

Article

# Mutations in EMT-Related Genes in *ALK* Positive Crizotinib Resistant Non-Small Cell Lung Cancers

Jiacong Wei <sup>1,2</sup> , Anthonie J. van der Wekken <sup>3</sup>, Ali Saber <sup>4</sup> , Miente M. Terpstra <sup>5</sup>, Ed Schuurin <sup>4</sup>, Wim Timens <sup>4</sup>, T. Jeroen N. Hiltermann <sup>3</sup>, Harry J. M. Groen <sup>3</sup>, Anke van den Berg <sup>4</sup> and Klaas Kok <sup>5,\*</sup>

<sup>1</sup> Department of Genetics, University of Groningen, University Medical Centre Groningen, 9700 RB Groningen, The Netherlands; weijiacong@126.com

<sup>2</sup> Department of Pathology, Shantou University, Shantou 515041, China

<sup>3</sup> Department of Pulmonary Diseases, University of Groningen, University Medical Centre Groningen, 9700 RB Groningen, The Netherlands; a.j.van.der.wekken@umcg.nl (A.J.v.d.W.); t.j.n.hiltermann@umcg.nl (T.J.N.H.); h.j.m.groen@umcg.nl (H.J.M.G.)

<sup>4</sup> Department of Pathology and Medical Biology, University of Groningen, University Medical Centre Groningen, 9700 RB Groningen, The Netherlands; a.saber814@gmail.com (A.S.); e.schuuring@umcg.nl (E.S.); w.timens@umcg.nl (W.T.); a.van.den.berg01@umcg.nl (A.v.d.B.)

<sup>5</sup> Department of Genetics, University of Groningen, University Medical Centre Groningen, 9700 RB Groningen, The Netherlands; m.m.terpstra.cluster@gmail.com

\* Correspondence: k.kok@umcg.nl; Tel.: +31-06-3893-0111

Received: 4 December 2017; Accepted: 29 December 2017; Published: 4 January 2018

**Abstract:** Crizotinib is an effective drug for patients with anaplastic lymphoma kinase (ALK)-positive non-small-cell lung cancer (NSCLC), but upon treatment, the tumors inevitably become crizotinib resistant in time. The resistance mechanisms are only partly understood. In this study, we aim to identify gene mutations associated with resistance in ALK-positive advanced non-squamous NSCLC treated with crizotinib. Four ALK positive patients with progressive disease following crizotinib treatment were identified with paired pre- and post-crizotinib tumor tissue from our previously published cohort. Somatic variants in these samples were detected by whole exome sequencing. In one of the four patients, an ALK-resistance associated mutation was identified. In the other three patients, no ALK-resistance associated mutations were present. In these patients we identified 89 relevant somatic mutations in 74 genes that were specific to the resistant tumors. These genes were enriched in 15 pathways. Four pathways, were related to epithelial-mesenchymal transition (EMT): proteoglycans in cancer, HIF-1 signaling, FoxO signaling pathway, and ECM-receptor interaction. Analysis of other EMT-related pathways revealed three additional genes with mutations specific to the crizotinib-resistant tumor samples. The enrichment of mutations in genes associated with EMT-related pathways indicates that loss of epithelial differentiation may represent a relevant resistance mechanism for crizotinib.

**Keywords:** whole exome sequencing; adenocarcinoma; *ALK*; crizotinib resistance

## 1. Introduction

Lung cancer is the leading cause of cancer-related deaths worldwide [1]. In the traditional clinical classification, there are two major types of lung cancer, small cell lung cancer, and non-small-cell lung carcinoma (NSCLC) [2]. NSCLC is further subdivided into the major categories of squamous cell carcinoma, adenocarcinoma, and large cell carcinoma, with adenocarcinoma being the most common subtype [2]. Over the last decade, clinical management and treatment of lung cancer patients has become more dependent on molecular classification using “driver” mutations that occur in genes,

such as anaplastic lymphoma kinase (*ALK*), epidermal growth factor receptor (*EGFR*), and c-Ros oncogene 1 (*ROS1*) [3–5]. Around 5% of all adenocarcinomas have a chromosomal inversion or translocation that produces a fusion product consisting of the kinase domain of *ALK* combined with *EML4* or another fusion partner [5,6]. Tumor cells with an *ALK* fusion are highly sensitive to tyrosine kinase inhibitors (TKIs) that target *ALK*, which include crizotinib and the second-generation *ALK* inhibitors, ceritinib, brigatinib, and alectinib [7–10]. However, most patients will inevitably acquire resistance to *ALK*-TKI treatment, usually within one year [11,12].

