Genomic alterations in lung adenocarcinoma

Siddhartha Devarakonda, Daniel Morgensztern, Ramaswamy Govindan

Treatment for non-small-cell lung cancer is evolving from the use of cytotoxic chemotherapy to personalised treatment based on molecular alterations. This past decade has witnessed substantial progress in the treatment of patients with EGFR mutations and ALK rearrangements, and it is now possible to study complex genomic alterations in cancer using next-generation sequencing. Sequencing data from large-scale consortia, such as The Cancer Genome Atlas, as well as several independent groups, have helped identify novel drivers and potentially targetable alterations in lung adenocarcinomas. These data clearly suggest that lung adenocarcinoma is associated with distinct genomic alterations compared with other lung cancer subtypes, and highlight the widespread molecular heterogeneity that underlies the disease. In this Review, we discuss some of the key findings from genomic studies of lung adenocarcinoma.

Lancet Oncol 2015; 16: e342–51

Division of Medical Oncology, Washington University School of Medicine, St Louis, MO, USA (S Devarakonda MD, D Morgensztern MD, Prof R Govindan MD); **and Alvin J Siteman Cancer Center, St Louis, MO, USA** (D Morgensztern, R Govindan)

Correspondence to:
Prof Ramaswamy Govindan, Division of Medical Oncology, Washington University School of Medicine, 660 S Euclid, Box 8056, St Louis, MO 63021, USA
rgovinda@dom.wustl.edu

Introduction

Lung cancer is the predominant cause of cancer-related mortality in the USA.[1] Lung adenocarcinoma is the most diagnosed histological subtype of non-small-cell lung cancer (NSCLC), followed by squamous cell carcinoma.[2] Although the advent of targeted therapies has improved outcomes in a subset of patients, most patients with metastatic adenocarcinoma are treated with empirical chemotherapy regimens.[3–6] The study of complex genomic alterations in cancer cells by use of next generation sequencing technologies is now possible. Research efforts, undertaken by large-scale consortia such as The Cancer Genome Atlas (TCGA), the International Cancer Genome Consortium (ICGC), and several other independent groups, have culminated in the generation of vast amounts of genomic data. These studies have identified many novel driver and potentially targetable alterations and considerable intertumour and intratumour heterogeneity.

Genomic landscape

TCGA is a collective effort that aims to characterise the genomic alterations driving various malignancies through large-scale sequencing.[7] TCGA lung cancer working group plans to do comprehensive genomic analyses of 1000 tumour specimens from patients with adenocarcinoma and squamous cell carcinoma. Genomic analyses of 178 squamous cell carcinoma samples and 230 adenocarcinoma samples from TCGA and other studies have enhanced our knowledge of molecular alterations in lung cancer.[8–11] The aim of this Review is to summarise some of the key findings from TCGA and other genomic studies in lung adenocarcinoma (figure 1), and discuss their potential clinical application.

Somatic copy number alterations

Somatic copy number alterations affect a large fraction of the cancer cell genome and are widely prevalent across various malignancies including lung cancer.[12,13] Copy number analysis of 528 snap frozen lung adenocarcinoma samples by Weir and colleagues[14] identified 26 large-scale significant copy-number gains and losses (defined as those including at least half a chromosome arm), and 31 recurrent focal (chromosomal segment) events. Copy-number gain of chromosome 5p was the most frequent alteration in lung adenocarcinoma in this study, and the investigators suggested that the telomerase catalytic subunit gene, telomerase reverse transcriptase (TERT), on 5p15 was the target of this amplification. Telomerase is a ribonucleoprotein enzyme that catalyses the synthesis of telomeric DNA, a process that is crucial in enabling replicative immortality in cancer cells.[15,16] The TERT enzyme catalyses this process with the RNA component of telomerase (TERC) as a template. Amplifications in *TERT* were noted in 18% of TCGA lung adenocarcinoma samples.[11,17,18] Similarly, *TERC*, located on the chromosome 3q, was amplified in 4% of TCGA lung adenocarcinoma samples.[17,18]

Focal amplifications including 14q13·3 are frequent in lung adenocarcinoma.[14] This region harbours NKX2–1 (*TTF1*), a transcription factor crucial for development of the lung, thyroid, and brain.[19,20] Localised expression of *NKX2–1* in the foregut marks one of the earliest steps in lung organogenesis.[21] This transcription factor also regulates development of peripheral airway cells that constitute the terminal respiratory unit.[20] Inhibition of *NKX2–1* by RNA interference is associated with impaired viability, colony formation, and anchorage-independent growth in lung cancer cell lines that express *NKX2–1*.[14] Amplifications in *NKX2–1* were noted in 14% of TCGA samples.[11,17,18] These findings support a role for *NKX2–1* as a lineage survival oncogene in lung adenocarcinoma, similar to *SOX2* in squamous cell carcinoma and small-cell lung cancers (SCLC).[22,23] Nevertheless, the therapeutic significance of *NKX2–1* in lung adenocarcinoma is unclear.

MYC transcription factors are among the downstream targets of several mitogenic pathways and regulate many cellular processes including cell division.[24,25] *MYC* amplification has been previously reported in lung adenocarcinomas and SCLC, and was seen in 9% of TCGA samples.[11,17,18,22,26] These amplifications were mutually exclusive with loss of function and nonsense mutations in *MGA* (mutated in 8% of samples). *MGA* codes for a Max-interacting protein, which functions as a transcriptional repressor capable of blocking MYC-dependent transformation.[27]

Overall, these results suggest that somatic copy number alterations are fairly prevalent in lung adenocarcinomas. Furthermore, data suggest that somatic copy

