Mutational Landscape of DDR2 Gene in Lung Squamous Cell Carcinoma Using Next-generation Sequencing

Charles Ricordel, Alexandra Lespagnol, Francisco Llamas-Gutierrez, Marie de Tayrac, Mallorie Kerjouan, Alice Fievet, Houda Hamdi-Rozé, Amyrat Aliouat, Benoit Desrues, Jean Mosser, et al.

To cite this version:

Charles Ricordel, Alexandra Lespagnol, Francisco Llamas-Gutierrez, Marie de Tayrac, Mallorie Kerjouan, et al.. Mutational Landscape of DDR2 Gene in Lung Squamous Cell Carcinoma Using Next-generation Sequencing. Clinical Lung Cancer, Elsevier, 2018, 19 (2), pp.163-169.e4. 10.1016/j.cllc.2017.10.006 . hal-01739958

HAL Id: hal-01739958

https://hal-univ-rennes1.archives-ouvertes.fr/hal-01739958

Submitted on 13 Apr 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Mutational landscape of DDR2 gene in lung squamous cell carcinoma using next-generation sequencing

C. Ricordel\textsuperscript{1,2}, A. Lespagnol\textsuperscript{3}, F. Llamas-Gutierrez \textsuperscript{4}, M. de Tayrac \textsuperscript{3}, M. Kerjouan\textsuperscript{1}, A. Fievet\textsuperscript{3}, H. Hamdi-Rozé\textsuperscript{3}, A. Aliouat\textsuperscript{5}, B. Desrues\textsuperscript{1,2}, J. Mosser\textsuperscript{3,5}, H. Léna\textsuperscript{1,2}

\textsuperscript{1} Department of Respiratory Medicine, Pontchaillou Hospital, Rennes 1 University, 2 rue Henri Le Guilloux, 35033 Rennes cedex 9, France
\textsuperscript{2} Chemistry, Oncogenesis and Stress Signaling, INSERM U1242, Centre Eugène Marquis, Avenue bataille Flandres Dunkerque, 35042 Rennes, France
\textsuperscript{3} Department of Molecular Genetics and Genomics, Pontchaillou Hospital, Rennes 1 University, 2 rue Henri Le Guilloux, 35033 Rennes cedex 9, France
\textsuperscript{4} Department of Pathology, Pontchaillou Hospital, Rennes 1 University, 2 rue Henri Le Guilloux, 35033 Rennes cedex 9, France
\textsuperscript{5} Functional Integrated Genomics & Biomarkers, IGDR, UMR6290, CNRS, Rennes 1 University, 35043 Rennes cedex, France

Word count micro-abstract: 60
Word count abstract: 230
Word count text: 3015
Figures: 3
Tables: 3
Appendix: 5

Correspondence to:
Charles Ricordel
Service de pneumologie, CHU de Rennes, 2 rue Henri-Le-Guilloux, 35033 Rennes cedex 9
Tel: 02.99.28.24.78 – Fax: 02.99.28.24.80
charles.ricordel@chu-rennes.fr
Financial support: None
Conflict of interest statement: None

Keywords: Discoidin domain receptor 2, Squamous cell lung carcinoma, Somatic mutation, Molecular biology, lung neoplasms
Micro-abstract:

*DDR2* alterations were identified as a promising therapeutic target in lung SCC. Using a large cohort, *DDR2* genetic landscape was assessed by next-generation-sequencing. *DDR2* mutations are present in 4% of SCC, are not exclusive with driver genes alterations, and have no prognosis impact. This study emphasizes the need for a better knowledge of DDR2 biology before developing dedicated targeted therapies.
Abstract:

Background: Lung cancer represents the leading cause of cancer-related death worldwide. Despite great advances in lung cancer management with the recent emergence of molecular targeted therapies for non-squamous non-small cell lung cancer, no dramatic improvements have been achieved in lung squamous cell carcinoma (SCC). Mutations in DDR2 gene were recently identified as promising molecular targets in this histology. The aim of this study is to describe the DDR2 mutational landscape of lung SCC and investigate the associated clinical factors.

Methods: Next generation sequencing of the DDR2 gene was performed on 271 samples of lung SCC. Patients followed in our institution from January 2011 to August 2014 were retrospectively selected for data collection. Other driver gene alterations (EGFR, KRAS, BRAF, HER2 and PI3KCA) were analyzed using pyrosequencing.

Results: A total of 11 patients harboring a DDR2 mutation was detected among the 271 sequenced lung SCC samples (4%). We describe 10 unreported mutations, comprising a novel DDR2 exon 7 splice mutant. DDR2 mutations were not mutually exclusive with other driver gene alterations. One hundred thirty-six patients were included for clinical comparison and logistic regression analysis. No difference was detected between DDR2-mutant and DDR2 wild-type lung SCC regarding clinical characteristics or survival.

