marked change in the treatment of patients with lung adenocarcinoma, the most common type of lung cancer[2–5]. More recent data have suggested that targeting mutations in *BRAF*, *AKT1*, *ERBB2* and *PIK3CA* and fusions that involve *ROS1* and *RET* may also be successful[6,7]. Unfortunately, activating mutations in *EGFR* and *ALK* fusions are typically not present in the second most common type of lung cancer, lung squamous cell carcinoma (SQCC)[8], and targeted agents developed for lung adenocarcinoma are largely ineffective against lung SQCC.

Although no comprehensive genomic analysis of lung SQCCs has been reported, single-platform studies have identified regions of somatic copy number alterations in lung SQCCs, including amplification of *SOX2*, *PDGFRA* and *FGFR1* and/or *WHSC1L1* and deletion of *CDKN2A*[9,10]. DNA sequencing studies of lung SQCCs have reported recurrent mutations in several genes, including *TP53*, *NFE2L2*, *KEAP1*, *BAI3*, *FBXW7*, *GRM8*, *MUC16*, *RUNX1T1*, *STK11* and *ERBB4* (refs 11, 12). *DDR2* mutations and *FGFR1* amplification have been nominated as therapeutic targets[13–15].

We have conducted a comprehensive study of lung SQCCs from a large cohort of patients as part of The Cancer Genome Atlas (TCGA) project. The twin aims are to characterize the genomic and epigenomic landscape of lung SQCC and to identify potential opportunities for therapy. We report an integrated analysis based on DNA copy number, somatic exonic mutations, messenger RNA sequencing, mRNA expression and promoter methylation for 178 histopathologically reviewed lung SQCCs, in addition to whole genome sequencing (WGS) of 19 samples and microRNA sequencing of 159 samples (Supplementary Table 1.1). Demographic and clinical data and results of the genomic analyses can be downloaded from the TCGA data portal (https://tcga-data.nci.nih.gov/docs/publications/lusc_2012/).

Samples and clinical data

Tumour samples were obtained from 178 patients with previously untreated stage I–IV lung SQCC. Germline DNA was obtained from adjacent, histologically normal tissues resected at the time of surgery ($n = 137$) or from peripheral blood ($n = 41$). All patients provided written informed consent to conduct genomic studies in accordance with local Institutional Review Boards. The demographic characteristics are described in Supplementary Table 1.2. The median follow-up for the cohort was 15.8 months, and 60% of patients were alive at the time of the last follow-up (data updated in November 2011). Ninety-six per cent of the patients had a history of tobacco use, similar to previous reports for North American patients with lung SQCC[16]. DNA and RNA were extracted from patient specimens and measured by several genomic assays, which included standard quality-control assessments (Supplementary Methods, sections 2–8). A committee of experts in lung cancer pathology performed a further review of all samples to confirm the histological subtype (Supplementary Fig. 1.1 and Supplementary Methods, section 1).

Somatic DNA alterations

The lung SQCCs analysed in this study display a large number and variety of DNA alterations, with a mean of 360 exonic mutations, 323 altered copy number segments and 165 genomic rearrangements per tumour.

Copy number alterations were analysed using several platforms. Analysis of single nucleotide polymorphism (SNP) 6.0 array data across the set of 178 lung SQCCs identified a high rate of copy number alteration (mean of 323 segments) when compared with other TCGA projects (as of 1 February 2012), including ovarian cancer (477 segments)[17], glioblastoma multiforme (282 segments)[18], colorectal carcinoma (213 segments), breast carcinoma (282 segments) and renal cell carcinoma (156 segments) ($P < 1 \times 10^{-15}$ by Fisher's exact test). These segments gave rise to regions of both focal and broad somatic copy number alterations (SCNAs), with a mean of 47 focal and 23 broad events per tumour (broad events defined as ≥50% of the length of the chromosome arm). There was strong concordance between the three independent copy number assays for all regions of SCNA (Supplementary Figs 2.1–2.4).

At the level of whole chromosome arm SCNAs, lung SQCCs exhibit many similarities to 205 cases of lung adenocarcinoma analysed by

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

TCGA (Supplementary Fig. 2.1a). The most notable difference between these cancers is selective amplification of chromosome 3q in lung SQCC, as has been reported[9,19]. Using the SNP 6.0 array platform and GISTIC 2.0 (refs 20, 21), we identified regions of significant copy number alteration (Supplementary Methods, section 2). There were 50 peaks of significant amplification or deletion ($Q < 0.05$), several of which included SCNAs previously seen in lung SQCCs including *SOX2*, *PDGFRA* and/or *KIT*, *EGFR*, *FGFR1* and/or *WHSC1L1*, *CCND1* and *CDKN2A*[9,10,19] (Supplementary Fig. 2.1b and Supplementary Data 2.1 and 2.2). Other peaks defined regions of SCNA reported for the first time, including amplifications of chromosomal segments containing *NFE2L2*, *MYC*, *CDK6*, *MDM2*, *BCL2L1* and *EYS* and deletions of *FOXP1*, *PTEN* and *NF1* (Supplementary Fig. 2.1b).

Whole exome sequencing of 178 lung SQCCs and matched germline DNA targeted 193,094 exons from 18,863 genes. The mean sequencing coverage across targeted bases was 121×, with 83% of target bases above 30× coverage. We identified a total of 48,690 non-silent mutations with a mean of 228 non-silent and 360 total exonic mutations per tumour, corresponding to a mean somatic mutation rate of 8.1 mutations per megabase (Mb) and median of 8.4 per Mb. That rate is higher than rates observed in other TCGA projects including acute myelogenous leukaemia (0.56 per Mb), breast carcinoma (1.0 per Mb), ovarian cancer[17] (2.1 per Mb), glioblastoma multiforme[18] (2.3 per Mb) and colorectal carcinoma (3.2 per Mb) (data as of 1 February 2012, $P < 2.2 \times 10^{-16}$ by *t*-test or Wilcoxon's rank sum test for lung SQCC versus all others). In lung SQCC, CpG transitions and transversions were the most commonly observed mutation types, with mean rates of 9.9 and 10.7 per sequenced megabase of CpG context, respectively, for a total mutation rate of 20.6 per Mb. At non-CpG sites, transversions at C:G sites were more common than transitions (7.3 versus 2.9 per Mb; total = 10.2 per Mb) and more common than transversions or transitions at A:T sites (1.5 versus 1.3 per Mb; total = 2.8 per Mb).

Significantly mutated genes were identified using a modified version of the MutSig algorithm (Supplementary Methods, section 3)[22,23]. We identified 10 genes with a false discovery rate (FDR) *Q* value < 0.1 (Supplementary Table 3.1): *TP53*, *CDKN2A*, *PTEN*, *PIK3CA*, *KEAP1*, *MLL2*, *HLA-A*, *NFE2L2*, *NOTCH1* and *RB1*, all of which demonstrated robust evidence of gene expression as defined by reads per kilobase of exon model per million mapped reads (RPKM) > 1 (Fig. 1). *TP53* mutation was observed in 81% of samples by automated analysis; visual review of sequencing reads identified a further 9% of samples with potential mutations in regions of sub-optimal coverage or in samples with low purity. Most observed mutations in *NOTCH1* (8 out of 17) were truncating alterations, suggesting loss-of-function, as has recently been reported for head and neck SQCCs[22,24]. Mutations in *HLA-A* were also almost exclusively nonsense or splice site events (7 out of 8).

To increase our statistical power to detect mutated genes in the setting of the observed high background mutation rate, we performed a secondary MutSig analysis only considering genes previously observed to be mutated in cancer according to the COSMIC database. This yielded 12 other genes with FDR < 0.1: *FAM123B* (also known as *WTX*), *HRAS*, *FBXW7*, *SMARCA4*, *NF1*, *SMAD4*, *EGFR*, *APC*, *TSC1*, *BRAF*, *TNFAIP3* and *CREBBP* (Supplementary Table 3.1). Both the spectrum and the frequency of *EGFR* mutations differed from those seen in lung adenocarcinomas. The two most common alterations in lung adenocarcinoma, Leu858Arg and inframe deletions in exon 19, were absent, whereas two Leu861Gln mutations were detected in *EGFR*.

