Resistance mechanisms in lung cancer patients with EGFR or ALK aberrations treated with kinase inhibitors
van der Wekken, Anthonie
Mutations in EMT related genes are an ALK independent resistance mechanism induced by crizotinib in ALK positive non-small cell lung cancer

Jiacong Wei
Anthonie J. van der Wekken
Ali Saber
M. Martijn Terpstra
Wim Timens
T. Jeroen N. Hiltermann
Harry J.M. Groen
Anke van den Berg
Klaas Kok

Submitted
Abstract

Background
Patients with ALK rearrangement positive lung cancer should receive crizotinib as the first line treatment. Although effective, most tumours develop resistance and show disease progression within one year, while a minority of patients do not respond at all. Mechanisms underlying resistance are only partly understood and studies on paired pre- and post-therapy biopsies are missing.

Methods
From our cohort of 29 ALK positive patients with advanced non-squamous NSCLC, we were able to retrieve sufficient tumour tissue from four responders before crizotinib treatment and upon resistance. For one non-responding patient, tumour tissue was obtained only before start of treatment. Somatic variants were detected by whole exome sequencing (WES) and pathway analysis was performed on resistant specific mutated genes in crizotinib-resistant samples.

Results
We identified 583 somatic mutations in crizotinib-resistant tumours, among which 137 mutations in 126 genes are resistant specific. These 126 genes were significantly enriched in 14 pathways, of which 9 genes related to the proteoglycans in cancer, HIF-1 signalling pathway, ECM-receptor interaction pathways, adherens junction, which are all related to epithelial-mesenchymal transition (EMT). Analysis of all EMT related pathways revealed 3 additional genes enriched in resistant tumours.

Conclusion
We observed a clear enrichment of mutations in genes associated with EMT related pathways, indicating that EMT may represent an important crizotinib resistance mechanism.
1. Introduction

Lung cancer is the leading cause of cancer-related deaths worldwide. Based on the traditional classification, there are two major types of lung cancer, small cell lung cancer and non-small cell lung cancer (NSCLC). NSCLC is divided into squamous, adenocarcinoma and large cell carcinoma, of which adenocarcinoma is the most common subtype. During the last decade, the clinical management and treatment of lung cancer patients has become more dependent on the molecular classification defined by ‘driver’ mutations that occur in genes like ALK, EGFR and ROS1. About 5% of all adenocarcinomas have a break in the ALK gene, resulting in a fusion product containing the kinase domain of the ALK protein fused to EML4 or other fusion proteins. Tumour cells with an ALK-break are highly sensitive to tyrosine kinase inhibitors (TKIs) targeting ALK, such as crizotinib, or second-generation ALK inhibitors, ceritinib and alectinib. However, most responders will inevitably acquire resistance against the TKI treatment within one year.

The mechanisms of resistance can be divided into ALK-dependent and ALK-independent. Approximately one third of the patients seem to have ALK-dependent resistant mechanisms, which include secondary ALK mutations and amplifications. Several ALK-independent mechanisms have been proposed based on studies in post-crizotinib tumour samples and studies in cell line models. These mechanisms include mutations in and amplifications of EGFR, KIT, MET and IGFR-1R, as well as activation of the MAP kinase pathway, PI3K/AKT pathways, Ras/MAPK pathways and JAK/STAT pathways. However, the level of proof is highly variable and the complete picture of all mechanisms involved in crizotinib resistance is not clear yet.

In this study, we sought to further characterize crizotinib treatment resistant-related somatic mutations using whole exome sequencing (WES) of paired tumour biopsies of advanced adenocarcinoma patients before and after crizotinib treatment.

2. Methods

2.1. Patient selection

In our previous study, we described 29 non-squamous NSCLC patients who were ALK break positive and were treated with crizotinib. For four out of 29 patients, we had sufficient pre- and post-treatment tumour samples (formalin fixed paraffin embedded (FFPE) or frozen) for WES in this study. For one non-responding patient we had sufficient FFPE material for
WES. The patients are numbered as ALK4, ALK6, ALK8, ALK14, and ALK16 in concordance with our previous study. ALK4 was shown to have an ALK G1269A mutation based on our previously reported RNA-seq data\(^\text{15}\). White blood cells were isolated from peripheral venous blood and used as normal control. Tumour samples before crizotinib treatment from patient ALK15 and upon resistance to crizotinib treatment from patients ALK4 and ALK6 were fresh frozen, whereas the other tumour specimens were FFPE. Macro dissection was applied to the tumour samples to achieve a tumour cell content of more than 60%.