Conclusion: DDR2 mutations were observed in 4% of lung SCC cases of European descent. DDR2-mutated tumors can exhibit other driver gene alterations. No clinical characteristics were significantly associated with DDR2 mutation.
Introduction

Lung cancer is the leading cause of cancer-related death worldwide, accounting for more than 1.5 million deaths in 2012.\textsuperscript{1} Non-small cell lung cancers (NSCLCs) represent approximately 85\% of lung neoplasms; among these, squamous cell carcinomas (SCC) account for approximately 30\% of cases. The landscape of thoracic oncology has greatly changed during the last decade with the discovery of tumor driver mutations predicting response to targeted therapies.\textsuperscript{2} Recent clinical trials demonstrated an unprecedented improvement in progression-free survival (PFS) and quality of life, especially in molecularly selected lung adenocarcinoma.\textsuperscript{3–5} Lung SCC lacks available molecular targets, and tyrosine kinase inhibitors are mostly ineffective in unselected populations\textsuperscript{6} despite many efforts made to understand its genomics.\textsuperscript{7} Among all of the somatic alterations explored to date in lung squamous histology, Discoidin Domain Receptor 2 (\textit{DDR2}) gene mutations are reported in approximately 3\% of cases.\textsuperscript{8,9} DDR2 is a tyrosine kinase receptor that belongs to the same family as Epidermal Growth Factor Receptor (EGFR).\textsuperscript{10} Numerous \textit{in vitro} studies on various tumoral cell lines demonstrated the important role of DDR2 in the regulation of cellular proliferation, migration, metastasis, and secretion of matrix metalloproteinase.\textsuperscript{11} In 2011, Hammerman \textit{et al} showed that a subset of \textit{DDR2} mutants are oncogenic in SCC cell lines \textit{in vitro}, delivering a strong rational for targeting \textit{DDR2} mutations.\textsuperscript{8} In addition, two patients with DDR2-mutated lung SCC were reported in the literature to experience a dramatic response to dasatinib, a potent Src inhibitor.\textsuperscript{8,12} Altogether, these data establish \textit{DDR2} mutation as a promising molecular target of tyrosine kinase inhibitors in lung SCC.

Few studies are available concerning the clinical characteristics of patients harboring \textit{DDR2} mutations. The first aim of this study was to describe \textit{DDR2} gene mutations in a large monocentric cohort of SCC. Further objectives of the study were to compare the clinical characteristics of \textit{DDR2}-mutant to \textit{DDR2} wild-type (WT) lung SCC and to investigate factors associated with \textit{DDR2} mutations.
Patients and methods

Tissues and patients

Samples were obtained from patients who were diagnosed with SCC by tissue biopsy or primary surgical resection from January 2011 to January 2015 in five hospital centers in Brittany (north-west of France). A total of 271 lung SCC samples were retrospectively enrolled for DDR2 gene sequencing on a centralized platform. For demographics, clinical and survival comparisons, we chose to only select patients diagnosed and followed from January 2011 to August 2014 in Rennes University Hospital to avoid site-dependent care bias and limit retrospective data missing in patient’s record. The choice of this recruiting period allows a minimal follow-up of at least 12 months for all patients. Subsequently, a total of 136 patients with lung SCC composed the investigation cohort (appendix 1). Central localization of the tumor was defined as proximal to the third bronchial division and advanced disease as unresectable stage IIIB or stage IV tumors according to the 7th IASLC TNM classification for lung cancer. Concerning the smoking exposure, patients were classified as active smoker, former smoker (no exposition during the last 12 months at diagnosis), never-smoker or unknown. The study was approved by the Rennes University Hospital Ethics Committee (n° 15.87).

DDR2 gene sequencing, sequence processing, variant calling and analysis of DDR2 mutations