As described in Supplementary Fig. 3.1, we verified somatic mutations by performing an independent hybrid-recapture of 76 genes in all samples. A total of 1,289 mutations were assayed, and we achieved satisfactory coverage to have power to verify at 1,283 positions. We validated 1,235 mutations (96.2%) (Supplementary Fig. 3.1 and Supplementary Methods, section 3). We also verified mutation calls using WGS and RNA sequencing data with similar results (Supplementary Figs 3.1, 4.3 and Supplementary Methods, sections 3 and 4).

WGS was performed for 19 tumour/normal pairs with a mean computed coverage of 54×. A mean of 165 somatic rearrangements was found per lung SQCC tumour pair (Supplementary Fig. 3.2), a value in excess of that reported for WGS studies of other tumour types including colorectal carcinoma (75)[25], prostate carcinoma (108)[26], multiple myeloma (21)[23] and breast cancer (90)[27]. Although most inframe coding fusions detected in WGS were validated by RNA sequencing, no recurrent rearrangements predicted to generate fusion proteins were identified (Supplementary Data 3.1 and 4.1).

Somatically altered pathways

Many of the somatic alterations we have identified in lung SQCCs seem to be drivers of pathways important to the initiation or progression of the cancer. Specifically, genes involved in the oxidative stress response and squamous differentiation were frequently altered by mutation or SCNA. We observed mutations and copy number alterations of *NFE2L2* and *KEAP1* and/or deletion or mutation of *CUL3* in 34% of cases (Fig. 2). *NFE2L2* and *KEAP1* code for proteins that bind to each other, have been shown to regulate the cell response to oxidative damage, chemo- and radiotherapy, and are somatically altered in a variety of cancer types[28,29]. We found mutations in *NFE2L2* almost exclusively in one of two KEAP1 interaction motifs, DLG or ETGE. Mutations in *KEAP1* and *CUL3* showed a pattern consistent with loss-of-function and were mutually exclusive with mutations in *NFE2L2* (Figs 1c and 2). PARADIGM SHIFT[30] analysis predicts that mutations in *NFE2L2* and *KEAP1* exert a considerable functional effect (Supplementary Fig. 7.C.1, 7.C.2 and Supplementary Methods, section 7).

We also found alterations in genes with known roles in squamous cell differentiation in 44% of samples, including overexpression and amplification of *SOX2* and *TP63*, loss-of-function mutations in *NOTCH1*, *NOTCH2* and *ASCL4* and focal deletions in *FOXP1* (Fig. 2). Although *NOTCH1* has been well characterized as an oncogene in haematological cancers[31], *NOTCH1* and *NOTCH2* truncating mutations have been reported in cutaneous SQCCs and lung SQCCs[32]. Truncating mutations in *ASCL4* are the first to be reported in human

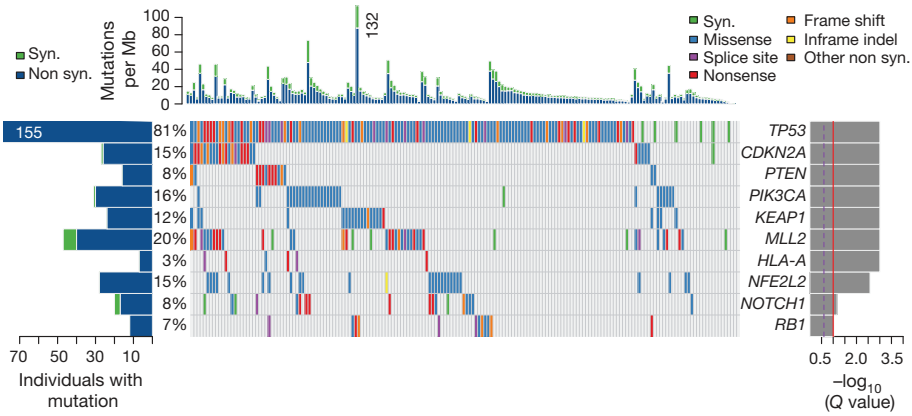

Figure 1 | Significantly mutated genes in lung SQCC. Significantly mutated genes (Q value < 0.1) identified by exome sequencing are listed vertically by Q value. The percentage of lung SQCC samples with a mutation detected by automated calling is noted at the left. Samples displayed as columns, with the overall number of mutations plotted at the top, and samples are arranged to emphasize mutual exclusivity among mutations. Syn., synonymous.

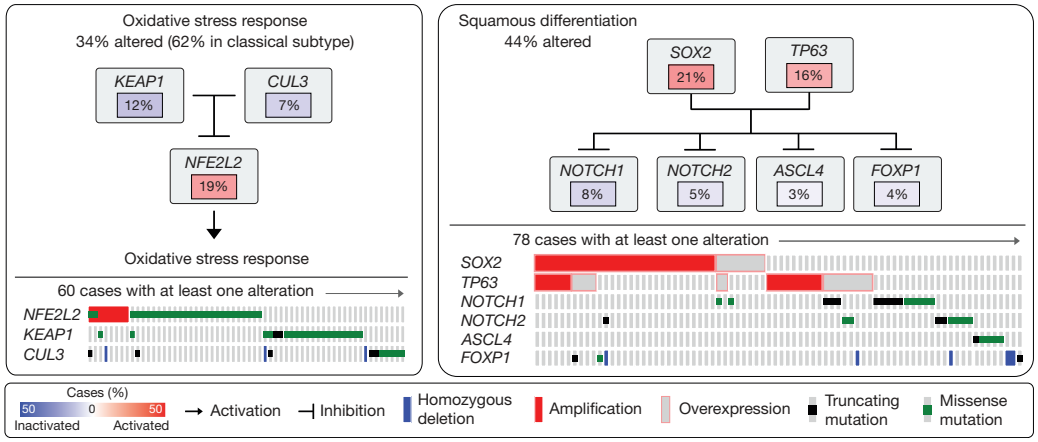

Figure 2 | Somatically altered pathways in squamous cell lung cancer. Left, alterations in oxidative stress response pathway genes as defined by somatic mutation, copy number alteration or up- or downregulation. Frequencies of alteration are expressed as a percentage of all cases, with background in red for activated genes and blue for inactivated genes. Right, alterations in genes that regulate squamous differentiation, as defined in the left panel.

cancer and may have a lineage role given the requirement for *ASCL1* for survival of small-cell lung cancer cells[33]. Alterations in *NOTCH1*, *NOTCH2* and *ASCL4* were mutually exclusive and exhibited minimal overlap with amplification of *TP63* and/or *SOX2* (Fig. 2), suggesting that aberrations in those modulators of squamous cell differentiation have overlapping functional consequences.

mRNA expression profiling and subtype classification

Whole-transcriptome expression profiles were generated by RNA sequencing for the entire cohort and by microarrays for a 121-sample subset. Of 20,502 genes analysed, the mean RNA coverage indices were 19× and 6,420 RPKM (Supplementary Fig. 4.1 and Supplementary Methods, section 4). Previously reported lung SQCC gene expression-subtype signatures[34] were applied to both of the expression platforms, yielding four subtypes designated as classical (36%), basal (25%), secretory (24%) and primitive (15%). The concordance of subtypes between the two platforms was high (94% agreement) (Supplementary Fig. 4.2). Considerable correlations were found between the expression subtypes and genomic alterations in copy number, mutation and methylation (Fig. 3). The classical subtype was characterized by alterations in *KEAP1*, *NFE2L2* and *PTEN*, as well as pronounced hypermethylation and chromosomal instability. The 3q26 amplicon was present in all of the subtypes, but it was most characteristic of the classical subtype, which also showed the greatest overexpression of three known oncogenes on 3q: *SOX2*, *TP63* and *PIK3CA*. RNA sequencing data suggested that high expression levels of *TP63*, in samples with and without amplification of *TP63*, were associated with dominant expression of the deltaN isoform (also called p40), which lacks the amino-terminal transactivation domain, compared with the longer isoform, called tap63 (89% of tumours overexpressed deltaN compared with tap63; $P < 2.2 \times 10^{-16}$). The short deltaN isoform is thought to function as an oncogene[35,36], and its expression was most enriched in the classical subtype. By contrast, the primitive expression subtype more commonly exhibited *RB1* and *PTEN* alterations, and the basal expression subtype showed *NF1* alterations (Fig. 3). Amplification of *FGFR1* and *WHSC1L1* was anticorrelated with the classical subtype and specifically with *NFE2L2* or *KEAP1* mutated samples. Although *CDKN2A* alterations are common in lung SQCCs, they are not associated with any particular expression subtype (Fig. 3).