### 2.2. DNA isolation

DNA from blood and frozen samples was isolated using a standard salt-chloroform DNA isolation protocol. For FFPE samples, DNA was isolated using the ReliaPrep™ FFPE gDNA Miniprep System kit (Promega, Madison, USA) following protocol of the manufacturer. The concentrations of the DNA samples were measured by NanoDrop (Thermo Fisher Scientific Inc., Waltham, USA), and the quality was evaluated on a 1% agarose gel.

### 2.3. Whole exome sequencing

WES was carried out by BGI (BGI Tech Solutions Co. Ltd, Hong Kong) on 0.6-2μg genomic DNA of normal and tumour derived DNA samples. Target enrichment was done using the Agilent SureSelect All Exon V5 kit (Agilent technologies, Santa Clara, USA). Paired-end sequencing with a read length of 2x100nt was performed on Illumina HiSeq2000. As part of the validation procedure we performed a second WES on the crizotinib resistant samples of patients ALK4 and ALK6 using the protocol and data analysis pipeline as published previously\(^\text{15}\). Also, RNA-sequencing data of resistant samples of ALK4 and ALK6 were re-analysed to confirm presence of mutations detected in the BGI data\(^\text{15}\).

### 2.4. Bioinformatics approach

The variant calling pipeline is an adaption of the GATK workflow and molgenis compute as workflow management software\(^\text{16-18}\). Alignment of reads was done using BWA and the Genome Analysis Toolkit (GATK), using the human genome reference build GRCH37 with decoys from the GATK bundle\(^\text{19}\). Picard Tools were used for format conversion and marking duplicates. As variant caller, this pipeline uses the HaplotypeCaller for all the samples of the same patient/cohort. Variants were annotated using SnpEff / SnpSift with the ensembl release 75 gene annotations and the dbNSFP2.0 database\(^\text{20-22}\). GATK was used to
identify variant annotated in dbsnp 138, Cosmic v72, 1000 genomes phase 3 and Exac 0.3 databases. The data were filtered for quality metrics similar to GATK recommendations and custom filters for population frequency and variant effect. Combined Annotation Dependent Depletion (CADD) scores were used to predict pathogenicity of the identified variants. Variants with a CADD score equal or more than twenty are defined as deleterious; those between 10 and 20 are possibly deleterious and those below 10 are non-deleterious.

To identify somatic mutations we first excluded variants for which the total number of reads in the normal sample was less than ten. Next, we excluded all variants for which one or more mutant reads were present in the normal sample. Then, variants with total reads less than ten in either the primary or resistant samples were excluded. The remaining variants with two or more mutant reads in either the pre-treatment tumour samples or post-treatment tumour samples were considered as somatic. Variants with mutant read frequencies (MRFs) ≥20% in resistant samples, and MRFs in the resistant sample at least two times more than those in the paired primary sample were marked as “treatment-related” variants. Pathway analysis was performed with Partek Genomics Suite 6.6 (Partek Inc., St Louis, MO) software using all genes with crizotinib resistance-related variants.

2.5. WES-based copy number variant (CNV) analysis

Pseudo probe data were generated with SAMtools, VarScan2 and DNAcopy. Briefly, for each sample the pseudo probe derived GC-normalized log2 copy number ratios were generated using the corresponding normal sample in case frozen tumour tissue was used for isolation of high quality DNA samples. For tumour samples with low quality DNA (all FFPE sample) we used WES data of a merged pool of normal samples as the reference. All alignments with a mapping quality greater than 40 in combination with a minimal segment size of 2kb and a maximal segment size of 5kb with a mean base-wise coverage of at least one were used to calculate the ratios. CNV plots of the resistant tumour sample were compared to the CNV plot of the paired primary tumour sample by a combination of the calculated ratios in identical bins based on changes in copy number level and by visual inspection.
3. Results

3.1. Patients

Four patients (ALK4, ALK6, ALK14, and ALK16) developed resistance to crizotinib after an initial response of approximately one year (Table 5.1 and Figure 5.1). One patient, ALK8, had no tumour response on crizotinib, and died after one month. ALK6 and ALK16 deceased 15 months and 3 years after end of crizotinib treatment, respectively. ALK4 and ALK14 patients are still alive at 12 and 9 years after initial diagnosis, respectively.