All tumor tissues were subjected to review by a pathologist to estimate the tumor cell content and 10 µM-thick formalin-fixed paraffin-embedded tissue sections were sampled. A tumor cell cut-off of at least 10% was used to proceed further with the sample. Normal tissue was also sampled when matched non-tumoral sequencing was indicated. Genomic DNA was extracted from routine FFPE samples using the MagDEA DNA 200 kit (Precision System Science®, Japan). Genomic DNAs from all samples were quantified with the Quan-iT PicoGreen dsDNA assay kit (Invitrogen). A panel of 36 specific primers targeting the DDR2 gene was created using AmplifX (see appendix 4 for sequence). All primers were validated by single-plex PCR and assessment of PCR products for expected size was performed on LabChip Gx (Perkin Elmer®). The experiments were performed according to the Multiplex Amplicon Tagging Protocol from the manufacturer (Fluidigm®) in three major steps: (i) a
multiplex specific target pre-amplification, (ii) a target enrichment PCR on the Access Array IFC and (iii) a barcode PCR. All PCRs were performed using FastStart High Fidelity PCR System, dNTPack (Roche®). The libraries’ quality and concentration were evaluated on a LabChip GX (Perkin Elmer®). Next, libraries were pooled together to create one PCR product library. The resulting sequencing-ready amplicon libraries were sequenced on a MiSeq sequencer (Illumina) using MiSeq Reagent Kit v2 (300 cycles). Each patient was sequenced twice with independent amplicon library preparation. The bioinformatics pipeline to detect low frequency variants developed for molecular testing in a diagnosis purpose was used to detected alterations in the DDR2 gene. Briefly, reads were aligned to the human reference genome GRCh37 (hg19) using BWA software (bwa-sw 0.5.9), recalibrated and realigned using GATK (GATK-2.1-11). NGS amplicon reads were split into non-overlapping bam files. These bam files were used to detect single-nucleotide variants and insertions/deletions based on SAMtools mpileup files (samtools-0.1.18). Custom scripts were used to identify sequence variants in the pileup files. For each patient at each variant position, variant allele fractions (VAFs) were estimated separately on forward and reverse reads, by amplicon and by index. When calling the variants, the minimum read depth was set to 150 in forward and reverse reads and in both indexes (i.e. a total read depth ≥ 600) and the variant allele must be present in > 10% of both reads. Variants were annotated with ANNOVAR and selected when present inside exons or at canonical splice sites. The genome version used was grch37/hg19 (February 2009) and the Refseq for DDR2 gene was NM_006182.3. Integrative Genomics Viewer (IGV) was used to visualize the read alignment and to confirm the variant calls. If the DDR2 variant allele was >50%, systematic DDR2 gene sequencing of normal tissue was performed to confirm the somatic status of the nucleotide variation. Potential protein-altering mutations were analyzed using CRAVAT 4.3 software.14,15 Pre-mRNA splicing potential defects were studied using the Human Splicing Finder 3.0 website.16

Detection of EGFR, PI3KCA, KRAS, BRAF and HER2 mutations

Genomic DNA was isolated from FFPE tissue specimen using the MagDEA DNA 200 kit (Precision System Science, Japan) according to the manufacturer’s instruction. DNA samples were quantified using a Nanodrop spectrophotometer and normalized to 10 ng/µl. Mutations of EGFR (exons 18–21),
HER2 (exon 20), BRAF (exon 15), KRAS (exon 2), and PIK3CA (exons 9 and 20) were detected on a routine basis on the biomolecular platform of Rennes University Hospital, using respective Refseq:

- EGFR: NM_005228.3
- HER2: NM_004448.3
- BRAF: NM_004333.4
- KRAS: NM_033360.3
- PIK3CA: NM_006218.2

The methodology was based on pyrosequencing assays, as previously described. ALK and FGFR genes fusions were not included in the analysis as mainly not available in our cohort (not routinely assessed at the time of inclusion).

**Statistical analysis**

Demographic and descriptive data are given as the median with the range. Categorical variables were compared with the Fischer exact test or Pearson chi-squared test, and quantitative variables were compared with the Mann-Whitney U test when appropriate. Overall survival was calculated in months from the date of diagnosis to the date of death (failure) or the date the patient was last known to be alive (censored). Two-tailed p-values were reported, with p<0.05 considered as statistically significant. The Kaplan-Meier method with log-rank test were used to perform survival analysis. Finally, a binomial logistic regression model was built with the DDR2 status (mutant or wild-type) as the dependent variable. GraphPad Prism® for Windows software (version 5.03; Graph Pad Software Inc., Los Angeles) was used for the statistical and survival analyses.
Results

Frequency, distribution and characteristics of lung SCC DDR2 mutation

Next generation sequencing was performed on 271 samples of lung SCC. Percentage of DDR2 gene coding sequence coverage was 100%, with a minimum of 600X depth for each sample (see appendix 5). DDR2 mutations were found in 4% (11/271) of the cases. Table 1 summarizes DDR2 mutations characteristics. The majority of DDR2 point mutations observed were nucleotide transversions (73%).