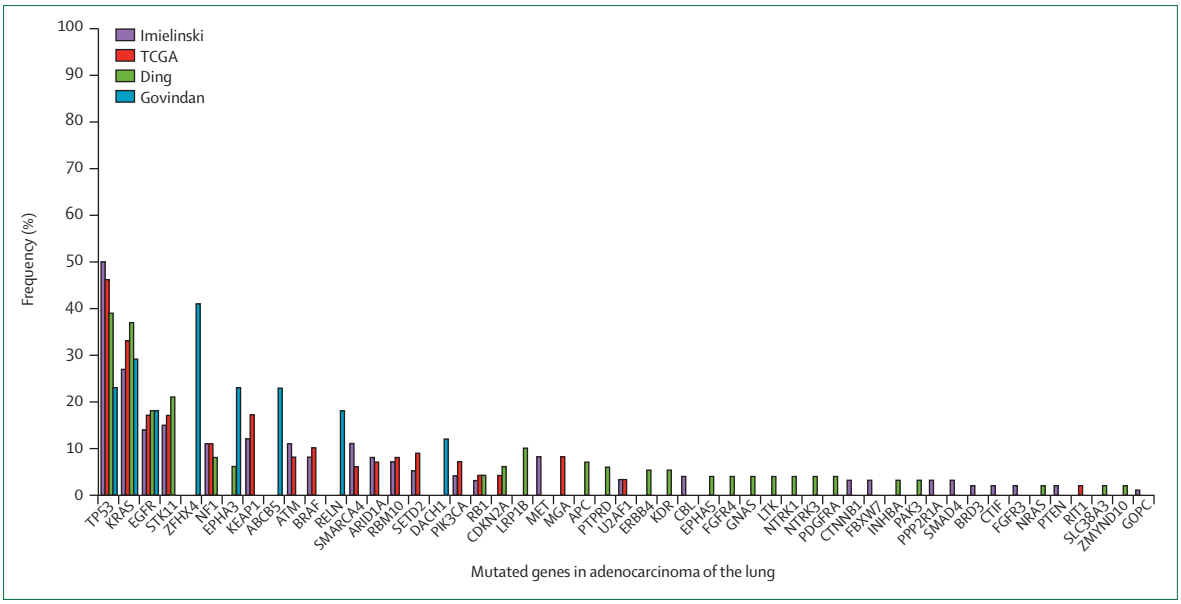

Figure 1: **Frequency of lung adenocarcinoma mutations at a statistically significant level across different studies**
For establishing significance, false discovery rate or Q threshold was <0·05 for The Cancer Genome Atlas study and Govindan and colleagues' study, and <0·25 for Imielinski and colleagues' study. Three different methods were used by Ding and colleagues to measure significance, and the false discovery rate cut off was 0·1.

number alterations occur both as early and late events during tumour evolution.[28,29] Some of the other recurrent somatic copy number alterations in lung adenocarcinoma include amplifications in *EGFR, MET, KRAS, ERBB2,* and *MDM2,* and deletions in *LRP1B, PTPRD,* and *CDKN2A*. Functional studies are needed to validate the role of some of these somatic copy number alterations in the pathogenesis of lung cancer to help the development of rational targeted therapies.

Lung adenocarcinoma mutations

Lung cancer is characterised by a high mutational burden compared with many other cancer types studied so far.[9–11,30–32] A mean somatic mutation rate of 8·87 per megabase, and a non-synonymous mutation rate of 6·86 per megabase of DNA were seen after whole-exome sequencing of 230 lung adenocarcinoma samples by TCGA. In a systematic analysis of 3281 tumour samples sequenced by TCGA across 12 cancer types, acute myeloid leukaemia was associated with the lowest mutational burden and squamous cell carcinoma of the lung with the highest mutational burden, followed by lung adenocarcinoma.[32] Adenocarcinoma and squamous cell carcinoma showed an increased frequency of cytosine to adenine mutations, which suggests exposure to tobacco smoke. The high mutational burden of lung cancer samples is explained by the mutagenic effect of carcinogen exposure (eg, cigarette smoke).[9] Melanoma samples similarly show a high mutational burden, as a result of exposure to ultraviolet radiation.[33] Adenocarcinoma genomes also show regional heterogeneity in the distribution of mutations. One of the factors that affects the distribution of mutations is the level of gene expression.[34] The mutation rate in highly expressed genes is postulated to be low because of transcription-coupled repair mechanisms. For instance, in one study,[9] highly expressed genes showed less than four mutations per megabase whereas unexpressed genes showed roughly 14 mutations per megabase.[9]

Significantly mutated oncogenes

Data from several lung cancer sequencing studies have consistently established that lung cancer is a molecularly heterogeneous disease. The mutational landscape of lung adenocarcinoma is substantially different from that of squamous cell carcinoma or SCLC. For instance, although mutations in receptor tyrosine kinases are frequent in lung adenocarcinomas, these alterations are rarely encountered in squamous cell carcinoma and SCLC (table 1).[8,11,22,35] Furthermore, studies have also established lung cancer in people who have never smoked to be demographically and molecularly distinct from smokers diagnosed with lung cancer.[36] Nearly 10% of all lung cancers worldwide are diagnosed in people who have never smoked. Lung cancer in this group is more frequent in women and almost always associated with lung adenocarcinoma histology. Molecularly, cytosine to thymidine transitions in general are more common in people who have never smoked, whereas tumours from smokers are usually enriched for cytosine to adenine transversions.[9,10] Although TCGA and other groups have identified genes such as *TP53, KRAS, STK11* (*LKB1*), *EGFR,* and *NF1* to be substantially mutated in adenocarcinomas, it is now evident that these mutations are differentially enriched in tumours classified by smoking status and gender. For instance, mutations in

TP53, KRAS, LKB1, NF1, and *RBM10* are enriched in transversion-high tumours and mutations in *EGFR, RB1,* and *PIK3CA,* and in-frame insertions in the receptor tyrosine kinases *EGFR* and *ERBB2,* are enriched in transversion-low tumours.[11] Similarly, mutations in *EGFR* and *RBM10* show an association with gender, with *EGFR* being more prevalent in women and *RBM10* more prevalent in men. These findings suggest that the pathogenesis of lung cancer varies based on smoking status and gender.

KRAS mutations

Mutations in the RAS family of genes, are frequently encountered in various malignancies.[37,38] *KRAS* mutations are observed in nearly 33% of TCGA lung adenocarcinoma samples, and almost never present in squamous cell carcinoma and SCLC.[8,11,22,35] *KRAS* mutations are frequently seen in smokers and are associated with codons 12, 13, or 61. Among these, codon 12 mutations are the most frequent and result in the substitution of glycine with either cytosine (Gly12Cys) or valine (Gly12Val) in most cases. These missense mutations maintain *KRAS* in a constitutively active state.

Cell-line studies and some clinical data show differential sensitivity of different *KRAS* genotypes to conventional therapies.[39,40] This effect is probably mediated by the ability of different *KRAS* mutants to engage in various downstream signalling pathways including RAF/MEK, PI3K, and RALGDS, some of which might be potentially targetable. Castellano and colleagues[41] showed partial regression of *KRAS*-driven lung tumours in mice by genetically disrupting PIK3CA-RAS interaction where concomitant MEK inhibition further improved this effect. These results suggest a role for PI3K inhibitors in *KRAS*-driven lung cancers. Kim and colleagues[42] reported addiction of *KRAS* and *LKB1* comutated NSCLC cell lines to the coatomer complex I dependent lysosomal acidification. The investigators showed that blocking this pathway inhibited survival in both *KRAS* and *LKB1* comutated cells. *KRAS* and *LKB1* comutations were recorded in nearly 9% of TCGA lung adenocarcinoma samples[17,18] The development of small-molecule inhibitors capable of irreversibly binding Gly12Cys mutant *KRAS*, the most common subtype of *KRAS* mutation in lung adenocarcinoma, has been reported.[43] Taken together, these data highlight the molecular heterogeneity that underlies *KRAS*-mutated tumours. The unique vulnerabilities observed in a subset of these tumours will need to be studied further through well-designed trials to improve outcomes in these patients.

EGFR mutations

EGFR belongs to the ERBB family of cell-surface tyrosine kinase receptors. Upon activation, these receptors homodimerise or heterodimerise with other ERBB members and trigger the RAS/RAF/MEK and PI3K/AKT/mTOR signalling pathways.[44] *EGFR* is mutated in about 16% of tumour specimens from patients with non-squamous NSCLC.[45] Tumours with *EGFR* mutations have increased sensitivity to reversible small-molecule receptor tyrosine kinase inhibitors, such as gefitinib and erlotinib, and irreversible inhibitors such as neratinib, dacomitinib, and afatinib, which inhibit the EGFR signalling and result in cancer-cell apoptosis.[46] Known oncogenic mutations were observed in nearly 11% of TCGA lung adenocarcinoma samples.[11] *EGFR* mutations are generally more frequent in non-smokers and in women.