Table 5.1. Characteristics of five ALK rearranged advanced NSCLC patients

<table>
<thead>
<tr>
<th>Patient</th>
<th>Gender</th>
<th>Age at diagnosis</th>
<th>PFS* (m)</th>
<th>Smoking status (PY**)</th>
<th>Primary tumour biopsy</th>
<th>Resistant tumour biopsy</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALK4</td>
<td>F</td>
<td>34</td>
<td>15.9</td>
<td>Non-smoker</td>
<td>Ovary</td>
<td>Liver</td>
</tr>
<tr>
<td>ALK6</td>
<td>M</td>
<td>55</td>
<td>9.5</td>
<td>Past-smoker (15)</td>
<td>Cervical lymph node</td>
<td>Glenoid</td>
</tr>
<tr>
<td>ALK8</td>
<td>M</td>
<td>76</td>
<td>1.6</td>
<td>Past-smoker (NA)</td>
<td>Lymph node mediastinal (4R)</td>
<td>ND</td>
</tr>
<tr>
<td>ALK14</td>
<td>F</td>
<td>62</td>
<td>8.4</td>
<td>Current smoker (20)</td>
<td>Brain occipital metastasis</td>
<td>Mediastinal lymph node (7)</td>
</tr>
<tr>
<td>ALK16</td>
<td>F</td>
<td>48</td>
<td>4.1</td>
<td>Past-smoker (18)</td>
<td>Bronchial biopsy</td>
<td>Lung</td>
</tr>
</tbody>
</table>

*PFS: progression free (time from start of treatment to the examination by CT scan where we have seen progressive disease or death). **PY: pack years. ND: not done.

3.2. Whole exome sequencing and its validation

WES generated an average of 60 x 10^6 unique reads that passed the Illumina quality filtering steps per sample (Supplementary Table 5.1). An average of 98% of the unique reads could be aligned to the human reference genome. The mean coverage per sample was 66x with 86% of the target region covered at least 20x.

We identified 583 variants in 519 genes across the five patients. Among these 583 variants, 20% (116 SNVs in 143 genes) had CADD scores equal or higher than 20; 45% (265 SNVs in 253 genes) had CADD scores between 10 and 20; 26% (151 SNVs in 101 genes) had CADD scores less than 10. For 9% of the variants (51 variants in 47 genes) no CADD scores were available; 47/51 variants were small insertions or deletions (INDELs).
We analysed validity of the WES data by two independent approaches. We first re-analysed the RNA-seq data from our published study, which included two crizotinib-resistant tumour samples of ALK4 and ALK6 patients. A coverage of ≥10x used as a criterion to allow a reliable confirmation was observed for 95 out of the 169 variants in ALK4 and 34 out of the 61 variants in ALK6; 90 out of the 95 variants (95%) in ALK4 and 29 out of 34 (85%) in ALK6 mutant reads were observed consistent with those identified in the WES data, indicating that the vast majority of the mutated genes were expressed in these tumour samples. The second approach was based on an independent WES, again focusing on variants with a coverage of ≥10x. Out of the 67 variants in ALK4 and the 42 variants in ALK6, we confirmed 61 (91%) variants in ALK4 and 40 (95%) variants in ALK6. In total, we independently validated 114 out of 123 (93%) variants in ALK4 and 45 out of 48 (94%) variants in ALK6.