Prediction of the functional impact on the DDR2 protein reveals that four mutations had a neutral effect whereas seven showed a deleterious effect. DDR2 mutations predominantly map onto the discoidin homology domains and the kinase domain, each comprising 36% of the mutations (Figure 1A). Only one DDR2 mutation (c.1857G>C, p.R619S) described in this study was previously reported in lung SCC. A new DDR2 splice mutation was identified: c.566-1G>C. This substitution occurring in the late intron upstream of exon 7 of the DDR2 gene alter the wild-type acceptor site. A new splicing acceptor site is predicted at position c.564 resulting in a reading frame shift, as schematized in Figure 1B. Of note, single nucleotide polymorphism (SNP) of the DDR2 gene was found in seventeen patients in our cohort (according to public database or paired sequenced non-tumoral tissue) (see appendix 2).

Mutational status of EGFR, PI3KCA, KRAS, HER2 and BRAF genes

In our panel, the mutational status for other known driver genes was routinely assessed for 265 patients. Seventeen patients had a mutation in one of these genes, including five patients with an Exon 2 c.34G>T KRAS mutation, one patient with an exon 20 c.2338C>A HER2 mutation, eleven patients with a mutation PI3KCA (ten with exon 10 c.1624G>A and one with exon 21 c.3140A>G). These results are summarized in Table 2. One tumor exhibits a concomitant deleterious p.G575R DDR2 mutation and a p.G12D KRAS mutation.

Investigation cohort characteristics
Demographic and descriptive statistics of the DDR2-mutant and the DDR2-WT groups of the investigation cohort are listed in Table 3 and clustered genomic profile of the investigation cohort is presented in figure 2. The two groups were composed of male heavy smokers (81% and 93% tobacco exposure in the DDR2-mutant and the DDR2-WT groups, respectively) with a high level of comorbidities, mainly cardiovascular disease, in approximately two-thirds of the cases in the entire cohort. No differences in age, gender, smoking history, pack-year, comorbidities, Eastern Cooperative Oncology Group (ECOG) performance status, disease stage or follow-up were observed between the two groups.

Factors associated with DDR2 mutation and prognosis impact of DDR2 mutations

Performing logistic regression, we investigated factors associated with DDR2 gene mutational status. In univariate analysis, no variables were significantly associated with DDR2 mutant lung SCC (see appendix 3). Global cohort survival analysis using the Kaplan-Meier estimator showed no difference between DDR2-mutant lung SCC and DDR2-WT lung SCC (p=0.71; HR : 0.52 CI 95% [0.25-1.08]) (Figure 3A). A subgroup comparison of advanced DDR2-mutant lung SCC with advanced DDR2-WT lung SCC revealed no difference concerning overall survival (p=0.77; HR : 1.66 CI 95% [0.75-3.71]) (Figure 3B) or progression-free survival (p = 0.65; HR: 0.76 CI 95% [0.28 to 2.08]) (data not shown).
Discussion

In this study, we provide one of the largest description of DDR2 genetic landscape in lung SCC to date. We identified eleven DDR2 mutations (4%) and report for the first time, a splicing mutant c.566-1G>C of the DDR2 gene in lung cancer. The following rates of mutation were previously observed: 2% (2/100), 3.2% (9/277), 4.6% (4/86), 1.1% (2/178), and 1.3% (3/178) in the lung SCC cohorts of Lee et al\textsuperscript{19}, Hammerman et al\textsuperscript{8}, Miao et al\textsuperscript{20} and cBioportal database (TCGA\textsuperscript{7} and TCGA provisional) respectively. In contrast, Kenmotsu et al\textsuperscript{21} (sequencing limited to p.S768R mutation), and Hashima et al\textsuperscript{22} (sequencing of 8 exons of DDR2 gene) failed to identify any DDR2 mutation in their Japanese cohort comprising more than one hundred cases. Of note, sequencing coverage of DDR2 gene is very different between these studies and may contribute to such prevalence inconsistency. Some authors have suggested that DDR2 mutations may be related to tobacco exposure.\textsuperscript{23} To support this notion, the largest percentage of DDR2 mutations described in this report are transversions, known to be smoking-related.\textsuperscript{24} Epidemiologic studies showing that East Asian populations are less susceptible to smoking-related lung cancer\textsuperscript{25} might also partially explain these geographic variations.