The sensitivity of *EGFR* mutant tumours to receptor tyrosine kinase inhibitors has been shown to depend on the genotype of *EGFR*.[44] Exon 19 deletions, exon 21 mutations, and exon 18 mutations are often associated with sensitivity to EGFR tyrosine kinase inhibitors. By contrast, exon 20 insertion mutations are associated with a variable and often decreased sensitivity to tyrosine kinase inhibitors.[47] Patients given EGFR tyrosine kinase inhibitors eventually develop resistance through various mechanisms including the *EGFR* Thr790Met mutation, *MET* amplification, *PTEN* loss,

	Adenocarcinoma	SCC	SCLC
Mutations	*TP53, KRAS, EGFR, NF1, BRAF, MET, RIT*	*TP53, CDKN2A, PIK3CA, NFE2L2, KEAP1, CUL3, PTEN, NF1, NOTCH1,2,* and *3, DDR2, EGFR*	*TP53, RB1, EP300, CREBBP, PTEN, SLIT2*
Fusions	*ALK, ROS1, RET*	*FGFRs*	..
SCNAs	Gains: *NKX2-1, TERT, EGFR, MET, KRAS, ERBB2, MDM2* Losses: *LRP1B, PTPRD, and CDKN2A*	Gains: Chr 3q 26 (*SOX2, PIK3CA, TP63* etc) Losses: *CDKN2A, PTEN*	Gains: *MYC, MYCN, MYCL1, SOX2, FGFR1, KIT* Losses: Chr 3p *(FHIT, FUS1, RASSF1A)*
Pathway alterations	RTK/RAS/RAF mTOR JAK-STAT DNA repair Cell cycle regulation Epigenetic deregulation	Squamous differentiation Oxidative stress response PIK3CA DNA repair Cell cycle regulation Epigenetic deregulation	Cell cycle regulation, epigenetic deregulation, Hedgehog, DNA repair, axonal guidance and neuroendocrine differentiation

SCNA=somatic copy number alteration.

Table 1: **Recurrent features in lung adenocarcinomas, squamous cell carcinoma (SCC), and small-cell lung cancer (SCLC)**

and histological transformation to SCLC.[48] Nearly half of the *EGFR* tyrosine kinase inhibitor-resistant tumours harbour the Thr790Met mutation. New tyrosine kinase inhibitors, such as rociletinib and mereletinib, have been reported to be active in these tumours.[49,50]

BRAF mutations

BRAF operates downstream of RAS proteins, and constitutes a crucial step in the RAS-MAPK pathway. *BRAF* mutations are frequent in melanoma, and vemurafenib has been shown to target the the Val600Glu variant in these cancers.[51] *BRAF* mutations are reported in roughly 7% of lung adenocarcinomas by TCGA.[11] The most frequently encountered *BRAF* mutation was Val600Glu (five samples), followed by Gly469 (four), Gly466 (three), Asp594 and Asn581 (two each) mutations. Although Val600Glu results in kinase activation and is capable of oncogenic transformation, the role of other, less common *BRAF* mutations in NSCLC is unclear. Although Val600Glu mutations are generally more prevalent in women and associated with worse outcomes, non-Val600Glu mutations are prevalent in smokers and are not associated with prognosis.[52] Tumours with Val600Glu mutations are usually negative for *KRAS* and *EGFR* mutations, suggesting a driver role for *BRAF*. Results with multi-kinase inhibitors with activity against BRAF, such as sorafenib, have been disappointing in non-biomarker-selected patients with NSCLC.[53] In a phase 2 study, dabrafenib showed promising activity in patients with lung adenocarcinoma harbouring BRAF Val600Glu mutations.[54]

MET mutations

MET activation through exon 14 skipping (analysis of both DNA and RNA sequencing data suggested that exon 14 was not expressed in these samples) and MET amplification have been reported in lung adeno-carcinomas.[11,55] Exon 14 skipping results in the loss of a negative regulatory site, and is often the result of splice-site mutations.[11,55] Mutations that caused exon 14 skipping were mutually exclusive with *EGFR* and *KRAS* mutations. Although missense mutations in *MET* were noted concordantly with *KRAS* and *EGFR* mutations, the oncogenic potential of these variants is unknown. *MET* amplification is a known mechanism of resistance to EGFR tyrosine kinase inhibitors.[48] Responses to crizotinib in *MET*-amplified NSCLC have been reported in case reports and small studies.[56,57] Therefore, additional data are definitely needed to support its routine use in patients with *MET* alterations.

ERBB2 mutations

ERBB2 (*HER2*), similar to *EGFR*, belongs to the ERBB family of receptor tyrosine kinases. Activation of ERBB2 through amplification and mutation has been reported in NSCLC.[58] Activating mutations in *ERBB2* are frequently exon 20 insertions.[59–61] Responses to trastuzumab and EGFR/ERBB2 dual inhibitors, afatinib and neratinib, have been reported in cell lines and tumours with exon 20 insertions.[59,62–64] *ERBB2* exon 20 mutations and amplifications were noted in 3% of TCGA samples.[11]

Roughly 62% of lung adenocarcinomas sequenced by TCGA had alterations in known driver oncogenes belonging to the RTK/RAS/RAF pathway. An effort was made by TCGA to identify possible oncogenic drivers in a subset of tumours without known oncogenic alterations in this pathway. Amplifications including *MET* and *ERBB2* and mutations in *TP53, KEAP1, NF1,* and *RIT1* were enriched in these tumours. Considering the role of *NF1* in RAS inhibition and *RIT1* in MEK and PI3K activation these alterations probably represent driver events.[65,66] Upon including *MET, ERBB2, RIT1,* and *NF1* alterations, driver events were identifiable in nearly 76% of samples (table 2). Although a select few cancer-causing genes (eg, *KRAS* in lung adenocarcinoma) are mutated at high frequencies (>20%) in a given cancer type, most cancer genes are mutated at intermediate frequencies (2–20%) or lower.[67] Large-scale sequencing studies could potentially uncover more low frequency driver events and the pathways they alter in lung cancer.