Independent clustering of miRNA and methylation data indicated association with expression subtypes. The highest overall methylation was seen in the classical subtype (Fig. 3, Supplementary Figs 5.1 and 6.1, Supplementary Methods, sections 5 and 6, Supplementary Data 6.1 and 6.2 and Supplementary Table 5.1). Integrative clustering (iCluster)[37] of mRNA, miRNA, methylation, SCNA and mutation data demonstrated concordance with the mRNA expression subtypes and associated alterations (Fig. 3, Supplementary Fig. 7.A.1 and Supplementary Methods, section 7). Independent correlation of somatic mutations, copy number alterations and gene expression signatures revealed notable subtype associations with alterations in the *TP53*, *PI3K*, *RB1* and *NFE2L2/KEAP1* pathways (Supplementary Fig. 7.B.1 and Supplementary Methods, section 7).

Analysis of the *CDKN2A* locus

Integrated multiplatform analyses showed that *CDKN2A*, a known tumour suppressor gene in lung SQCC[38] that encodes the $p16^{INK4A}$ and $p14^{ARF}$ proteins, is inactivated in 72% of cases of lung SQCC (Fig. 4a and Supplementary Data 7.1)—by epigenetic silencing by methylation (21%), inactivating mutation (18%), exon 1β skipping (4%) and homozygous deletion (29%).

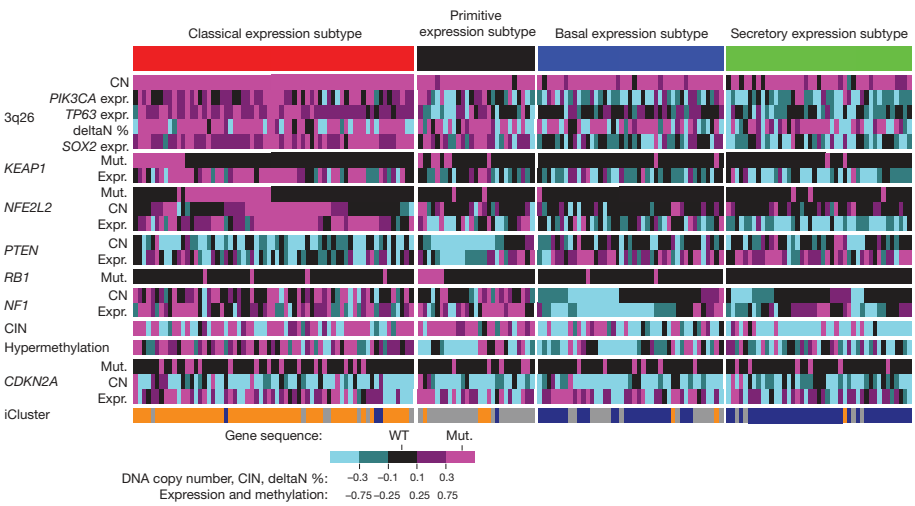

Figure 3 | Gene expression subtypes integrated with genomic alterations. Tumours are displayed as columns, grouped by gene expression subtype. Subtypes were compared by Kruskal–Wallis tests for continuous features and by Fisher's exact tests for categorical features. Displayed features showed significant association with gene expression subtype ($P < 0.05$), except for *CDKN2A* alterations. deltaN percentage represents transcript isoform usage between the *TP63* isoforms, deltaN and tap63, as determined by RNA sequencing. Chromosomal instability (CIN) is defined by the mean of the absolute values of chromosome arm copy numbers (CN) from the GISTIC[23,24] output. Absolute values are used so that amplification and deletion alterations are counted equally. Hypermethylation scores and iCluster assignments are described in Supplementary Figs 6.1 and 7.A1, respectively. CIN, methylation, gene expression and deltaN values were standardized for display using *z*-score transformation. Expr., expression; mut., mutation; WT, wild type.

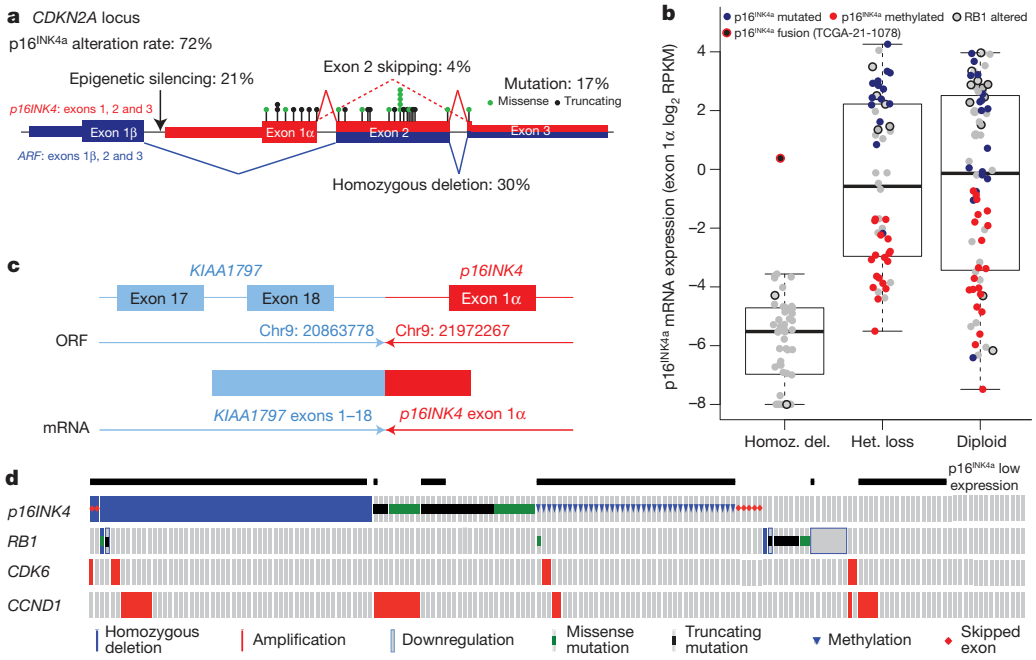

Figure 4 | Multi-faceted characterization of mechanisms of *CDKN2A* loss. **a**, Schematic view of the exon structure of *CDKN2A* demonstrating the types of alterations identified in the study. The locations of point mutation are denoted by black and green circles. **b**, *CDKN2A* expression (*y* axis) versus *CDKN2A* copy number (*x* axis). Samples are represented by circles and colour-coded by specific type of *CDKN2A* alteration. Del., deletion; het., heterozygous; homoz., homozygous. **c**, Diagram of the *KIAA1797-p16INK4* fusion identified by WGS. ORF, open reading frame. **d**, *CDKN2A* alterations and expression levels (binary) in each sample.