We identified a new DDR2 truncating mutation p.E85X that is predicted to have a major impact on the receptor function. Nevertheless, to date, some discrepancies exist in the literature concerning the oncogenic potential induced by DDR2 aberrant signaling. Functional validations assays are needed to assess the oncogenic properties of this variant. Moreover, DDR2 mutations appear not to be exclusive from other driver gene alterations. This is supported by the observation of co-occurrence of such mutations with a KRAS p.G12D mutation in our cohort. However, known-driver mutations frequencies appear lower in our cohort than expected, comparing to TCGA\textsuperscript{7}. This could be explained by our methodology based on hotspot exon sequencing which contrary to TCGA, do not cover all gene coding-sequence. In their founding publication, Hammerman et al demonstrated that ectopic expression of a subset of DDR2 mutants (p.L63V and p.I638F) in non tumoral cells is able to promote proliferation \textit{in vitro} in a DDR2-dependent manner.\textsuperscript{8} These results were strengthened by \textit{in vivo} experiments with a recent publication revealing that inducible DDR2 p.L63V mutation combined with
conditional loss of TP53 is oncogenic in a murine model. However, a recent publication showed the presence of DDR2 p.L63V and p.G505S alterations not only in tumor specimens but also in matched normal tissues, suggesting that these variants are inherited. Some of these variants, also identified in our cohort, have already been reported in the NHLBI Exome Sequencing Project but misclassified as mutations in the literature. Further research should now investigate to what extent these constitutional DDR2 variants could be associated with lung SCC initiation; similar to the ongoing research on EGFR germline mutations. A better deciphering of DDR2 somatic mutations by deep-sequencing is now critical to better select patients in clinical trials, as new potent and selective DDR2 inhibitors are currently in development.

Some limitations to our study should be noted. First, the sample size of our study was limited. As DDR2 mutations are rare in lung SCC, it is possible that the power of our study was insufficient to detect potential associated factors. Second, a significant proportion of the 271 sequenced patients had to be withdrawn from the data collection to avoid missing information. Consequently, potential selection bias could have influenced our findings. Third, this study was monocentric, and reproducibility of the molecular results could not be assessed on multiple genomic platforms. Fourth, we cannot definitively confirmed the non-germinal status of one of the new p.P492S DDR2 variant described in this cohort, as samples with healthy tissue were not available for this patient.

Conclusion

In conclusion, DDR2 mutations were detected in 4% of cases in our lung SCC cohort. Our study identified ten new DDR2 mutations that have not been described to date, including a unique DDR2 splice mutant and shows that DDR2-mutated tumors can exhibit other driver gene alterations. No clinical characteristics or survival difference was detected between DDR2-mutant and DDR2-WT lung SCC, and no clinical factors were significantly associated with DDR2 mutation. A better understanding of DDR2 biology and its mutants will be critical for developing future dedicated therapy in molecularly selected DDR2-driven tumors.

Clinical practice points
- DDR2 mutations occur in 4% of squamous cell lung cancer
- DDR2-mutated tumors are not mutually exclusive with other driver gene alterations
- DDR2 gene mutations are not associated with specific clinical characteristics and show no prognosis impact
- Further studies deciphering DDR2 biology in lung SCC are warranted for developing future dedicated targeted therapies.
Figure and Tables legends:

Table 1. Overview of the 11 DDR2 somatic mutations observed in a panel of 271 lung SCCs. (ID: Identification Number; Refseq: NCBI Reference Sequence Database; Ref: reference acid nucleic; AA: aminoacid; VAFs: Variant Allele Frequencies)

Figure 1. (A) Distribution of mutations in the domain structure of the DDR2 protein. Novel mutations are listed in the upper part of the figure. IXJM: intra-cellular juxta-membrane domain. EXJM: extra-cellular juxta-membrane domain. TM: trans-membrane domain. Mutations also reported as SNP in literature are indicated with *. Mutations reported predictive to clinical response on DDR2- inhibiting drugs are noted with †. (B) Diagram of DDR2 exon 7 alteration in relation to the 5’ splice site.

Table 2. Proportion of driver gene alterations in 271 lung SCCs.

Table 3. Baseline characteristics of the investigation cohort.

Figure 2. Heatmap of the genomic landscape clustered by DDR2 mutational status (investigation cohort). Each column represent a patient (n=136).

Figure 3. Kaplan-Meier analysis. (A) Global cohort survival of the DDR2-mutant (n = 11) and DDR2-WT groups (n=125). (B) Overall survival in the advanced DDR2-mutant (n = 7) and DDR2-WT (n = 48) lung SCC subgroups. Log-rank test results are shown on the graphs along with the number of patients at risk during follow-up.
List of Appendix

Appendix 1. Flow Chart of the study

Appendix 2. Annotation of the identified DDR2 SNV not called as mutation in the validation cohort

Appendix 3. Logistic regression analysis (DDR2 as the dependent variable). Univariate analysis.

Appendix 4: Sequences of the primers used for DDR2 gene sequencing and mean coverage achieved for each primer.