Gene fusions

Several studies have reported oncogenic fusions and in-frame rearrangements with several kinases in lung adenocarincomas. Of these, the discovery of fusions including the tyrosine kinases has substantially changed the setting of lung cancer treatment. About 3–8% of lung adenocarcinomas show ALK rearrangements. Crizotinib and ceritinib are approved for use in this subset of NSCLC.[6,68–70] Several mechanisms of resistance to crizotinib, including mutations in *ALK* that alter the drug-binding site (eg, Leu1196Met), *ALK* amplification, and activation of bypass pathways, have been described so far.[48] Ceritinib is active in crizotinib-resistant ALK-rearranged tumours.[68] Data also support the use of crizotinib in *ROS1*-rearranged lung adenocarcinoma.[71] Responses with tyrosine kinase inhibitors such as vandetanib and cabozantinib in *RET* fusion-positive tumours have also been reported.[72] Overall, *ALK, ROS1,* and *RET* fusions were recorded in three, four, and two TCGA samples, respectively.[11]

With next-generation sequencing and FISH probes, Vaishnavi and colleagues[73] reported an oncogenic fusion including the tyrosine receptor kinase *NTRK1* in three of 91 samples from patients with lung adenocarcinoma without known oncogenic alterations. Proliferation of Ba/F3 cells expressing *NTRK1* fusion proteins was inhibited by drugs such as lestaurtinib, ARRY-470, and to a lesser extent, crizotinib, suggesting that routine screening of lung adenocarcinomas for these fusions could be clinically relevant. Analysis of the transcriptome data from 230 TCGA lung adenocarcinoma samples, however, did not show the presence of *NTRK1* fusions. In-frame

rearrangements including kinases *SIK2* and *ROCK1* in lung adenocarcinomas without known oncogenic mutations, novel non-kinase gene fusions including *RASSF1A*, and the DNA-binding protein *FZR1* were reported in previous next-generation sequencing studies.[9,10] A larger study including several thousands of samples is probably needed to fully identify rare oncogenic fusion kinases that play a part in the molecular pathogenesis of lung adenocarcinoma since many gene fusions identified so far in lung adenocarcinoma are rarely recurrent.[9]

Transcriptome analyses

Combined analysis of genome and transcriptome sequencing data allows determination of the concordance between identified variants. Furthermore, RNA-seq data can be used to further classify variants into categories such as expressed or silent based on their expression.[9] Expressed heterozygous variants can be further classified as mutant or wild-type biased, depending on the variant allele frequency of the expressed variant in RNA-seq data. The variant allele frequency at a specific tumour location for a mutated gene is calculated by dividing the number of times the mutant allele is encountered by the total number of alleles, both mutant and wildtype, encountered at that location. For instance, Govindan and colleagues[9] considered variants showing a variant allele frequency that was greater than 20% higher in the RNA-sequencing data than in whole genome sequencing data as having a mutant-biased expression, and variants with at least 20% lower variant allele frequencies in the RNA-sequencing data than in whole genome sequencing data as having a wild type-biased expression. In this study, genomes of patients who never smoked showed a higher proportion of expressed variants (49·4%) than lung cancer genomes from smokers. Genes such as *KRAS* and *TP53* show mutant-biased expression patterns in lung adenocarcinoma. Furthermore, analysis of transcriptome data also helps in the examination of the association between mutation frequency and gene expression, deregulation of splicing, and tumour classification.

Deregulated splicing

Splicing is a mechanism by which introns from pre-mRNA are removed and exons are joined together to synthesise an mRNA molecule that codes for a protein. Cells can achieve proteomic diversity through a tightly regulated mechanism known as alternative splicing by which one gene can code for several protein isoforms by varying the exons included in the final mRNA (figure 2). Deregulated alternative splicing has been implicated in oncogenesis.[74] The combined analysis of exome sequencing and transcriptome data allows identification of both the effects of gene-splice site mutations on mRNA splicing and differentially expressed splice forms between healthy and tumour samples.[75]

MET activation by exon 14 skipping, which removes a negative regulatory site from the protein, was identified in ten of 230 TCGA samples.[11] Nine of these samples harboured 3′ or 5′ splice-site mutations. Similarly, recurrent S34F missense mutations in *U2AF1*, a splicing factor that enables 3′ splice site identification, were associated with several splicing events in TCGA samples including alternative splicing of a known proto-oncogene *CTNNB1*. Recurrent *U2AF1* S34F mutations were previously reported by Imielinski and colleagues[10] in lung adenocarcinoma. These findings are consistent with previous reports of widespread cancer-specific alternative splicing events in lung adenocarcinoma.[75] Combined analysis of the whole genome sequencing and RNA-seq data of 19 NSCLC cell lines and three tumour samples by Liu and colleagues[75] showed 438 mutations affecting splice donor and acceptor sites (the first two or last two base pairs of an intron) in 433 genes. Most of these mutations were associated with altered splicing patterns. Further analysis helped the identification of 153 genes with differentially expressed splicing isoforms between cancer cell lines and normal lung tissue. These data suggest a potential role for deregulated alternate splicing in lung cancers.

	Frequency
Mutations	
KRAS	32·2%
EGFR	11·3%
NF1	8·3%
BRAF	7·0%
MET exon 14 skipping	4·3%
RIT1	2·2%
ERBB2	1·7%
HRAS, NRAS, MAP2K1	1·7%
Translocations	
ROS1	1·7%
ALK	1·3%
RET	0·9%
Amplifications	
MET	2·2%
ERBB2	0·9%

***Table 2*: Driver alterations in lung adenocarcinoma**

Subtyping based on gene expression

Expression profiling-based prognostic assays have been successfully applied to breast cancers and diffuse large B-cell lymphomas, providing important prognostic information and guiding treatment.[76–79] The expression subtypes of lung adenocarcinomas and their associated clinical and mutational features are concordant between different studies.[80,81] Based on these data, lung adenocarcinomas can be classified into bronchioid or terminal respiratory unit, magnoid or proximal-proliferative, and squamoid or proximal-inflammatory subtypes. Data from TCGA showed that terminal respiratory unit-subtype tumours are associated with a low mutation rate, low