Analysis of mRNA expression across the *CDKN2A* locus revealed four distinct patterns of expression: complete absence of both *p16INK4* and *ARF* (33%); expression of high levels of both *p16INK4* and *ARF* (31%); high expression of *ARF* and absence of *p16INK4* (31%); or expression of a transcript that represents a splicing of exon 1β from *ARF* with the shared exon 3 of *ARF* and *p16INK4*, generating a premature stop codon (4%) (Supplementary Fig. 4.4). Almost all of the cases completely lacking *p16INK4* and *ARF* expression showed homozygous deletion (Fig. 4b and Supplementary Data 7.1). In one case, *p16INK4* expression was detected but analysis of WGS data demonstrated an intergenic fusion event that resulted in detectable transcription between exon 1α *p16INK4* and exon 18 of *KIAA1797* (Fig. 4b, c). Interestingly, combined analysis of WGS and RNA sequencing data identified tumour suppressor gene inactivation by intra- or interchromosomal rearrangement in *PTEN*, *NOTCH1*, *ARID1A*, *CTNNA2*, *VHL* and *NF1*, in eight further cases (Supplementary Data 3.1 and 4.1).

In addition to homozygous deletion, there are frequent mutational events in *CDKN2A* (Fig. 4b and Supplementary Data 7.1). These account for 45% of the 56 cases with high *p16INK4* and *ARF* expression. Furthermore, methylation of the exon 1α promoter accounts for many other cases of *CDKN2A* inactivation (70% of lung SQCCs with *ARF* expression in the absence of detectable *p16INK4*). Seven other tumours in the high-*ARF*/low-*INK4A* group had documented mutations of *INK4A*, primarily nonsense mutations, suggesting nonsense-mediated decay as a mechanism. Of the 28% of tumours without *CDKN2A* alterations, *RB1* mutations were identified in eight cases and *CDK6* amplification in one case (Fig. 4d).

Therapeutic targets

Molecularly targeted agents are now commonly used in patients with adenocarcinoma of the lung, whereas no effective targeted agents have been developed specifically for lung SQCCs[13]. We analysed our genomic data for evidence of the two common genomic alterations in adenocarcinomas of the lung: *EGFR* and *KRAS* mutations. Only one sample had a *KRAS* codon 61 mutation, and there were no exon 19 deletions or Leu858Arg mutations in *EGFR*. However, amplifications of *EGFR* were found in 7% of cases, as were two instances of the Leu861Gln *EGFR* mutation, which confers sensitivity to erlotinib and gefitinib[39].

The presence of new potential therapeutic targets in lung SQCC was suggested by the observation that 96% (171 out of 178) of tumours contain one or more mutations in tyrosine kinases, serine/threonine kinases, phosphatidylinositol-3-OH kinase (PI(3)K) catalytic and regulatory subunits, nuclear hormone receptors, G-protein-coupled receptors, proteases and tyrosine phosphatases (Supplementary Fig. 7.D.1a and Supplementary Data 7.2 and 7.3). From 50 to 77% of the mutations were predicted to have a medium or high functional effect as determined by the mutation assessor score[40] (Supplementary Fig. 7.D.1a), and 39% of tyrosine and 42% of serine/threonine kinase mutations were located in the kinase domain. Many of the alterations were in known oncogenes and tumour suppressors, as defined in the COSMIC database (Supplementary Data 7.3).

We selected potential therapeutic targets based on several features, including (1) availability of a US Food and Drug Administration (FDA)-approved targeted therapeutic agent or one under study in current clinical trials (Supplementary Data 7.2); (2) confirmation of the altered allele in RNA sequencing; and (3) the mutation assessor score[40]. Using those criteria, we identified 114 cases with somatic alteration of a potentially targetable gene (64%) (Supplementary Fig. 7.D.1b and Supplementary Data 7.4). Among these, we identified three families of tyrosine kinases, the erythroblastic leukaemia viral oncogene homologues (ERBBs), fibroblast growth factor receptors (FGFRs) and Janus kinases (JAKs), all of which were found to be mutated and/or amplified[41]. As discussed for *EGFR*, the mutational spectra in these potential therapeutic targets differed from those in lung adenocarcinoma (Supplementary Fig. 7.D.2)[42].

To complement a gene-centred search for potential therapeutic targets, we analysed core cellular pathways known to represent potential therapeutic vulnerabilities: PI(3)K/AKT, receptor tyrosine kinase (RTK) and RAS. Analysis of the 178 lung SQCCs revealed alteration in at least one of those pathways in 69% of samples after restriction of the analysis to mutations confirmed by RNA sequencing and to amplifications associated with overexpression of the target gene (Fig. 5). Mutational events that have been curated in COSMIC are also shown in Supplementary Fig. 7D.2, as is the distribution of mutations, amplifications and overexpression of the genes depicted in Fig. 5. (A summary of all samples and their significant mutations and copy number alterations, including alterations in Fig. 5, is shown in Supplementary Data 7.5.) Specifically, one of the components of the PI(3)K/AKT pathway was altered in 47% of tumours and RTK signalling probably affected by events such as *EGFR* amplification, *BRAF* mutation or *FGFR* amplification or mutation in 26% of tumours

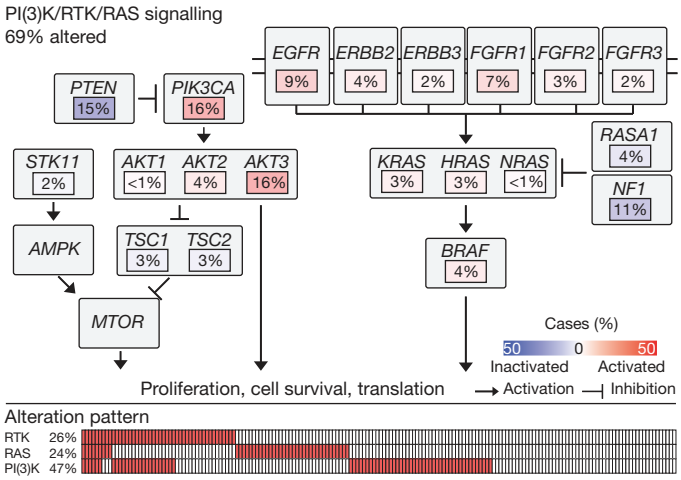

Figure 5 | Alterations in targetable oncogenic pathways in lung SQCCs. Pathway diagram showing the percentage of samples with alterations in the PI(3)K/RTK/RAS pathways. Alterations are defined by somatic mutations, homozygous deletions, high-level, focal amplifications, and, in some cases, by significant up- or downregulation of gene expression (*AKT3*, *FGFR1*, *PTEN*).

(Fig. 5 and Supplementary Fig. 7.D.3). Alterations in the PI(3)K/AKT pathway genes were mutually exclusive with *EGFR* alterations as identified by MEMo[43] (Supplementary Fig. 7.D.4.). Although the dependence of lung SQCC on many of these individual alterations remains to be defined functionally, this analysis suggests new areas for potential therapeutic development in this cancer.

Discussion

Lung SQCCs are characterized by a high overall mutation rate of 8.1 mutations per megabase and marked genomic complexity. Similar to high-grade serous ovarian carcinoma[17], almost all lung SQCCs display somatic mutation of *TP53*. There were also frequent alterations in the following pathways: *CDKN2A/RB1*, *NFE2L2/KEAP1/CUL3*, *PI3K/AKT* and *SOX2/TP63/NOTCH1* pathways, providing evidence of common dysfunction in cell cycle control, response to oxidative stress, apoptotic signalling and/or squamous cell differentiation. Pathway alterations clustered according to expression-subtype in many cases, suggesting that those subtypes have a biological basis.

A role for somatic mutation in the cancer hallmark of avoiding immune destruction[44] is suggested by the presence of inactivating mutations in the *HLA-A* gene. Somatic loss-of-function alterations of *HLA-A* have not been reported previously in genomic studies of lung cancer. Given the recently reported efficacy of anti-programmed death 1 (PD1)[45] and anti-cytotoxic T-lymphocyte antigen 4 (CTLA4) antibodies in non-small-cell lung cancer[46], these *HLA-A* mutations suggest a possible role for genotypic selection of patients for immunotherapies.