Appendix 5: Sequencing coverage of DDR2 gene
Author contributions

AL, HL, JM, BD and CR contributed to conception and design. CR, AL, HL and BD were responsible for the literature review. FL was responsible of pathology reviewing. CR and MK were responsible for figures. CR was responsible for clinical data collection. AL, MDT, AF, AA, HH and JM were responsible for molecular biology data collection, quality control and analysis. CR and MDT were responsible for statistical analysis. All authors were responsible for manuscript editing and final approval of the article. CR takes responsibility for the paper as a whole.

Acknowledgment

Authors want to thank Dr. D. Doultsinos for valuable comments and suggestions to improve the quality of the paper.

Conflict of interest statement related to this work: None
Bibliography

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. _CA Cancer J Clin_. 2015;65(2):87-108. doi:10.3322/caac.21262.

2. Thunnissen E, van der Oord K, den Bakker M. Prognostic and predictive biomarkers in lung cancer. A review. _Virchows Arch Int J Pathol_. 2014;464(3):347-358. doi:10.1007/s00428-014-1535-4.

3. Mok TS, Wu Y-L, Ahn M-J, et al. Osimertinib or Platinum-Pemetrexed in EGFR T790M-Positive Lung Cancer. _N Engl J Med_. December 2016. doi:10.1056/NEJMo1612674.

4. Solomon BJ, Mok T, Kim D-W, et al. First-line crizotinib versus chemotherapy in ALK-positive lung cancer. _N Engl J Med_. 2014;371(23):2167-2177. doi:10.1056/NEJMo1408440.

5. Shaw AT, Ou S-HI, Bang Y-J, et al. Crizotinib in ROS1-rearranged non-small-cell lung cancer. _N Engl J Med_. 2014;371(21):1963-1971. doi:10.1056/NEJMo1406766.

6. Garassino MC, Martelli O, Broggini M, et al. Erlotinib versus docetaxel as second-line treatment of patients with advanced non-small-cell lung cancer and wild-type EGFR tumours (TAILOR): a randomised controlled trial. _Lancet Oncol_. 2013;14(10):981-988. doi:10.1016/S1470-2045(13)70310-3.

7. Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. _Nature_. 2012;489(7417):519-525. doi:10.1038/nature11404.

8. Hammerman PS, Sos ML, Ramos AH, et al. Mutations in the DDR2 kinase gene identify a novel therapeutic target in squamous cell lung cancer. _Cancer Discov_. 2011;1(1):78-89. doi:10.1158/2159-8274.CD-11-0005.

9. König K, Peifer M, Fassunke J, et al. Implementation of Amplicon Parallel Sequencing Leads to Improvement of Diagnosis and Therapy of Lung Cancer Patients. _J Thorac Oncol_. 2015;10(7):1049-1057. doi:10.1097/JTO.0000000000000570.

10. Fu H-L, Valiathan RR, Arkwright R, et al. Discoidin domain receptors: unique receptor tyrosine kinases in collagen-mediated signaling. _J Biol Chem_. 2013;288(11):7430-7437. doi:10.1074/jbc.R112.444158.

11. Valiathan RR, Marco M, Leitinger B, Kleer CG, Fridman R. Discoidin domain receptor tyrosine kinases: new players in cancer progression. _Cancer Metastasis Rev_. 2012;31(1-2):295-321. doi:10.1007/s10555-012-9346-z.

12. Pitini V, Arrigo C, Di Mirto C, Mondello P, Altavilla G. Response to dasatinib in a patient with SQCC of the lung harboring a discoid-receptor-2 and synchronous chronic myelogenous leukemia. _Lung Cancer Amst Neth_. 2013;82(1):171-172. doi:10.1016/j.lungcan.2013.07.004.

13. Rami-Porta R, Crowley JJ, Goldstraw P. The revised TNM staging system for lung cancer. _Ann Thorac Cardiovasc Surg Off J Assoc Thorac Cardiovasc Surg Asia_. 2009;15(1):4-9.

14. Douville C, Masica DL, Stenson PD, et al. Assessing the Pathogenicity of Insertion and Deletion Variants with the Variant Effect Scoring Tool (VEST-Indel). _Hum Mutat_. 2016;37(1):28-35. doi:10.1002/humu.22911.

15. Douville C, Carter H, Kim R, et al. CRAVAT: cancer-related analysis of variants toolkit. _Bioinforma Oxf Engl_. 2013;29(5):647-648. doi:10.1093/bioinformatics/btt017.
16. Desmet F-O, Hamroun D, Lalande M, Collod-Béroud G, Claustres M, Béroud C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* 2009;37(9):e67. doi:10.1093/nar/gkp215.