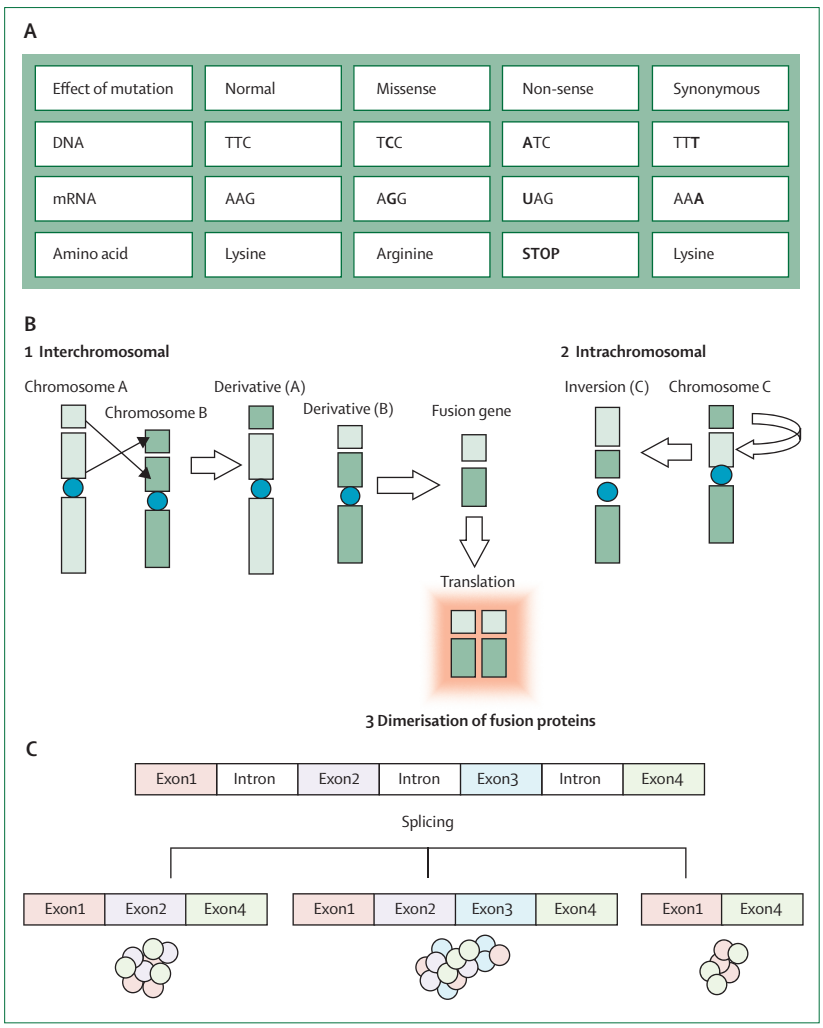

Figure 2: **Schematic representation of the various types of molecular alterations encountered in cancer cells and alternative splicing**

(A) The effects of different types of point mutations on RNA transcription and protein translation. (B) The formation of (1) derivative chromosomes through interchromosomal translocations and (2) chromosomal inversion intra-chromosomal translocations. These rearrangements can result in the formation of fusion proteins that result in constitutive growth signalling through abnormal dimerisation (3). (C) Alternative splicing resulting in the translation of three isoforms of the same protein.

ploidy, and better prognosis. Terminal respiratory unit-subtype tumours were also enriched for *EGFR* mutations and kinase fusions. RAS pathway activation was predominantly dependent on receptor tyrosine kinase activation in terminal respiratory unit tumours, loss of *NF1* in proximal-inflammatory tumours, and *KRAS* mutations in proximal-proliferative tumours. While comutation of *NF1* and *TP53* was frequent in proximal-inflammatory tumours, proximal-proliferative tumours showed a high frequency of *KRAS* mutation and *LKB1* inactivation.

Similar to the breast cancer model, in which gene expression profiling-based intrinsic subtyping (luminal A, luminal B, HER2-enriched, and basal-like subtypes) is prognostic and predictive of the efficacy of neoadjuvant chemotherapy and risk of disease relapse, it is worth exploring the use of molecular sub-classification in predicting outcomes after complete resection in early-stage lung adenocarcinoma in carefully designed studies.[82,83]

Epigenetic alterations

Chromatin modification and DNA methylation are well-known mechanisms of epigenetic regulation. Through deregulation of these epigenetic pathways, cancer cells aberrantly overexpress oncogenes and silence tumour suppressors.[84] The methylation status of several CpG sites across the genome can be investigated through the use of methylation-specific probes (eg, the HM450 assay includes probes for more than 480 000 CpG sites).[85] Investigators from TCGA[11] classified lung adenocarcinoma samples into various CpG island methylator phenotypes (CIMP) based on the extent of CpG methylation. Tumours were classified as either high CIMP, normal-like or low, or intermediate. The Wnt pathway was over-represented in genes that were hypermethylated, and underexpressed in CIMP-high tumours. Tumours showing *CDKN2A* hypermethylation were enriched for mutations in *SETD2*, a gene that regulates chromatin modification. CIMP-high tumours were also associated with an overexpression of *MYC*, although the underlying mechanism driving this association was unclear. A more thorough understanding of epigenetic alterations in lung adenocarcinoma will be crucial to identify patients with adenocarcinoma who might benefit from epigenetic therapies.

Intratumour heterogeneity and clonal evolution

Cancers are composed of related clones by virtue of their common ancestry.[86] The forces of natural selection constantly shape the evolution of these neoplastic clones. Cancer cells grow in hostile environments in which they compete for resources such as nutrition and oxygen, and must survive attacks by host immunity and antineoplastic treatments. Moreover, cancer cells also constantly accrue mutations that either confer survival advantages, or reduce their fitness compared to other non-neoplastic cells. The selection pressures to which cancer cells are exposed therefore shape their clonal architecture over time, and results in the emergence of cells with increased fitness.

The estimation of variant allele frequencies for different genes across many regions of a tumour (multiregion sequencing) aids the study of intratumoral heterogeneity and its clonal composition.[9,28,29,87] Recently, two independent groups reported substantial intratumoral heterogeneity within lung adenocarcinomas.[28,29] The mutations present in all tumour regions were classified as ubiquitous and those that were spatially restricted in their distribution were classified as heterogeneous. Ubiquitous alterations are acquired early in the course of tumour evolution and represent molecular events that characterise the founder clone. Alternatively, heterogeneous alterations represent events that occur later in the course of tumour evolution at a subclonal level. When the clonal structure of a tumour is

visually represented as a phylogenetic tree, the founder alterations (ubiquitous) map to the trunk, whereas subclonal (heterogeneous) events map to the branches. The clonal architecture of tumours can be either linear, in which progressively fitter clones replace the founder clone cells, or branched, in which several subclones coexist simultaneously.

Clonal analyses by de Bruin and colleagues,[28] and Zhang and colleagues,[29] suggest that mutations in known oncogenic drivers occur early in the course of lung cancer evolution and, in a few cases, possibly long before the development of clinical disease. An examination of the mutational patterns within lung cancers also suggests that smoking-related genomic events (characterised by cytosine to adenine transversions) occur early in the course of tumour evolution and that the proportion of these events decreases substantially with time despite maintained exposure to tobacco smoke. This decrease is accompanied by an increase in the proportion of cytosine to guanine and cytosine to thymidine mutations characteristic of apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like (APOBEC) family enzyme activity. de Bruin and colleagues[28] noted that about 31% of non-silent heterogenous mutations and 11% of ubiquitous mutations occur in a APOBEC-mediated context.

Characterisation of the driver molecular events in the founder clone and also the branch points in subclonally diverse tumours are crucial for guiding therapy. A study of the evolution patterns at relapse in acute myeloid leukaemia suggests that cancer relapses can occur either as a result of founder clone cells acquiring additional alterations that allow them to survive therapy or through the survival of therapy-resistant subclones that eventually evolve by gaining additional mutations.[88] These findings also suggest that intratumoral heterogeneity by itself can serve as a prognostic biomarker in patients with lung cancer. Zhang and colleagues[29] reported an increased chance of postsurgical relapse in patients with lung cancer whose tumours showed prominent subclonal diversity.

Potential novel targets

Next-generation sequencing can aid the identification of potential new drivers and targetable alterations in lung cancer.[8,9,11] Furthermore, clonal analysis suggests that some targetable alterations are enriched at the subclonal level, which has important implications for therapy and achievement of durable responses. Although drugs targeting genes such as *EGFR*, *ALK*, *ERBB2*, and *MET* are areas of intense ongoing research, the feasibility of targeting other novel molecular alterations warrants further study. We provide a few examples here.