The mechanisms of resistance can be *ALK*-dependent or *ALK*-independent [13]. Secondary *ALK* mutations are observed in 25–33% of patients progressing to crizotinib treatment, and the number of mutations increases to approximately 50% after using a second-generation TKI [14]. Several *ALK*-independent resistance mechanisms have been proposed based on studies in post-crizotinib tumor samples and cell line models [15–20]. These mechanisms include alterations of *EGFR*, proto-oncogene receptor tyrosine kinase (*KIT*), and Insulin like growth factor 1 receptor (*IGF1R*), as well as activation of the MAP kinase, PI3K/AKT, RAS/MAPK, and JAK/STAT pathways. Studies based on resistant and sensitive cell lines and those based solely on post-resistant tumor tissues have suggested that the epithelial-mesenchymal transition (EMT) potentially mediates resistance against *ALK* inhibitors [21–23]. However, the level of evidence is variable and our understanding of the mechanisms that are involved in crizotinib-resistance remains incomplete.

In this study, we aimed to identify somatic mutations related to crizotinib resistance using whole exome sequencing (WES) of paired tumor biopsies from advanced adenocarcinoma patients, taken both before crizotinib treatment and upon disease progression to crizotinib.

## 2. Results

### 2.1. Patients

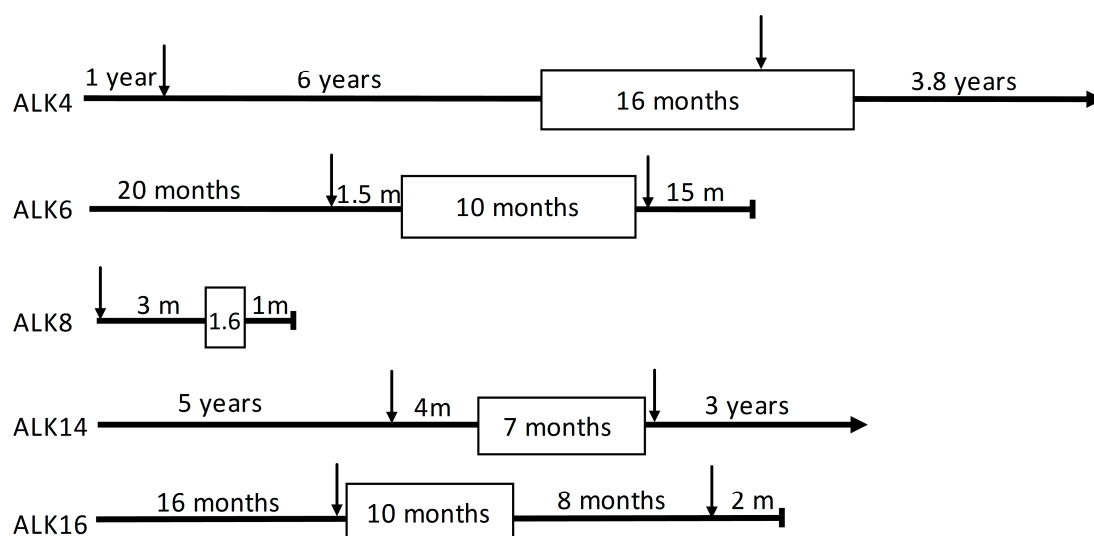
Four patients (*ALK4*, *ALK6*, *ALK14*, and *ALK16*) developed crizotinib resistance after initially responding to treatment for approximately one year (Table 1 and Figure 1). One patient, *ALK8*, did not respond to crizotinib treatment at all and died 2.5 months after initiating crizotinib. *ALK6* and *ALK16* died 15 months and 10 months, respectively, after crizotinib treatment. At the time of this study, *ALK4* and *ALK14* were still alive 3.8 and 3 years, respectively, after developing crizotinib resistance. Upon developing resistance, *ALK4* was treated with ceritinib and *ALK14* with alectinib. Patients *ALK6* and *ALK16* had not been treated with 2nd or 3rd generation *ALK* inhibitors, because those drugs were not available at that time.

**Table 1.** Characteristics of the five anaplastic lymphoma kinase (*ALK*) positive advanced non-small-cell lung cancer (NSCLC) patients.