Targeted kinase inhibitors have been successfully used for the treatment of lung adenocarcinoma but minimally so in lung SQCC. The observations reported here suggest that a detailed understanding of the possible targets in lung SQCCs may identify targeted therapeutic approaches. Whereas *EGFR* and *KRAS* mutations, the two most common oncogenic aberrations in lung adenocarcinoma, are extremely rare in lung SQCC, alterations in the FGFR kinase family are common. Lung SQCCs also share many alterations in common with head and neck squamous cell carcinomas without evidence of human papilloma virus infection, including mutation in *PIK3CA*, *PTEN*, *TP53*, *CDKN2A*, *NOTCH1* and *HRAS*[22,24], suggesting that the biology of these two diseases may be similar.

The current study has identified a potentially targetable gene or pathway alteration in most lung SQCC samples studied. The data presented here can help to organize efforts to analyse lung SQCC clinical tumour specimens for a panel of specific, actionable mutations to select patients for appropriately targeted clinical trials. These data could thereby help to facilitate effective personalized therapy 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 commercial technology 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[23,24] 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[20,21,34,38,41,44].

Received 9 March; accepted 9 July 2012.
Published online 9 September 2012.

1. World Health Organization. Cancer, fact sheet no. 297 ⟨http://www.who.int/mediacentre/factsheets/fs297/en/⟩ (accessed February 2012).
2. Soda, M. *et al.* Identification of the transforming *EML4–ALK* fusion gene in non-small-cell lung cancer. *Nature* **448,** 561–566 (2007).
3. Paez, J. G. *et al.* EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* **304,** 1497–1500 (2004).
4. Lynch, T. J. *et al.* Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* **350,** 2129–2139 (2004).
5. Pao, W. *et al.* EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc. Natl Acad. Sci. USA* **101,** 13306–13311 (2004).
6. Felip, E., Gridelli, C., Baas, P., Rosell, R. & Stahel, R. Metastatic non-small-cell lung cancer: consensus on pathology and molecular tests, first-line, second-line, and third-line therapy. *Ann. Oncol.* **22,** 1507–1519 (2011).
7. Ju, Y. S. *et al.* A transforming *KIF5B* and *RET* gene fusion in lung adenocarcinoma revealed from whole-genome and transcriptome sequencing. *Genome Res.* **22,** 436–445 (2012).
8. Rekhtman, N. *et al.* Clarifying the spectrum of driver oncogene mutations in biomarker-verified squamous carcinoma of lung: lack of EGFR/KRAS and presence of PIK3CA/AKT1 mutations. *Clin. Cancer. Res.* **18,** 1167–1176 (2012).
9. Bass, A. J. *et al. SOX2* is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. *Nature Genet.* **41,** 1238–1242 (2009).
10. Ramos, A. H. *et al.* Amplification of chromosomal segment 4q12 in non-small cell lung cancer. *Cancer Biol. Ther.* **8,** 2042–2050 (2009).
11. 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).
12. Kan, Z. *et al.* Diverse somatic mutation patterns and pathway alterations in human cancers. *Nature* **466,** 869–873 (2010).
13. Hammerman, P. S., Sos, M. L., Ramos, A. & Xu, C. Mutations in the *DDR2* kinase gene identify a novel therapeutic target in squamous cell lung cancer. *Cancer Discovery* **1,** 78 (2011).
14. Weiss, J. *et al.* Frequent and focal *FGFR1* amplification associates with therapeutically tractable FGFR1 dependency in squamous cell lung cancer. *Sci. Transl. Med.* **2,** 62ra93 (2010).
15. Dutt, A. *et al.* Inhibitor-sensitive *FGFR1* amplification in human non-small cell lung cancer. *PLoS One* **6,** e20351 (2011).
16. Kenfield, S. A., Wei, E. K., Stampfer, M. J., Rosner, B. A. & Colditz, G. A. Comparison of aspects of smoking among the four histological types of lung cancer. *Tob. Control* **17,** 198–204 (2008).
17. The Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature* **474,** 609–615 (2011).
18. The Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **455,** 1061–1068 (2008).
19. Tonon, G. *et al.* High-resolution genomic profiles of human lung cancer. *Proc. Natl Acad. Sci. USA* **102,** 9625–9630 (2005).
20. Beroukhim, R. *et al.* The landscape of somatic copy-number alteration across human cancers. *Nature* **463,** 899–905 (2010).
21. Mermel, C. H. *et al.* GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* **12,** R41 (2011).
22. Stransky, N. *et al.* The mutational landscape of head and neck squamous cell carcinoma. *Science* **333,** 1157–1160 (2011).
23. Chapman, M. A. *et al.* Initial genome sequencing and analysis of multiple myeloma. *Nature* **471,** 467–472 (2011).
24. Agrawal, N. *et al.* Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in *NOTCH1*. *Science* **333,** 1154–1157 (2011).
25. Bass, A. J. *et al.* Genomic sequencing of colorectal adenocarcinomas identifies a recurrent *VTI1A-TCF7L2* fusion. *Nature Genet.* **43,** 964–968 (2011).
26. Berger, M. F. *et al.* The genomic complexity of primary human prostate cancer. *Nature* **470,** 214–220 (2011).
27. Stephens, P. J. *et al.* Complex landscapes of somatic rearrangement in human breast cancer genomes. *Nature* **462,** 1005–1010 (2009).

28. Singh, A. *et al.* Dysfunctional KEAP1–NRF2 interaction in non-small-cell lung cancer. *PLoS Med.* **3,** e420 (2006).
29. Singh, A., Bodas, M., Wakabayashi, N., Bunz, F. & Biswal, S. Gain of Nrf2 function in non-small-cell lung cancer cells confers radioresistance. *Antioxid. Redox Signal.* **13,** 1627–1637 (2010).
30. Vaske, C. J. *et al.* Inference of patient-specific pathway activities from multi-dimensional cancer genomics data using PARADIGM. *Bioinformatics* **26,** i237–i245 (2010).
31. Aster, J. C., Blacklow, S. C. & Pear, W. S. Notch signalling in T-cell lymphoblastic leukaemia/lymphoma and other haematological malignancies. *J. Pathol.* **223,** 263–274 (2011).
32. Wang, N. J. *et al.* Loss-of-function mutations in Notch receptors in cutaneous and lung squamous cell carcinoma. *Proc. Natl Acad. Sci. USA* **108,** 17761–17766 (2011).
33. Osada, H., Tatematsu, Y., Yatabe, Y., Horio, Y. & Takahashi, T. *ASH1* gene is a specific therapeutic target for lung cancers with neuroendocrine features. *Cancer Res.* **65,** 10680–10685 (2005).
34. Wilkerson, M. D. *et al.* Lung squamous cell carcinoma mRNA expression subtypes are reproducible, clinically important, and correspond to normal cell types. *Clin. Cancer Res.* **16,** 4864–4875 (2010).
35. Bishop, J. A. *et al.* p40 (ΔNp63) is superior to p63 for the diagnosis of pulmonary squamous cell carcinoma. *Mod. Pathol.* **25,** 405–415 (2011).
36. Massion, P. P. *et al.* Significance of p63 amplification and overexpression in lung cancer development and prognosis. *Cancer Res.* **63,** 7113–7121 (2003).
37. Shen, R., Olshen, A. B. & Ladanyi, M. Integrative clustering of multiple genomic data types using a joint latent variable model with application to breast and lung cancer subtype analysis. *Bioinformatics* **25,** 2906–2912 (2009).
38. Wikman, H. & Kettunen, E. Regulation of the G1/S phase of the cell cycle and alterations in the RB pathway in human lung cancer. *Expert Rev. Anticancer Ther.* **6,** 515–530 (2006).
39. Kancha, R. K., Peschel, C. & Duyster, J. The epidermal growth factor receptor-L861Q mutation increases kinase activity without leading to enhanced sensitivity toward epidermal growth factor receptor kinase inhibitors. *J. Thorac. Oncol.* **6,** 387–392 (2011).
40. Reva, B., Antipin, Y. & Sander, C. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res.* **39,** e118 (2011).
41. Govindan, R. Summary of the proceedings from the 10th annual meeting of molecularly targeted therapy in non-small cell lung cancer. *J. Thorac. Oncol.* **5,** S433 (2010).
42. Ding, L. *et al.* Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* **455,** 1069–1075 (2008).
43. Ciriello, G., Cerami, E., Sander, C. & Schultz, N. Mutual exclusivity analysis identifies oncogenic network modules. *Genome Res.* **22,** 398–406 (2012).
44. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144,** 646–674 (2011).
45. Brahmer, J. R. *et al.* Phase I study of single-agent anti-programmed death-1 (MDX-1106) in refractory solid tumors: safety, clinical activity, pharmacodynamics, and immunologic correlates. *J. Clin. Oncol.* **28,** 3167–3175 (2010).
46. Lynch, T. J. *et al.* Phase II trial of ipilimumab (IPI) and paclitaxel/carboplatin (P/C) in first-line stage IIIb/IV non-small cell lung cancer (NSCLC). *J. Clin. Oncol.* **28,** 7531 (2010).