KEAP1 and CUL3 have a key role in the oxidative stress response pathway and cause NRF2 (NFE2L2) degradation.[89–91] NFE2L2 is an important mediator of the oxidative stress response pathway that helps xenobiotic metabolism and drug efflux.[92] Inactivation of *KEAP1* and *CUL3* through mutation can result in the impaired regulation and stabilisation of NFE2L2. Deregulation of the oxidative stress response pathway has been implicated in the resistance of tumours to chemotherapy.[93] Alterations in this pathway have been reported in lung adenocarcinoma and squamous cell carcinoma.[31,91] *KEAP1, CUL3,* and *NF2EL2* mutations have been observed in 12%, 7% and 19% of squamous cell carcinomas and in 19%, less than 1%, and 3% of lung adenocarcinomas, respectively.[8,11] The potential role of all transretinioic acid and retinoic acid receptor α agonists, both capable of blocking NFE2L2-mediated transcription, and brusatol, which enhances NFE2L2 degradation, in tumours harbouring oxidative stress response pathway alterations requires further investigation.[94,95]

MYC-mediated regional transcriptional activation is characterised by acetylation of histones.[96] Histone acetylation promotes the assembly of transcriptional complexes with acetyl-lysine-binding domains or bromodomains. Members of the bromodomain and extra-terminal subfamily proteins BRD2, BRD3, and BRD4 play an important part in this process. A role for bromodomain inhibition has been shown in tumours harbouring *MYC* amplification.[96] Evidence from murine studies suggests MYC dependence in *KRAS* mutant lung cancers and sensitivity of these cancers to bromodomain inhibitors such as JQ1.[97] These data suggest a role for bromodomain inhibitors in *MGA*-mutated, *MYC*-amplified, and *KRAS*-mutated lung cancers.

Next-generation sequencing in the clinic

We reported the feasibility of doing next-generation sequencing in patients with NSCLC in a clinical setting.[98] Of 381 patient samples that were referred for targeted next-generation sequencing of genes that are known to be commonly altered in NSCLC, 209 (55%) were successfully sequenced. Excisional, endoscopic, and core biopsies yielded sufficient genetic material to carry out next-generation sequencing testing in 95%, 66%, and 40% of patients that underwent these procedures, respectively. The median time to laboratory results was 21 days (range 9–51), with a trend towards improvement of this time with more rapid sequencing platforms.

The Lung Cancer Mutation Consortium and the Clinical Lung Cancer Genome Project studies showed the feasibility of tumour genotyping to guide therapy in lung adenocarcinoma patients.[26,69] As part of the Lung Cancer Mutation Consortium, patients received matched targeted treatments after multiplexed testing for alterations in ten genes. Of 1007 patients who were screened for at least one genetic alteration, 275 (28%) tumours had actionable alterations that led to the use of targeted treatments. Median survival was longer in patients who received genotype directed treatments. Testing for key genomic alterations was also feasible in 3863 paraffin-embedded tumour samples in a molecular screening outreach programme run by the Clinical Lung Cancer Genome Project that prospectively

Search strategy and selection criteria

We searched PubMed for articles published in English between June 1, 2007, and Nov 1, 2014 using the search terms: "lung adenocarcinoma", "NSCLC", "genomics", and "next-generation sequencing". Articles were also identified through searches of the authors' own files. The final reference list was generated on the basis of relevance to the topic of the Review.

assessed 5145 lung cancer patients.[26] This information was used to guide matched therapy with approved and investigational drugs.

Several genomic-driven clinical trials are currently being designed to study investigational treatments. Some examples include the National Cancer Institute supported trials such as Lung-MAP (Master Protocol, NCT02154490), ALCHEMIST (Adjuvant Lung Cancer Enrichment Marker Identification and Sequencing Trials, NCT02194738), National Cancer Institute-MPACT (Molecular Profiling-Based Assignment of Cancer Therapy for patients with advanced solid tumours, NCT01827384), and MATCH (Molecular Analysis for Therapy Choice). Furthermore, unbiased characterisation of genomes from tumours that show exceptional responses to these therapies can also aid the identification of occult biomarkers that predict response to therapy (National Cancer Institute exceptional responder study, NCT02243592). Analyses of matched primary and metastatic tumour samples might improve our understanding of the molecular underpinning of metastatic process. Future studies should include comprehensive genomic characterisation of tumour specimens at the time of disease progression to fully understand the clonal evolution of tumours under treatment-induced selection pressure.

Contributors

All authors contributed equally.

Declaration of interests

DM is a speaker for Boehringer Ingelheim and is on the advisory board for Celgene. RG has received consultant fees and honoraria from Pfizer, Merck, Boehringer Ingelheim, Clovis, Helsinn Healthcare, Genentech, AbbVie, and GlaxoSmithKline. SD declares no competing interests.

References

1 Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. *CA Cancer J Clin* 2015; **65:** 5–29.

2 Travis WD. Pathology of lung cancer. *Clin Chest Med* 2011; **32:** 669–92.

3 Zhou C, Wu YL, Chen G, et al. Erlotinib versus chemotherapy as first-line treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer (OPTIMAL, CTONG-0802): a multicentre, open-label, randomised, phase 3 study. *Lancet Oncol* 2011; **12:** 735–42.

4 Maemondo M, Inoue A, Kobayashi K, et al, and the North-East Japan Study Group. Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR. *N Engl J Med* 2010; **362:** 2380–88.

5 Paez JG, Jänne PA, Lee JC, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* 2004; **304:** 1497–500.

6 Kwak EL, Bang YJ, Camidge DR, et al. Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N Engl J Med* 2010; **363:** 1693–703.

7 Ledford H. Big science: The cancer genome challenge. *Nature* 2010; **464:** 972–74.

8 Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* 2012; **489:** 519–25.

9 Govindan R, Ding L, Griffith M, et al. Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell* 2012; **150:** 1121–34.

10 Imielinski M, Berger AH, Hammerman PS, et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* 2012; **150:** 1107–20.

11 Network CGAR, and the Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* 2014; **511:** 543–50.

12 Zack TI, Schumacher SE, Carter SL, et al. Pan-cancer patterns of somatic copy number alteration. *Nat Genet* 2013; **45:** 1134–40.

13 Beroukhim R, Mermel CH, Porter D, et al. The landscape of somatic copy-number alteration across human cancers. *Nature* 2010; **463:** 899–905.

14 Weir BA, Woo MS, Getz G, et al. Characterizing the cancer genome in lung adenocarcinoma. *Nature* 2007; **450:** 893–98.

15 Li Y, Tergaonkar V. Noncanonical functions of telomerase: implications in telomerase-targeted cancer therapies. *Cancer Res* 2014; **74:** 1639–44.

16 Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000; **100:** 57–70.

17 Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2012; **2:** 401–04.

18 Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* 2013; **6:** pl1.

19 Yamaguchi T, Hosono Y, Yanagisawa K, Takahashi T. NKX2-1/TTF-1: an enigmatic oncogene that functions as a double-edged sword for cancer cell survival and progression. *Cancer Cell* 2013; **23:** 718–23.

20 Yuan B, Li C, Kimura S, Engelhardt RT, Smith BR, Minoo P. Inhibition of distal lung morphogenesis in Nkx2.1(-/-) embryos. *Dev Dyn* 2000; **217:** 180–90.