17. Barlesi F, Mazieres J, Merlio J-P, et al. Routine molecular profiling of patients with advanced non-small-cell lung cancer: results of a 1-year nationwide programme of the French Cooperative Thoracic Intergroup (IFCT). *Lancet Lond Engl.* 2016;387(10026):1415-1426. doi:10.1016/S0140-6736(16)00004-0.

18. Bai Y, Kim J-Y, Watters JM, et al. Adaptive responses to dasatinib-treated lung squamous cell cancer cells harboring DDR2 mutations. *Cancer Res.* 2014;74(24):7217-7228. doi:10.1158/0008-5472.CAN-14-0505.

19. Lee M-S, Jung EA, An SB, et al. Prevalence of Mutations in Discoidin Domain-Containing Receptor Tyrosine Kinase 2 (DDR2) in Squamous Cell Lung Cancers in Korean Patients. *Cancer Res Treat.* January 2017. doi:10.4143/crt.2016.347.

20. Miao L, Wang Y, Zhu S, et al. Identification of novel driver mutations of the discoidin domain receptor 2 (DDR2) gene in squamous cell lung cancer of Chinese patients. *BMC Cancer.* 2014;14:369. doi:10.1186/1471-2407-14-369.

21. Kenmotsu H, Serizawa M, Koh Y, et al. Prospective genetic profiling of squamous cell lung cancer and adenosquamous carcinoma in Japanese patients by multitarget assays. *BMC Cancer.* 2014;14:786. doi:10.1186/1471-2407-14-786.

22. Yashima H, Shimizu K, Araki T, et al. Assessment of DDR2, BRAF, EGFR and KRAS mutations as therapeutic targets in non-adenocarcinoma lung cancer patients. *Mol Clin Oncol.* 2014;2(5):714-718. doi:10.3892/mco.2014.302.

23. An S-J, Chen Z-H, Su J, et al. Identification of enriched driver gene alterations in subgroups of non-small cell lung cancer patients based on histology and smoking status. *PloS One.* 2012;7(6):e40109. doi:10.1371/journal.pone.0040109.

24. Govindan R, Ding L, Griffith M, et al. Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell.* 2012;150(6):1121-1134. doi:10.1016/j.cell.2012.08.024.

25. Matsuo K, Ito H, Yatabe Y, et al. Risk factors differ for non-small-cell lung cancers with and without EGFR mutation: assessment of smoking and sex by a case-control study in Japanese. *Cancer Sci.* 2007;98(1):96-101. doi:10.1111/j.1349-7006.2006.00347.x.

26. Xu C, Buczkowski KA, Zhang Y, et al. NSCLC Driven by DDR2 Mutation is Sensitive to Dasatinib and JQ1 Combination Therapy. *Mol Cancer Ther.* July 2015. doi:10.1158/1535-7163.MCT-15-0077.

27. Bollig-Fischer A, Chen W, Gadgeel SM, et al. Racial diversity of actionable mutations in non-small cell lung cancer. *J Thorac Oncol.* 2015;10(2):250-255. doi:10.1097/JTO.0000000000000420.

28. Exome Variant Server , NHLBI GO Exome Sequencing Project (ESP), Seattle, WA (URL: http://evs.gs.washington.edu/EVS/). Accessed September 4, 2015.

29. Prim N, Legrain M, Guerin E, et al. Germ-line exon 21 EGFR mutations, V843I and P848L, in nonsmall cell lung cancer patients. *Eur Respir Rev.* 2014;23(133):390-392. doi:10.1183/09059180.0009313.
30. Murray CW, Berdini V, Buck IM, et al. Fragment-Based Discovery of Potent and Selective DDR1/2 Inhibitors. *ACS Med Chem Lett*. 2015;6(7):798-803. doi:10.1021/acsmedchemlett.5b00143.
Table 1. Overview of the 11 DDR2 somatic mutations observed in a panel of 271 lung SCCs