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

Acknowledgements This study was supported by NIH grants U24 CA126561, U24 CA126551, U24 CA126554, U24 CA126543, U24 CA126546, U24 CA126563, U24 CA126544, U24 CA143845, U24 CA143858, U24 CA144025, U24 CA143882, U24 CA143866, U24 CA143867, U24 CA143848, U24 CA143840, U24 CA143835, U24 CA143799, U24 CA143883, U24 CA143843, U54 HG003067, U54 HG003079 and U54 HG003273.

Author Contributions The TCGA 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 centre. Project activities were coordinated by the National Cancer Institute and National Human Genome Research Institute project teams. We also acknowledge the following TCGA investigators who made substantial contributions to the project: P.S.H. and D.N.H. (manuscript coordinators); M.D.W. (data coordinator); P.S.H. and N.S. (analysis coordinators); P.S.H., M.S.L., A. Sivachecnko, B.H. and G.G. (DNA sequence analysis); M.D.W., J.L. and D.N.H. (mRNA sequence analysis); L. Cope, J.G.H. and L. Danilova (DNA methylation analysis); A.C., G.S., N.H.P., R.K. and M.L. (copy number analysis); N.S., R. Bose, C.J.C., R. Sinha, C.M., S.N., E.A.C., R. Shen, J.N.W. and C. Sander (pathway analysis); A.C. and G.R. (miRNA sequence analysis); W.D.T., B.E.J., D.A.W. and M.-S.T. (pathology and clinical expertise); S.B.B., 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 (https://tcga-data.nci.nih.gov/tcga/tcgaDownload.jsp, https://cghub.ucsc.edu/ and https://tcga-data.nci.nih.gov/docs/publications/lusc_2012/). Reprints and permissions information is available at www.nature.com/reprints. This paper is distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share Alike licence, and the online version of the paper is freely available to all readers. 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 M. Meyerson (matthew_meyerson@dfci.harvard.edu).

The Cancer Genome Atlas Research Network (Participants are arranged by area of contribution and then by institution.)

Genome sequencing centres: Broad Institute Peter S. Hammerman[1,2], Michael S. Lawrence[1], Douglas Voet[1], Rui Jing[1], Kristian Cibulskis[1], Andrey Sivachenko[1], Petar Stojanov[1], Aaron McKenna[1], Eric S. Lander[1,3,4], Stacey Gabriel[5], Gad Getz[1,5], Carrie Sougnez[5], Marcin Imielinski[1,6], Elena Helman[1], Bryan Hernandez[1], Nam H. Pho[1], Matthew Meyerson[1,2,6]

Genome characterization centres: BC Cancer Agency Andy Chu[7], Hye-Jung E. Chun[7], Andrew J. Mungall[7], Erin Pleasance[7], A. Gordon Robertson[7], Payal Sipahimalani[7], Dominik Stoll[7], Miruna Balasundaram[7], Inanc Birol[7], Yaron S. N. Butterfield[7], Eric Chuah[7], Robin J. N. Coope[7], Richard Corbett[7], Noreen Dhalla[7], Ranabir Guin[7], An He[7], Carrie Hirst[7], Martin Hirst[7], Robert A. Holt[7], Darlene Lee[7], Haiyan I. Li[7], Michael Mayo[7], Richard A. Moore[7], Karen Mungall[7], Ka Ming Nip[7], Adam Olshen[8], Jacqueline E. Schein[7], Jared R. Slobodan[7], Angela Tam[7], Nina Thiessen[7], Richard Varhol[7], Thomas Zeng[7], Yongjun Zhao[7], Steven J. M. Jones[7], Marco A. Marra[7]; **Broad Institute** Gordon Saksena[1], Andrew D. Cherniack[1], Stephen E. Schumacher[1,2], Barbara Tabak[1,2], Scott L. Carter[1], Nam H. Pho[1], Huy Nguyen[1], Robert C. Onofrio[5], Andrew Crenshaw[1], Kristin Ardlie[5], Rameen Beroukhim[1,2], Wendy Winckler[1,5], Peter S. Hammerman[1,2], Gad Getz[1,5], Matthew Meyerson[1,2,6]; **Brigham & Women's Hospital/Harvard Medical School** Alexei Protopopov[9,10], Jianhua Zhang[9,10], Angela Hadjipanayis[11,12], Semin Lee[13], Ruibin Xi[13], Lixing Yang[13], Xiaojia Ren[9,11,12], Hailei Zhang[1,9], Sachet Shukla[1,9], Peng-Chieh Chen[11,12], Psalm Haseley[12,13], Eunjung Lee[12,13], Lynda Chin[1,2,9,10,14], Peter J. Park[12,13,15], Raju Kucherlapati[11,12]; **Memorial Sloan-Kettering Cancer Center (TCGA pilot phase only)** Nicholas D. Socci[16], Yupu Liang[16], Nikolaus Schultz[16], Laetitia Borsu[16], Alex E. Lash[16], Agnes Viale[16], Chris Sander[16], Marc Ladanyi[17,18]; **University of North Carolina at Chapel Hill** J. Todd Auman[19,20], Katherine A. Hoadley[21,22,23], Matthew D. Wilkerson[23], Yan Shi[23], Christina Liquori[23], Shaowu Meng[23], Ling Li[23], Yidi J. Turman[23], Michael D. Topal[22,23], Donghui Tan[24], Scot Waring[23], Elizabeth Buda[23], Jesse Walsh[23], Corbin D. Jones[25], Piotr A. Mieczkowski[21], Darshan Singh[23], Junyuan Wu[23], Anisha Gulabani[23], Peter Dolina[23], Tom Bodenheimer[23], Alan P. Hoyle[23], Janae V. Simons[23], Matthew G. Soloway[23], Lisle E. Mose[22], Stuart R. Jefferys[22], Saianand Balu[23], Brian D. O'Connor[23], Jan F. Prins[26], Jinze Liu[27], Derek Y. Chiang[21,23], D. Neil Hayes[23,28], Charles M. Perou[21,22,23]; **University of Southern California/Johns Hopkins** Leslie Cope[29], Ludmila Danilova[29], Daniel J. Weisenberger[30], Dennis T. Maglinte[30], Fei Pan[30], David J. Van Den Berg[30], Timothy Triche Jr[30], James G. Herman[29], Stephen B. Baylin[29], Peter W. Laird[30]