21 Morrisey EE, Hogan BL. Preparing for the first breath: genetic and cellular mechanisms in lung development. *Dev Cell* 2010; **18:** 8–23.

22 Rudin CM, Durinck S, Stawiski EW, et al. Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. *Nat Genet* 2012; **44:** 1111–16.

23 Bass AJ, Watanabe H, Mermel CH, et al. SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. *Nat Genet* 2009; **41:** 1238–42.

24 Meyer N, Penn LZ. Reflecting on 25 years with MYC. *Nat Rev Cancer* 2008; **8:** 976–90.

25 Patel JH, Loboda AP, Showe MK, Showe LC, McMahon SB. Analysis of genomic targets reveals complex functions of MYC. *Nat Rev Cancer* 2004; **4:** 562–68.

26 Clinical Lung Cancer Genome Project (CLCGP) Network Genomic Medicine (NGM). A genomics-based classification of human lung tumors. *Sci Transl Med* 2013; **5:** 209ra153.

27 Hurlin PJ, Huang J. The MAX-interacting transcription factor network. *Semin Cancer Biol* 2006; **16:** 265–74.

28 de Bruin EC, McGranahan N, Mitter R, et al. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science* 2014; **346:** 251–56.

29 Zhang J, Fujimoto J, Zhang J, et al. Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. *Science* 2014; **346:** 256–59.

30 Pleasance ED, Stephens PJ, O'Meara S, et al. A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature* 2010; **463:** 184–90.

31 Hammerman PS, Hayes DN, Wilkerson MD, et al, and the Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* 2012; **489:** 519–25.

32 Kandoth C, McLellan MD, Vandin F, et al. Mutational landscape and significance across 12 major cancer types. *Nature* 2013; **502:** 333–39.

33 Hodis E, Watson IR, Kryukov GV, et al. A landscape of driver mutations in melanoma. *Cell* 2012; **150:** 251–63.

34 Lawrence MS, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 2013; **499:** 214–18.

35 Peifer M, Fernández-Cuesta L, Sos ML, et al. Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. *Nat Genet* 2012; **44:** 1104–10.

36 Subramanian J, Govindan R. Lung cancer in never smokers: a review. *J Clin Oncol* 2007; **25:** 561–70.

37 Wennerberg K, Rossman KL, Der CJ. The Ras superfamily at a glance. *J Cell Sci* 2005; **118:** 843–46.

38 Pylayeva-Gupta Y, Grabocka E, Bar-Sagi D. RAS oncogenes: weaving a tumorigenic web. *Nat Rev Cancer* 2011; **11:** 761–74.

39 Ihle NT, Byers LA, Kim ES, et al. Effect of KRAS oncogene substitutions on protein behavior: implications for signaling and clinical outcome. *J Natl Cancer Inst* 2012; **104:** 228–39.

40 Garassino MC, Marabese M, Rusconi P, et al. Different types of K-Ras mutations could affect drug sensitivity and tumour behaviour in non-small-cell lung cancer. *Ann Oncol* 2011; **22:** 235–37.

41 Castellano E, Sheridan C, Thin MZ, et al. Requirement for interaction of PI3-kinase p110α with RAS in lung tumor maintenance. *Cancer Cell* 2013; **24:** 617–30.

42 Kim HS, Mendiratta S, Kim J, et al. Systematic identification of molecular subtype-selective vulnerabilities in non-small-cell lung cancer. *Cell* 2013; **155:** 552–66.

43 Ostrem JM, Peters U, Sos ML, Wells JA, Shokat KM. K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature* 2013; **503:** 548–51.

44 Sharma SV, Bell DW, Settleman J, Haber DA. Epidermal growth factor receptor mutations in lung cancer. *Nat Rev Cancer* 2007; **7:** 169–81.

45 Rosell R, Moran T, Queralt C, et al, and the Spanish Lung Cancer Group. Screening for epidermal growth factor receptor mutations in lung cancer. *N Engl J Med* 2009; **361:** 958–67.

46 Ohashi K, Maruvka YE, Michor F, Pao W. Epidermal growth factor receptor tyrosine kinase inhibitor-resistant disease. *J Clin Oncol* 2013; **31:** 1070–80.

47 Yasuda H, Kobayashi S, Costa DB. EGFR exon 20 insertion mutations in non-small-cell lung cancer: preclinical data and clinical implications. *Lancet Oncol* 2012; **13:** e23–31.

48 Camidge DR, Pao W, Sequist LV. Acquired resistance to TKIs in solid tumours: learning from lung cancer. *Nat Rev Clin Oncol* 2014; **11:** 473–81.

49 Sequist LV, Soria JC, Goldman JW, et al. Rociletinib in EGFR-mutated non-small-cell lung cancer. *N Engl J Med* 2015; **372:** 1700–09.

50 Jänne PA, Yang JC, Kim DW, et al. AZD9291 in EGFR inhibitor-resistant non-small-cell lung cancer. *N Engl J Med* 2015; **372:** 1689–99.

51 Chapman PB, Hauschild A, Robert C, et al, and the BRIM-3 Study Group. Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *N Engl J Med* 2011; **364:** 2507–16.

52 Marchetti A, Felicioni L, Malatesta S, et al. Clinical features and outcome of patients with non-small-cell lung cancer harboring BRAF mutations. *J Clin Oncol* 2011; **29:** 3574–79.

53 Paz-Ares LG, Biesma B, Heigener D, et al, and the NSCLC [non–small-cell lung cancer] Research Experience Utilizing Sorafenib (NExUS) Investigators Study Group. Phase III, randomized, double-blind, placebo-controlled trial of gemcitabine/cisplatin alone or with sorafenib for the first-line treatment of advanced, nonsquamous non-small-cell lung cancer. *J Clin Oncol* 2012; **30:** 3084–92.

54 Planchard D, Mazieres J, Riely GJ, et al. Interim results of phase II study BRF113928 of dabrafenib in BRAF V600E mutation-positive non-small cell lung cancer (NSCLC) patients. ASCO Meeting Abstracts. *J Clin Oncol* 2013; **31:** 8009 (abstr).

55 Seo JS, Ju YS, Lee WC, et al. The transcriptional landscape and mutational profile of lung adenocarcinoma. *Genome Res* 2012; **22:** 2109–19.

56 Camidge DR, Ou SHI, Shapiro G, et al. Efficacy and safety of crizotinib in patients with advanced *c-MET*-amplified non-small cell lung cancer (NSCLC). *J Clin Oncol* 2014; **32** (suppl 5s)**:** 8001 (abstr).

57 Ou SH, Kwak EL, Siwak-Tapp C, et al. Activity of crizotinib (PF02341066), a dual mesenchymal-epithelial transition (MET) and anaplastic lymphoma kinase (ALK) inhibitor, in a non-small cell lung cancer patient with de novo MET amplification. *J Thorac Oncol* 2011; **6:** 942–46.

58 Pao W, Girard N. New driver mutations in non-small-cell lung cancer. *Lancet Oncol* 2011; **12:** 175–80.

59 Mazières J, Peters S, Lepage B, et al. Lung cancer that harbors an HER2 mutation: epidemiologic characteristics and therapeutic perspectives. *J Clin Oncol* 2013; **31:** 1997–2003.