### 3.3. Crizotinib treatment-related variants

A comparison of the primary to the resistant tumour in all four paired patients revealed in total 176 different “treatment-related” variants in 156 genes (Figure 5.2). Among these variants 16% (21 SNVs in 15 genes) had CADD scores above 20, 43% (75 SNVs in 71 genes) had CADD scores between 10 and 20, 30% (53 SNVs in 49 genes) had CADD scores less than 10 and 11% (20 INDELs in 19 genes) had no CADD scores. The distribution of the variants
over the different CADD score groups was similar to the distribution of all somatic variants. One gene, BZRAP1, was mutated in more than one patient only. The 156 mutated genes were present in 167 pathways. A significant enrichment was observed for 20 pathways. To focus on ALK-independent resistance mechanisms, we next excluded gate keeper genes mutated in patient ALK4, which was known to carry an ALK mutation that explains the crizotinib resistance. The three paired patients had a total of 90 “treatment-related” variants in 74 genes. All genes were mutated only in one of the three patients. These 74 genes were present in 105 pathways of whom 15 pathways were significantly enriched (Supplementary Table 5.2). Five of these were related to “hormonal pathways”, three to “specific cancer subtypes”, one to “hepatitis B”, one to “GABAergic synapse” and one to “arrhythmogenic right ventricular cardiomyopathy”.

Figure 5.2. Comparison of mutant read frequencies (MRF) in primary (x-axis) and resistant samples (y-axis). Each spot is a single variant. Variants with mutant read frequencies (MRFs) ≥20% in resistant samples and MRFs in the resistant sample at least two times more than those in the paired primary sample (number in figure) were marked as “treatment-related” variants.

A. ALK4  
B. ALK6  
C. ALK15  
D. ALK17
Four of the pathways including 9 genes were linked to EMT, i.e. proteoglycans in cancer, HIF-1 signalling pathway, FoxO signalling pathway and ECM-receptor interaction (Table 5.2). Four of the genes, i.e. ARNT, PTPN11, SMAD4, VEGFA, variants had CADD scores of more than 20, whereas no CADD scores was available for one gene, i.e. LAMA2, having an out of frame INDEL (LAMA2 has two nucleotides deleted).

In ALK8, the patient without a response to crizotinib, somatic variants in three EMT related genes, i.e. ITGAM, CACNA1E, and RUVBL1, were observed. These three genes were involved in the cell adhesion molecules (CAMs), MAPK signalling, Wnt signalling, and regulation of actin cytoskeleton. Pathway analysis per patient on genes with treatment-related variants revealed only one pathway that was shared by all five patients, i.e. metabolism pathway.

<table>
<thead>
<tr>
<th>Pathway name</th>
<th>Genes mutated</th>
<th>Enrichment score</th>
<th>Enrichment p-value</th>
<th>Genes mutated / genes not mutated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proteoglycans in cancer</td>
<td>ANK2, FASLG, HSPG2, PTPN11, STAT3, VEGFA</td>
<td>8.0</td>
<td>0.00034</td>
<td>6/1049</td>
</tr>
<tr>
<td>HIF-1 signalling</td>
<td>ARNT, STAT3, VEGFA</td>
<td>4.3</td>
<td>0.01391</td>
<td>3/548</td>
</tr>
<tr>
<td>FoxO signalling</td>
<td>FASLG, SMAD4, STAT3</td>
<td>3.7</td>
<td>0.02598</td>
<td>3/696</td>
</tr>
<tr>
<td>ECM-receptor interaction</td>
<td>HSPG2, LAMA2</td>
<td>3.3</td>
<td>0.03764</td>
<td>2/331</td>
</tr>
</tbody>
</table>

3.4. Treatment-related CNVs

We compared WES-based CNV plots of the primary tumours to the paired resistant samples. Although we did identify several differences in copy number variations between primary and resistant samples (Figure 5.3), we did not see amplification of the ALK, MET or KIT gene loci known to be associated with resistance in any of the patients. Moreover, no recurrent CNVs were observed that were shared between the 4 patients in other parts of the genome.
Figure 5.3. CNV plots across four patients with ALK rearranged advanced NSCLC with matched primary and resistant tumour samples. The chromosomes with conspicuous changes between primary and resistant samples in each patient are marked in red box.
4. Discussion