Patient	Gender	Age at Diagnosis (Years)	PFS <sup>1</sup> (m)	Smoking Status (py <sup>2</sup> )	Tumor Location	
					Primary <sup>4</sup>	Resistant <sup>5</sup>
<i>ALK4</i>	female	34	15.9	non-smoker	ovary	liver
<i>ALK6</i>	male	55	9.5	past-smoker (15)	cervical lymph node	glenoid
<i>ALK8</i>	male	76	1.6	past-smoker (NA <sup>3</sup> )	lymph node mediastinal 4 right	not done
<i>ALK14</i>	female	62	8.4	current smoker (20)	brain occipital metastasis	mediastinal lymph node 7
<i>ALK16</i>	female	48	5.1	past-smoker (18)	lung	lung

<sup>1</sup> PFS: progression-free survival (time from start of treatment to disease progression estimated by CT); <sup>2</sup> py: pack years;

<sup>3</sup> NA: not available; <sup>4</sup> Primary: primary tumor biopsy; <sup>5</sup> Resistant: resistant tumor biopsy.



**Figure 1.** Schematic representation of the total period of treatment and follow-up per patient, indicated by a horizontal black line. ALK4 and ALK14 were alive at the end of study; ALK6, ALK8 and ALK16 died before the end of study). Blocks indicate the duration of crizotinib treatment. Arrows indicate when pre- and post-crizotinib tissue samples were collected. Notes: ALK4 was treated with several lines of chemotherapy and radiotherapy. ALK4's second biopsy was taken from a local relapse, and crizotinib was continued. Ceritinib was started at further disease progression. Patient ALK6, ALK14, ALK16 received first-line chemo (radio) therapy before the first biopsy. For patient ALK8, crizotinib treatment was stopped after 1.6 months upon disease progression. ALK14 had ongoing response to alectinib subsequent to the crizotinib treatment. ALK16 was treated with ceritinib for eight months following crizotinib treatment.

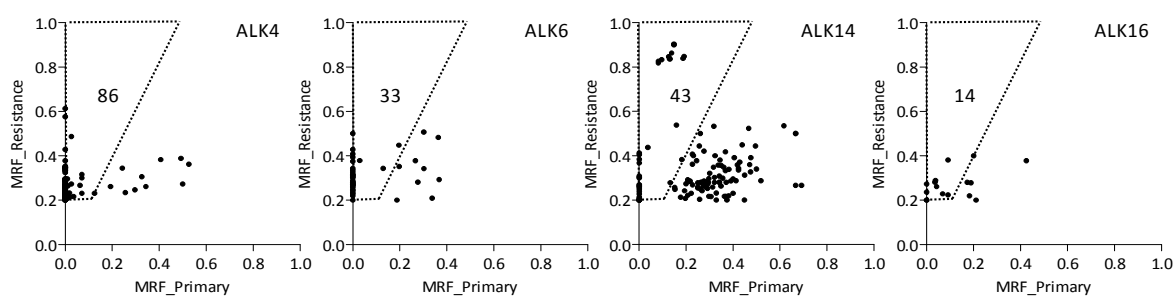
## 2.2. Whole Exome Sequencing

WES generated an average of  $60 \times 10^6$  unique reads per sample that passed the Illumina quality filtering steps (Supplementary Table S1). On average, 98% of the unique reads could be aligned to the human reference genome. The mean coverage per sample was  $66\times$  with 86% of the target region being covered at least  $20\times$ . We identified 582 variants in 519 genes across the four patients who had developed resistance. Of these variants, 20% (116 single-nucleotide variants (SNVs) in 143 genes) had Combined Annotation Dependent Depletion (CADD) scores  $\geq 20$ , 45% (265 SNVs in 253 genes) had CADD scores between 10 and 20, and 26% (151 SNVs in 101 genes) had CADD scores  $< 10$ . For 9% of the variants (50 variants in 47 genes), no CADD score was available. Most of the variants without CADD scores ( $n = 46$ ) were small insertions or deletions (INDELS), of which 25 were out of frame INDELS and should be considered as being deleterious.