| ID   | % tumoral cells | Exon | Ref | Alteration | Substitution Type | Chromosome | Starting codon | Nucleotide Change | AA Change | VAFs (%) | VEST pathogenecity p-value | Functional impact |
|------|-----------------|------|-----|------------|------------------|------------|----------------|-------------------|------------|----------|----------------------------|------------------|
| 7915 D | >50%            | 3    | C   | G          | Transversion     | 1          | 162688915      | c.62C>G         | p.A21G     | NS       | 43.21%                      | NEUTRAL          |
| 7135 D | >50%            | 15   | G   | C          | Transversion     | 1          | 162745442      | c.1857G>C       | p.R619S    | NS       | 60.54%                      | DELETERIOUS      |
| 6313 D | >50%            | 12   | G   | C          | Transversion     | 1          | 162740194      | c.1396G>C       | R          | NS       | 17.69%                      | NEUTRAL          |
| 5914 D | 25-50%          | 8    | G   | T          | Transversion     | 1          | 162729635      | c.721G>T        | p.R752     | NS       | 28.23%                      | DELETERIOUS      |
| 4789 D | 25-50%          | 16   | G   | A          | Transition       | 1          | 162746132      | c.2255G>A       | p.R752     | NS       | 50.15%                      | NEUTRAL          |
| 3605 D | >50%            | 12   | C   | T          | Transition       | 1          | 162740272      | c.1474C>T       | p.P492S    | NS       | 39.36%                      | DELETERIOUS      |
| 1439 D | >50%            | 16   | G   | A          | Transition       | 1          | 162742032      | c.1723G>A       | R          | NS       | 15.79%                      | NEUTRAL          |
| 1340 D | >50%            | 5    | C   | A          | Transversion     | 1          | 162724630      | c.402C>A        | K          | NS       | 12.32%                      | DELETERIOUS      |
| 1242 D | 25-50%          | 7    | G   | C          | Transversion     | 1          | 162725453      | c.566-1G>C      | NA         | SV       | 30.59%                      | DELETERIOUS      |
| 1223 D | 25-50%          | 5    | G   | T          | Transversion     | 1          | 162724481      | c.253G>T        | p.E85X     | SG       | 29.58%                      | DELETERIOUS      |
| 1011 D | >50%            | 16   | C   | A          | Transversion     | 1          | 162745961      | c.2084C>A       | p.A695     | D        | 25.13%                      | DELETERIOUS      |
| Gene                                | Analyzed samples | Mutated samples | %     |
|-------------------------------------|------------------|----------------|-------|
| DDR2 mutation                       | 271              | 11             | 4,06  |
| KRAS exon 2 mutation                | 265              | 5              | 1,89  |
| PI3KCA exon 10 and 21 mutation      | 145              | 11             | 7,59  |
| HER2 exon 20 insertion              | 234              | 1              | 0,43  |
| EGFR activating mutation            | 264              | 0              | 0     |
| BRAF exon 15 mutation               | 263              | 0              | 0     |
| No alteration found                 | 265              | -              | 97,78 |
Table 3. Baseline characteristics of the investigation cohort.

|                             | DDR2 mutated | DDR2 wild-type | p    |
|-----------------------------|--------------|----------------|------|
| N                           | 11           | 125            |      |
| Age                         | Median 65    | 68             | 0.1528 |
|                             | Range (51-78)| (39-92)        |      |
| Gender                      | Male 11 (100%)| 117 (94%)      | 0.999 |
|                             | Female 0 (0%) | 8 (6%)         |      |
| Smoking status              | Active smoker 5 (45%) | 54 (44%) | 0.2793 |
|                             | Former smoker 4 (36%) | 62 (49%) |      |
|                             | None 0 (0%) | 5 (4%)         |      |
|                             | Unknown 2 (18%) | 4 (3%) |      |
| Pack-Years                  | Median 53.62 | 42.04          | 0.1282 |
|                             | Range (30;80) | (2;100)        |      |
| Comorbidities               | Cardio-vascular 8 (72%) | 77 (61%) | 0.7481 |
|                             | COPD 5 (45%) | 25 (20%)       | 0.1277 |
|                             | Malignancies 4 (36%) | 25 (20%) | 0.2654 |
| ECOG Performance status     | 0 3 (26.5%) | 35 (28%)       | 0.701 |
|                             | 1 5 (45%) | 64 (51%)       |      |
|                             | 2 3 (26.5%) | 21 (17%)       |      |
|                             | ≥3 0 (0%) | 5 (4%)         |      |
| Stage TNM (2009)            | I-II 3 (27%) | 47 (37%)       | 0.1191 |
|                             | IIIA-resectable IIB 1 (9%) | 30 (24%) |      |
|                             | unresectable IIIB-IV 7 (63%) | 48 (38%) |      |
| Tumor localization          | Central 6 (54.5%) | 36 (28%) | 0.0939 |
|                             | Distal 5 (45.5%) | 89 (71%) |      |
| Line of chemotherapy        | Mean 1.4 | 1.4 | 0.8103 |
|                             | Range (1;2) | (1;4) |      |
| Follow-up                   | Median 13 | 13 | 0.918 |
|                             | Range (1;42) | (1;54) |      |
| Identifier | Value |
|------------|-------|
| some_value | 12345 |
| another_value | 67890 |

Diagram showing data distribution with various categories and colors.