In this study, we identified 176 variants in 156 genes that are specific for or more enriched in crizotinib resistant samples as compared with the matched tumour sample before crizotinib treatment. Pathways analysis based on genes with treatment related mutations revealed a significant enrichment of 9 genes in four pathways, i.e. proteoglycans in cancer (ANK2, FASLG, HSPG2, PTPN11, STAT3, VEGFA), HIF-1 signalling pathway (ARNT, STAT3, VEGFA), FoxO signalling pathway (FASLG, SMAD4, STAT3), and ECM-receptor (extracellular matrix) interaction pathway (HSPG2, LAMA2). These four pathways are all involved in EMT and gene mutations were found in two of the three patients with ALK-independent resistance mechanisms.

To our knowledge, we are the first to explore ALK-independent mechanisms of crizotinib resistance in advanced NSCLC patients by comparing variants observed in the crizotinib resistant tumour to the variants present in the tumour sample before crizotinib treatment. Many crizotinib resistance mechanisms have been proposed based on patients samples and cell line studies. However, these studies were performed on resistant samples without comparison to the pre-treatment samples. For example, activation of ALK-independent bypass mechanisms, such as activation of EGFR, KRAS, SRC and MAPK signaling, were shown in resistant samples from NSCLC patients. Although these studies generated valuable data, it does not allow to pinpoint the true treatment induced alterations. Although the total number of matched samples was still quite limited, we present the first comparative analysis on pre- and post-crizotinib treatment tumour specimens in ALK+ NSCLC patients. A main problem in lung cancer patients is the scarceness of re-biopsies, which generally are relatively small and with low tumour cell percentages. These biopsies are frequently used in total for the initial diagnostic tests to guide further therapy. In future studies, we may be able to use liquid-biopsy based material, which will be easier to obtain and will allow with rapid improvements in NGS techniques genome wide analysis.

The limited DNA quality of the FFPE tumour samples resulted in a suboptimal coverage and this precluded detection of mutations in sub clones of the tumour. Although such coverage might have resulted in an underestimation of somatic variants, it should not have affected the identification of treatment-related mutations because of the use of resistant tumour samples. Nevertheless, we also experienced limited biopsy sizes and therefore extensive validations of the observed variants was limited.

Based on previous studies, the resistant mechanisms against ALK inhibitors can be classified into ALK-dependent and ALK-independent mechanisms. Like many other studies, we
found the known \( \text{ALK} \) G1269A mutation present in the resistant tumour while absent in the primary tumour in patient \( \text{ALK}4 \), which is in accordance with the RNA-seq result in our previous and other studies\(^7,15,34,35\).

For the remaining three patients, no \( \text{ALK} \) gate keeper mutations were found, indicating possible activation of \( \text{ALK} \)-independent bypass mechanisms. Since there was no overlap between the genes mutated in these three patients, it might be that \( \text{ALK} \)-independent mechanisms are quite diverse. This would be consistent with the broad variation in resistance mechanisms proposed in the literature\(^9,34\). To find a possible common mode of action, we proceeded with pathway analysis on the potentially treatment-related genes in these three patients. Among the significantly enriched pathways, we found four EMT related pathways and the subsequent analysis of all other EMT related pathways revealed a total of 12 mutated genes that are involved in EMT-related pathways. Several studies either on cell lines studies have proven the potential role of EMT as a mediator of resistance against \( \text{ALK} \) inhibitors. Silencing of vimentin restored the responsiveness of \( \text{ALK} \) break positive, crizotinib-resistant cells to \( \text{ALK} \) inhibitors, indicating a causal relation between EMT and crizotinib resistance\(^36\). Five out of eleven \( \text{ALK}^+ \) NSCLC patients who were treated with the second-generation \( \text{ALK} \) inhibitor ceritinib, showed EMT based on immunostaining for E-cadherin and vimentin\(^7\). In our study, seven of the 14 genes involved in EMT had variants with a potential impact on the protein based on a CADD score of more than 20 or presence of out of frame INDELs.