60 Arcila ME, Chaft JE, Nafa K, et al. Prevalence, clinicopathologic associations, and molecular spectrum of ERBB2 (HER2) tyrosine kinase mutations in lung adenocarcinomas. *Clin Cancer Res* 2012; **18:** 4910–18.

61 Stephens P, Hunter C, Bignell G, et al. Lung cancer: intragenic ERBB2 kinase mutations in tumours. *Nature* 2004; **431:** 525–26.

62 De Grève J, Teugels E, Geers C, et al. Clinical activity of afatinib (BIBW 2992) in patients with lung adenocarcinoma with mutations in the kinase domain of HER2/neu. *Lung Cancer* 2012; **76:** 123–27.

63 Shimamura T, Ji H, Minami Y, et al. Non-small-cell lung cancer and Ba/F3 transformed cells harboring the ERBB2 G776insV_G/C mutation are sensitive to the dual-specific epidermal growth factor receptor and ERBB2 inhibitor HKI-272. *Cancer Res* 2006; **66:** 6487–91.

64 Li D, Ambrogio L, Shimamura T, et al. BIBW2992, an irreversible EGFR/HER2 inhibitor highly effective in preclinical lung cancer models. *Oncogene* 2008; **27:** 4702–11.

65 Gómez-Seguí I, Makishima H, Jerez A, et al. Novel recurrent mutations in the RAS-like GTP-binding gene RIT1 in myeloid malignancies. *Leukemia* 2013; **27:** 1943–46.

66 Berger AH, Imielinski M, Duke F, et al. Oncogenic RIT1 mutations in lung adenocarcinoma. *Oncogene* 2014; **33:** 4418–23.

67 Lawrence MS, Stojanov P, Mermel CH, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* 2014; **505:** 495–501.

68 Shaw AT, Kim DW, Mehra R, et al. Ceritinib in ALK-rearranged non-small-cell lung cancer. *N Engl J Med* 2014; **370:** 1189–97.

69 Shaw AT, Kim DW, Nakagawa K, et al. Crizotinib versus chemotherapy in advanced ALK-positive lung cancer. *N Engl J Med* 2013; **368:** 2385–94.

70 Kris MG, Johnson BE, Berry LD, et al. Using multiplexed assays of oncogenic drivers in lung cancers to select targeted drugs. *JAMA* 2014; **311:** 1998–2006.

71 Shaw AT, Ou SH, Bang YJ, et al. Crizotinib in ROS1-rearranged non-small-cell lung cancer. *N Engl J Med* 2014; **371:** 1963–71.

72 Gainor JF, Shaw AT. Novel targets in non-small cell lung cancer: ROS1 and RET fusions. *Oncologist* 2013; **18:** 865–75.

73 Vaishnavi A, Capelletti M, Le AT, et al. Oncogenic and drug-sensitive NTRK1 rearrangements in lung cancer. *Nat Med* 2013; **19:** 1469–72.

74 Zhang J, Manley JL. Misregulation of pre-mRNA alternative splicing in cancer. *Cancer Discov* 2013; **3:** 1228–37.

75 Liu J, Lee W, Jiang Z, et al. Genome and transcriptome sequencing of lung cancers reveal diverse mutational and splicing events. *Genome Res* 2012; **22:** 2315–27.

76 van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002; **415:** 530–36.

77 Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 2004; **351:** 2817–26.

78 Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000; **403:** 503–11.

79 Dunleavy K, Roschewski M, Wilson WH. Precision treatment of distinct molecular subtypes of diffuse large B-cell lymphoma: ascribing treatment based on the molecular phenotype. *Clin Cancer Res* 2014; **20:** 5182–93.

80 Wilkerson MD, Yin X, Walter V, et al. Differential pathogenesis of lung adenocarcinoma subtypes involving sequence mutations, copy number, chromosomal instability, and methylation. *PLoS One* 2012; **7:** e36530.

81 Hayes DN, Monti S, Parmigiani G, et al. Gene expression profiling reveals reproducible human lung adenocarcinoma subtypes in multiple independent patient cohorts. *J Clin Oncol* 2006; **24:** 5079–90.

82 Parker JL, Lushina N, Bal PS, Petrella T, Dent R, Lopes G. Impact of biomarkers on clinical trial risk in breast cancer. *Breast Cancer Res Treat* 2012; **136:** 179–85.

83 Senkus E, Kyriakides S, Penault-Llorca F, et al, and the ESMO Guidelines Working Group. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2013; **24** (suppl 6)**:** vi7–23.

84 Dawson MA, Kouzarides T. Cancer epigenetics: from mechanism to therapy. *Cell* 2012; **150:** 12–27.

85 Wong NC, Ng J, Hall NE, et al. Exploring the utility of human DNA methylation arrays for profiling mouse genomic DNA. *Genomics* 2013; **102:** 38–46.

86 Yates LR, Campbell PJ. Evolution of the cancer genome. *Nat Rev Genet* 2012; **13:** 795–806.

87 Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 2012; **366:** 883–92.

88 Ding L, Ley TJ, Larson DE, et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* 2012; **481:** 506–10.

89 DeNicola GM, Karreth FA, Humpton TJ, et al. Oncogene-induced Nrf2 transcription promotes ROS detoxification and tumorigenesis. *Nature* 2011; **475:** 106–09.

90 Shibata T, Ohta T, Tong KI, et al. Cancer related mutations in NRF2 impair its recognition by Keap1-Cul3 E3 ligase and promote malignancy. *Proc Natl Acad Sci USA* 2008; **105:** 13568–73.

91 Singh A, Misra V, Thimmulappa RK, et al. Dysfunctional KEAP1-NRF2 interaction in non-small-cell lung cancer. *PLoS Med* 2006; **3:** e420.

92 Hur W, Sun Z, Jiang T, et al. A small-molecule inducer of the antioxidant response element. *Chem Biol* 2010; **17:** 537–47.

93 Wang XJ, Sun Z, Villeneuve NF, et al. Nrf2 enhances resistance of cancer cells to chemotherapeutic drugs, the dark side of Nrf2. *Carcinogenesis* 2008; **29:** 1235–43.

94 Wang XJ, Hayes JD, Henderson CJ, Wolf CR. Identification of retinoic acid as an inhibitor of transcription factor Nrf2 through activation of retinoic acid receptor alpha. *Proc Natl Acad Sci USA* 2007; **104:** 19589–94.

95 Ren D, Villeneuve NF, Jiang T, et al. Brusatol enhances the efficacy of chemotherapy by inhibiting the Nrf2-mediated defense mechanism. *Proc Natl Acad Sci USA* 2011; **108:** 1433–38.

96 Delmore JE, Issa GC, Lemieux ME, et al. BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell* 2011; **146:** 904–17.

97 Shimamura T, Chen Z, Soucheray M, et al. Efficacy of BET bromodomain inhibition in Kras-mutant non-small cell lung cancer. *Clin Cancer Res* 2013; **19:** 6183–92.

98 Hagemann IS, Devarakonda S, Lockwood CM, et al. Clinical next-generation sequencing in patients with non-small cell lung cancer. *Cancer* 2015; **121:** 631–39.