Genome data analysis centres: Broad Institute Gad Getz[1,5], Michael Noble[1], Doug Voet[1], Gordon Saksena[1], Nils Gehlenborg[1,13], Daniel DiCara[1], Jinhua Zhang[9,10], Hailei Zhang[1], Chang-Jiun Wu[2,10], Spring Yingchun Liu[1], Michael S. Lawrence[1], Lihua Zou[1], Andrey Sivachenko[1], Pei Lin[1], Petar Stojanov[1], Rui Jing[1], Juok Cho[1], Marc-Danie Nazaire[1], Jim Robinson[1], Helga Thorvaldsdottir[1], Jill Mesirov[1], Peter J. Park[12,13,15], Lynda Chin[1,2,9,10,14]; **Memorial Sloan-Kettering Cancer Center** Nikolaus Schultz[16], Rileen Sinha[16], Giovanni Ciriello[16], Ethan Cerami[16], Benjamin Gross[16], Anders Jacobsen[16], Jianjiong Gao[16], B. Arman Aksoy[16], Nils Weinhold[16], Ricardo Ramirez[16], Barry S. Taylor[16], Yevgeniy Antipin[16], Boris Reva[16], Ronglai Shen[31], Qianxing Mo[31], Venkatraman Seshan[31], Paul K. Paik[32], Marc Ladanyi[17,18], Chris Sander[16]; **The University of Texas MD Anderson Cancer Center** Rehan Akbani[33], Nianxiang Zhang[33], Bradley M. Broom[33], Tod Casasent[33], Anna Unruh[33], Chris Wakefield[33], R. Craig Cason[34], Keith A. Baggerly[33], John N. Weinstein[33,35]; **University of California Santa Cruz/Buck Institute** David Haussler[36,37], Christopher C. Benz[38], Joshua M. Stuart[36], Jingchun Zhu[36], Christopher Szeto[36], Gary K. Scott[38], Christina Yau[38], Sam Ng[36], Ted Goldstein[36], Peter Waltman[36], Artem Sokolov[36], Kyle Ellrott[36], Eric A. Collisson[39], Daniel Zerbino[36], Christopher Wilks[36], Singer Ma[36], Brian Craft[36]; **University of North Carolina at Chapel Hill** Matthew D. Wilkerson[23], J. Todd Auman[19,20], Katherine A. Hoadley[21,22,23], Ying Du[23], Christopher Cabanski[23], Vonn Walter[23], Darshan Singh[23], Junyuan Wu[23], Anisha Gulabani[23], Tom Bodenheimer[23], Alan P. Hoyle[23], Janae V. Simons[23], Matthew G. Soloway[23], Lisle E. Mose[22], Stuart R. Jefferys[22], Saianand Balu[23], J. S. Marron[40], Yufeng Liu[24], Kai Wang[27], Jinze Liu[27], Jan F. Prins[23], D. Neil Hayes[23,28], Charles M. Perou[21,22,23]; **Baylor College of Medicine** Chad J. Creighton[41], Yiqun Zhang[41]

Pathology committee William D. Travis[42], Natasha Rekhtman[42], Joanne Yi[43], Marie C. Aubry[43], Richard Cheney[44], Sanja Dacic[45], Douglas Flieder[46], William Funkhouser[47], Peter Illei[48], Jerome Myers[49], Ming-Sound Tsao[50]

Biospecimen core resources: International Genomics Consortium Robert Penny[51], David Mallery[51], Troy Shelton[51], Martha Hatfield[51], Scott Morris[51], Peggy Yena[51], Candace Shelton[51], Mark Sherman[51], Joseph Paulauskis[51]

Disease working group Matthew Meyerson[1,2,6], Stephen B. Baylin[29], Ramaswamy Govindan[52], Rehan Akbani[33], Ijeoma Azodo[53], David Beer[54], Ron Bose[52], Lauren A. Byers[55], David Carbone[56], Li-Wei Chang[52], Derek Chiang[21,23], Andy Chu[7], Elizabeth Chun[7], Eric Collisson[39], Leslie Cope[29], Chad J. Creighton[41], Ludmila Danilova[29], Li Ding[52], Gad Getz[1,5], Peter S. Hammerman[1,2], D. Neil Hayes[23,28], Bryan Hernandez[1], James G. Herman[29], John Heymach[55], Cristiane Ida[43], Marcin Imielinski[1,6], Bruce Johnson[2], Igor Jurisica[57], Jacob Kaufman[56], Farhad Kosari[53], Raju Kucherlapati[11,12], David Kwiatkowski[2], Marc Ladanyi[17,18], Michael S. Lawrence[1], Christopher A. Maher[52], Andy Mungall[7], Sam Ng[36], William Pao[56], Martin Peifer[58,59], Robert Penny[51], Gordon Robertson[7], Valerie Rusch[60], Chris Sander[16], Nikolaus Schultz[16], Ronglai Shen[31], Jill Siegfried[61], Rileen Sinha[16], Andrey Sivachenko[1], Carrie Sougnez[4], Dominik Stoll[7], Joshua Stuart[36], Roman K. Thomas[58,59,62], Sandra Tomaszek[53], Ming-Sound Tsao[50],

William D. Travis[42], Charles Vaske[36], John N. Weinstein[33,35], Daniel Weisenberger[30], David Wheeler[63], Dennis A. Wigle[53], Matthew D. Wilkerson[23], Christopher Wilks[30], Ping Yang[53], Jianjua John Zhang[9,10]

Data coordination centre Mark A. Jensen[64], Robert Sfeir[64], Ari B. Kahn[64], Anna L. Chu[64], Prachi Kothiyal[64], Zhining Wang[64], Eric E. Snyder[64], Joan Pontius[64], Todd D. Pihl[64], Brenda Ayala[64], Mark Backus[64], Jessica Walton[64], Julien Baboud[64], Dominique L. Berton[64], Matthew C. Nicholls[64], Deepak Srinivasan[64], Rohini Raman[64], Stanley Girshik[64], Peter A. Kigonya[64], Shelley Alonso[64], Rashmi N. Sanbhadti[64], Sean P. Barletta[64], John M. Greene[64], David A. Pot[64]

Tissue source sites Ming-Sound Tsao[50], Bizhan Bandarchi-Chamkhaleh[50], Jeff Boyd[46], JoEllen Weaver[46], Dennis A. Wigle[53], Ijeoma A. Azodo[53], Sandra C. Tomaszek[53], Marie Christine Aubry[65], Christiane M. Ida[65], Ping Yang[66], Farhad Kosari[53], Malcolm V. Brock[67], Kristen Rodgers[67], Marian Rutledge[68], Travis Brown[67], Beverly Lee[68], James Shin[69], Dante Trusty[69], Rajiv Dhir[70], Jill M. Siegfried[61], Olga Potapova[71], Konstantin V. Fedosenko[72], Elena Nemirovich-Danchenko[71], Valerie Rusch[60], Maureen Zakowski[73], Mary V. Iacocca[74], Jennifer Brown[74], Brenda Rabeno[74], Christine Czerwinski[74], Nicholas Petrelli[74], Zhen Fan[75], Nicole Todaro[75], John Eckman[49], Jerome Myers[49], W. Kimryn Rathmell[23], Leigh B. Thorne[76], Mei Huang[76], Lori Boice[76], Ashley Hill[23], Robert Penny[51], David Mallery[51], Erin Curley[51], Candace Shelton[51], Peggy Yena[51], Carl Morrison[44], Carmelo Gaudioso[44], John M. S. Bartlett[77], Sugy Kodeeswaran[77], Brent Zanke[77], Harman Sekhon[78], Kerstin David[79], Hartmut Juhl[80], Xuan Van Le[81], Bernard Kohl[81], Richard Thorp[81], Nguyen Viet Tien[82], Nguyen Van Bang[83], Howard Sussman[84], Bui Duc Phu[83], Richard Hajek[85], Nguyen Phi Hung[86], Khurram Z. Khan[87], Thomas Muley[88]

Project team: National Cancer Institute Kenna R. Mills Shaw[89], Margi Sheth[89], Liming Yang[89], Ken Buetow[90], Tanja Davidsen[90], John A. Demchok[89], Greg Eley[90], Martin Ferguson[91], Laura A. L. Dillon[89], Carl Schaefer[90]; **National Human Genome Research Institute** Mark S. Guyer[92], Bradley A. Ozenberger[92], Jacqueline D. Palchik[92], Jane Peterson[92], Heidi J. Sofia[92], Elizabeth Thomson[92]