\( \text{SMAD}4 \) has been proven to be one of the most crucial genes in regulating the TGF-beta signalling pathway. The TGF-beta pathway is one of the most important pathways leading to EMT by e.g. TGF-beta-SMAD signalling, induction of microRNA expression, decreasing expression of epithelial splicing regulatory proteins (ESRPs), PI3K–AKT–mTORC1 signalling, and cytoskeletal changes\(^37\). HIF-1 has been shown to activate the expression of SNAI1 by binding to two hypoxia response elements (HREs) in the promoter. SNAI1 represses expression of the epithelial marker E-cadherin and enhances expression of the mesenchymal markers \( \beta \)-catenin and vimentin\(^38,39\). By knockdown and/or transfection experiments on cell line studies related to several cancer types including prostate cancer, gastric cancer and breast cancer, three out of the four FOXO family members (FOXO1, FOXO3a, FOXO4, except for FOXO6) have been proven to regulate EMT by repressing E-cadherin expression and promoting SNAI1 expression\(^40-42\). Some proteoglycans like syndecans are required to maintain the epithelial characteristic of basolateral surfaces and within adhesive junctions. Deficiencies in cell surface syndecan-1 lead to markedly reduced E-cadherin expression in
normal gland epithelia\textsuperscript{43,44}. ECM has been shown to promote EMT by weakening cell-cell adhesions\textsuperscript{37,45}.

The only recurrently mutated gene in more than one crizotinib resistant patients is BZRAP1, which is a synaptic transmission regulator. This gene has never been described in cancer related resistance.

In case of crizotinib resistance, there are second and third generation ALK inhibitors such as ceritinib, alectinib, brigatinib and lorlatinib that can be used as next treatment option\textsuperscript{46}. These drugs will be effective especially for ALK dependent resistant mechanism, based on their improved binding to the ATP binding site. It is still unclear whether these drugs also show tumour response in ALK independent resistant mechanism. This should be a further focus of future clinical trials in patients with ALK-independent crizotinib-resistance mechanisms.

In conclusion, the most significant finding in this study is the identification of mutations in genes involved in EMT-related pathways. From a clinical perspective, the mutational status of patients may provide therapeutic guidance for clinical management and the future use of EMT blocking agents in NSCLC patients should be studied further. This could therefore be a new target for treatment.
5. References


### 6. Supplementary files

Resistance mechanisms associated with crizotinib in ALK+ NSCLC patients

#### Supplementary Table 5.1. EMT related significantly enriched pathways based on crizotinib induced mutated genes in patients

<table>
<thead>
<tr>
<th>Sample</th>
<th>ALK4</th>
<th>ALK6</th>
<th>ALK8</th>
<th>ALK14</th>
<th>ALK16</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>T</td>
<td>R</td>
<td>N</td>
<td>T</td>
</tr>
<tr>
<td>Clean reads x10⁶</td>
<td>61</td>
<td>74</td>
<td>64</td>
<td>60</td>
<td>66</td>
</tr>
<tr>
<td>Unique reads x10⁶</td>
<td>58</td>
<td>62</td>
<td>60</td>
<td>57</td>
<td>59</td>
</tr>
<tr>
<td>Percentage of unique reads</td>
<td>95%</td>
<td>84%</td>
<td>94%</td>
<td>95%</td>
<td>89%</td>
</tr>
<tr>
<td>Unique aligned reads (x10⁶)</td>
<td>57</td>
<td>60</td>
<td>59</td>
<td>56</td>
<td>58</td>
</tr>
<tr>
<td>Percentage of unique aligned reads</td>
<td>98%</td>
<td>97%</td>
<td>98%</td>
<td>98%</td>
<td>98%</td>
</tr>
<tr>
<td>Average target coverage</td>
<td>68</td>
<td>70</td>
<td>66</td>
<td>65</td>
<td>71</td>
</tr>
<tr>
<td>20X coverage</td>
<td>91%</td>
<td>90%</td>
<td>91%</td>
<td>91%</td>
<td>90%</td>
</tr>
<tr>
<td>Number of variants</td>
<td>0</td>
<td>82</td>
<td>169</td>
<td>0</td>
<td>44</td>
</tr>
<tr>
<td>CADD scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;10</td>
<td>0</td>
<td>20</td>
<td>28</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>[10,20)</td>
<td>0</td>
<td>34</td>
<td>89</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>≥20</td>
<td>0</td>
<td>18</td>
<td>46</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>NA</td>
<td>0</td>
<td>10</td>
<td>6</td>
<td>0</td>
<td>6</td>
</tr>
</tbody>
</table>

N: normal sample; T: tumour sample before crizotinib treatment; R: tumour sample when resistance was observed either during or after crizotinib treatment. CADD: Combined Annotation Dependent Depletion scores used to predict pathogenicity of the identified variants.
## Supplementary Table 5.2. Significantly enriched pathways harboring treatment-related variants in ALK6, ALK14, and ALK16

<table>
<thead>
<tr>
<th>Pathway name</th>
<th>Genes mutated</th>
<th>Enrichment score</th>
<th>Enrichment p-value</th>
<th>Genes mutated / genes not mutated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proteoglycans in cancer</td>
<td>ANK2, FASLG, HSPG2, PTPN11, STAT3, VEGFA</td>
<td>8</td>
<td>0.00034</td>
<td>6/1049</td>
</tr>
<tr>
<td>Pathways in cancer</td>
<td>ARNT, CTNNA3, FASLG, LAMA2, SMAD4, STAT3, VEGFA</td>
<td>6.5</td>
<td>0.00154</td>
<td>7/1932</td>
</tr>
<tr>
<td>AGE-RAGE signaling pathway in diabetic complications</td>
<td>PLCD1, SMAD4, STAT3, VEGFA</td>
<td>6.4</td>
<td>0.00169</td>
<td>4/560</td>
</tr>
<tr>
<td>Renal cell carcinoma</td>
<td>ARNT, PTPN11, VEGFA</td>
<td>5.4</td>
<td>0.00455</td>
<td>3/363</td>
</tr>
<tr>
<td>Hepatitis B</td>
<td>FASLG, HSPG2, SMAD4, STAT3</td>
<td>5.3</td>
<td>0.00489</td>
<td>4/754</td>
</tr>
<tr>
<td>Pancreatic cancer</td>
<td>SMAD4, STAT3, VEGFA</td>
<td>5</td>
<td>0.00662</td>
<td>3/416</td>
</tr>
<tr>
<td>HIF-1 signaling pathway</td>
<td>ARNT, STAT3, VEGFA</td>
<td>4.3</td>
<td>0.01391</td>
<td>3/548</td>
</tr>
<tr>
<td>Type I diabetes mellitus</td>
<td>FASLG, GAD2</td>
<td>4.2</td>
<td>0.01474</td>
<td>2/199</td>
</tr>
<tr>
<td>FoxO signaling pathway</td>
<td>FASLG, SMAD4, STAT3</td>
<td>3.7</td>
<td>0.02598</td>
<td>3/696</td>
</tr>
<tr>
<td>GABAergic synapse</td>
<td>GAD2, SLC6A1</td>
<td>3.4</td>
<td>0.03323</td>
<td>2/309</td>
</tr>
<tr>
<td>Arrhythmogenic right ventricular cardiomyopathy (ARVC)</td>
<td>CTNNA3, LAMA2</td>
<td>3.3</td>
<td>0.03601</td>
<td>2/323</td>
</tr>
<tr>
<td>Taurine and hypotaurine metabolism</td>
<td>GAD2</td>
<td>3.3</td>
<td>0.0363</td>
<td>1/1939</td>
</tr>
<tr>
<td>ECM-receptor interaction</td>
<td>HSPG2, LAMA2</td>
<td>3.3</td>
<td>0.03764</td>
<td>2/331</td>
</tr>
<tr>
<td>Aldosterone synthesis and secretion</td>
<td>DAGLA, SCARB1</td>
<td>3.1</td>
<td>0.04574</td>
<td>2/369</td>
</tr>
<tr>
<td>Adipocytokine signaling pathway</td>
<td>PTPN11, STAT3</td>
<td>3.1</td>
<td>0.04708</td>
<td>2/375</td>
</tr>
</tbody>
</table>
Resistance mechanisms associated with crizotinib in ALK+ NSCLC patients