Writing committee Peter S. Hammerman[1,2], D. Neil Hayes[23,28], Matthew D. Wilkerson[23], Nikolaus Schultz[16], Ron Bose[52], Andy Chu[7], Eric A. Collisson[39], Leslie Cope[29], Chad J. Creighton[41], Gad Getz[1,5], James G. Herman[29], Bruce E. Johnson[2], Raju Kucherlapati[11,12], Marc Ladanyi[17,18], Christopher A. Maher[52], Gordon Robertson[7], Chris Sander[16], Ronglai Shen[16], Rileen Sinha[16], Andrey Sivachenko[1], Roman K. Thomas[58,59,62], William D. Travis[42], Ming-Sound Tsao[50], John N. Weinstein[33,35], Dennis A. Wigle[53], Stephen B. Baylin[29], Ramaswamy Govindan[52], Matthew Meyerson[1,2,6]

[1]The Eli and Edythe L. Broad Institute of Massachusetts Institute of Technology and Harvard University Cambridge, Massachusetts 02142, USA. [2]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts 02215, USA. [3]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA. [4]Department of Systems Biology, Harvard University, Boston, Massachusetts 02115, USA. [5]Genetic Analysis Platform, The Eli and Edythe L. Broad Institute of Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts 02142, USA. [6]Department of Pathology, Harvard Medical School, Boston, Massachusetts 02115, USA. [7]Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency, Vancouver, British Columbia V5Z, Canada. [8]Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, California 94143, USA. [9]Belfer Institute for Applied Cancer Science, Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA. [10]Institute for Applied Cancer Science, Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. [11]Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA. [12]Division of Genetics, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. [13]The Center for Biomedical Informatics, Harvard Medical School, Boston, Massachusetts 02115, USA. [14]Department of Dermatology, Harvard Medical School, Boston, Massachusetts 02115, USA. [15]Informatics Program, Children's Hospital, Boston, Massachusetts 02115, USA. [16]Computational Biology Center, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. [17]Department of Molecular Oncology, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. [18]Department of Pathology and Human Oncology & Pathogenesis Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. [19]Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [20]Institute for Pharmacogenetics and Individualized Therapy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [21]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [22]Department of Pathology and Laboratory Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [23]Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [24]Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [25]Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [26]Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [27]Department of Computer Science, University of Kentucky, Lexington, Kentucky 40506, USA. [28]Department of Internal Medicine, Division of Medical Oncology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [29]Cancer Biology Division, The Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins University, Baltimore, Maryland 21231, USA. [30]University of Southern California Epigenome Center, University of Southern California, Los Angeles, California 90033, USA. [31]Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. [32]Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. [33]Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. [34]Division of Pathology and Laboratory Medicine, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. [35]Department of Systems Biology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. [36]Department of Biomolecular Engineering and Center for Biomolecular Science and Engineering, University of California Santa Cruz, Santa Cruz, California 95064, USA. [37]Howard Hughes Medical Institute, University of California Santa Cruz, Santa Cruz, California 95064, USA. [38]Buck Institute for Age Research, Novato, California 94945, USA. [39]Division of Hematology/Oncology, University of California San Francisco, San Francisco, California 94143, USA [40]Department of Statistics and Operations Research, University of North Carolina Medical Center, Chapel Hill, North Carolina 27599, USA. [41]Human Genome Sequencing Center and Dan L. Duncan Cancer Center Division of Biostatistics, Baylor College of Medicine, Houston, Texas 77030, USA. [42]Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, New York 10065 USA. [43]Department of Pathology, Mayo Clinic, Rochester, Minnesota 55905, USA. [44]Department of Pathology, Roswell Park Cancer Institute, Buffalo, New York 14263, USA. [45]Department of Pathology, University of Pittsburgh Cancer Center, Pittsburgh, Pennsylvania 15213, USA. [46]Department of Pathology, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111, USA. [47]Department of Pathology, University of North Carolina Medical Center, Chapel Hill, North Carolina 27599, USA. [48]Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, Maryland 21287, USA. [49]Department of Pathology, Penrose-St. Francis Health System, Colorado Springs, Colorado 80907, USA. [50]Department of Pathology and Medical Biophysics, Ontario Cancer Institute and Princess Margaret Hospital, Toronto, Ontario M5G 2MY, Canada. [51]International Genomics Consortium, Phoenix, Arizona 85004, USA. [52]Division of Oncology, Department of Medicine and The Genome Institute, Washington University School of Medicine, St. Louis, Missouri 63110, USA. [53]Center for Individualized Medicine, Mayo Clinic, Rochester, Minnesota 55905, USA. [54]Department of Surgery, University of Michigan, Ann Arbor, Michigan 48109, USA. [55]The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. [56]Departments of Hematology/Oncology and Cancer Biology, Vanderbilt University School of Medicine, Nashville, Tennessee 37232, USA. [57]Ontario Cancer Institute, IBM Life Sciences Discovery Centre, Toronto, Ontario M5G 1L7, Canada. [58]Department of Translational Genomics, University of Cologne, Cologne D-50931, Germany. [59]Max Planck Institute for Neurological Research, Cologne D-50866, Germany. [60]Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, New York 10065, USA. [61]Department of Pharmacology and Chemical Biology, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania 15232, USA. [62]Department of Translational Cancer Genomics, Center of Integrated Oncology, University of Cologne, Cologne D-50924, Germany. [63]Human Genome Sequencing Center, Baylor College of Medcine, Houston, Texas 77030, USA. [64]SRA International, Fairfax, Virginia 22033, USA. [65]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota 55905, USA. [66]Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota 55905, USA. [67]Department of Surgery, Johns Hopkins School of Medicine, 600 North Wolfe Street, Baltimore, Maryland 21287, USA. [68]Department of Oncology, Johns Hopkins School of Medicine, 600 North Wolfe Street, Baltimore, Maryland 21287, USA. [69]Department of Pathology, Johns Hopkins School of Medicine, 600 North Wolfe Street, Baltimore, Maryland 21287, USA. [70]Department of Pathology, University of Pittsburgh, Pittsburgh, Pennsylvania 15213, USA. [71]Cureline, South San Francisco, California 94080, USA. [72]City Clinical Oncology Dispensary, St Petersburg 197022, Russia. [73]Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, New York 10065, USA. [74]Helen F. Graham Cancer Center, Newark, Delaware 19713, USA. [75]St Joseph Medical Center, Towson, Maryland 21204, USA. [76]UNC Tissue Procurement Facility, Department of Pathology, UNC Lineberger Cancer Center, Chapel Hill, North Carolina 27599, USA. [77]Ontario Tumour Bank, Ontario Institute for Cancer Research, Toronto, Ontario M5G 0A3, Canada. [78]Ontario Tumour Bank – Ottawa site, The Ottawa Hospital, Ottawa, Ontario K1H 8L6, Canada. [79]Indivumed GmbH, Hamburg, Falkenried 88, Haus D D-20251, Germany. [80]Indivumed Inc, Kensington, Maryland 20895, USA. [81]ILSBio, LLC, Chestertown, Maryland 21620, USA. [82]Ministry of Health, 138A Giang Vo Street, Hanoi, Vietnam. [83]Hue Central Hospital, Hue City, 16 Le Loi, Hue, Vietnam. [84]Stanford University Medical Center, Stanford, California 94305, USA. [85]Center for Minority Health Research, University of Texas, M.D. Anderson Cancer Center, Houston, Texas 77030, USA. [86]National Cancer Institute, 43 Quan Su Street, Hanoi, Vietnam. [87]ILSBio LLC, Chestertown, Maryland 21620, USA. [88]ThoraxKlinik, Heidelberg University Hospital, Heidelberg 69126, Germany. [89]The Cancer Genome Atlas Program Office, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA. [90]Center for Biomedical Informatics and Information Technology (CBIIT), National Cancer Institute, National Institutes of Health, Rockville, Maryland 20852, USA. [91]MLF Consulting, Arlington, Maryland 02474, USA. [92]National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892, USA.

CORRIGENDUM

doi:10.1038/nature11666

Corrigendum: Comprehensive genomic characterization of squamous cell lung cancers

The Cancer Genome Atlas Research Network

Nature **489,** 519–525 (2012); doi:10.1038/nature11404

In this Article, author Kristen Rodgers was spelt incorrectly. This error has been corrected in the HTML and PDF of the original paper.