We analyzed the validity of the WES data using two independent approaches. First, we reanalyzed RNA-seq data from our published study that included the crizotinib-resistant tumor samples of ALK4 and ALK6 patients [24]. A  $\geq 10\times$  coverage at the variant position was observed for 95 out of the 169 WES variants in ALK4 and 34 of the 61 WES variants in ALK6. Ninety out of the 95 variants (95%) in ALK4 and 29 of the 34 variants (85%) in ALK6 were consistent with those that were identified in the WES data, indicating that the majority of the mutant alleles were also expressed in these tumor samples. Our second validation approach was based on an independent WES analysis. After filtering for  $\geq 10\times$  coverage at the variant position, we could evaluate 67 variants in ALK4 and 42 variants in ALK6. We confirmed 61 variants (91%) in ALK4 and 40 variants (95%) in ALK6. When combining both strategies, we independently validated 114 of 123 (93%) variants in ALK4 and 45 of 48 (94%) variants in ALK6.

### 2.3. Variants Related to Crizotinib Treatment

When comparing the primary tumors with the resistant tumors for the four patients with paired samples revealed 175 putatively “treatment-related” variants in 156 genes, of which 136 variants (in 129 genes) were only found in the resistant tumor (Figure 2) (Supplementary Table S2). CADD scores > 20, indicative of deleterious variants, were observed for 16% of these variants (21 SNVs in 15 genes), whereas 43% (75 SNVs in 71 genes) had CADD scores between 10 and 20, 30% (53 SNVs in 49 genes) had CADD scores < 10, and 11% (19 INDELS in 19 genes) had no CADD scores (Supplementary Table S2). Fifteen INDELS caused a frameshift, and can thus be considered deleterious. The distribution of the variants over the different CADD score groups was similar to the distribution observed for all the somatic variants.



**Figure 2.** Comparison of mutant read frequencies (MRF) in primary tumor samples (x-axis) and resistant tumor samples (y-axis) from four patients. Each dot represents a single variant. Variants shown in the dashed trapezium have a MRF  $\geq 0.20$  in resistant samples and at least a two times higher MRF in the treatment-related sample as compared to the primary sample. The total number of “treatment-related” variants in each patient is indicated in the trapezium. In total, 176 variants were identified in 156 genes.

To focus on *ALK*-independent resistance mechanisms, we excluded genes that were mutated in *ALK4* because the resistant tumor carried a known *ALK*-resistance-associated mutation (*ALK* G1269A) that explained the crizotinib resistance. The three remaining paired pre- and post-treatment tumors had a total of 90 putatively “treatment-related” variants in 74 genes. Each gene was mutated in only one of the three patients. The 74 genes were involved in 105 pathways. A significant enrichment was observed for 15 pathways in Partek Genomics Suite 6.6 (Supplementary Table S3). Four of these pathways, harboring nine treatment-related mutated genes, were linked to a common biological theme: EMT. These EMT pathways were proteoglycans in cancer, HIF-1 signaling, FoxO signaling, and ECM-receptor interaction (Table 2). Using the Database for Annotation, Visualization and Integrated Discovery (DAVID) pathway analysis tool, we confirmed significant enrichment of proteoglycans in cancer and the HIF-1 signaling pathways, and found borderline significance for the FoxO signaling pathway. Using the Gene Network analysis tool, we found a borderline significant enrichment for the ECM-receptor interaction pathway but also that the other three pathways were not present in Gene Network. Analysis of the other EMT-related pathways in Partek [23] revealed three additional genes with treatment-related mutations (Table 2). In total we found 12 genes uniquely mutated in the resistant tumors that are linked to 13 EMT-related pathways. Variants of six of these genes (*ARNT*, *CTNNA3*, *PTPN11*, *MLTK*, *SMAD4*, and *VEGFA*) had CADD scores higher than 20. Another two genes, *LAMA2* and *MLTK*, had out-of-frame INDELS, which should be damaging.

**Table 2.** Overview of all epithelial-mesenchymal transition (EMT)-related pathways that include genes with resistance-related mutated in patients ALK6, ALK14, and ALK16.

Pathway Name	Patient(s)	Genes Mutated	Enrichment Score	Enrichment <i>p</i> -Value <sup>1</sup>
Proteoglycans in cancer <sup>2</sup>	ALK6 ALK14	<i>ANK2, FASLG, HSPG2, PTPN11, STAT3, VEGFA</i>	8.0	0.00
HIF-1 signaling pathway <sup>2</sup>	ALK6 ALK14	<i>ARNT, STAT3, VEGFA</i>	4.3	0.01
FoxO signaling pathway <sup>2</sup>	ALK6 ALK14	<i>FASLG, SMAD4, STAT3</i>	3.7	0.03
ECM-receptor interaction <sup>3</sup>	ALK6 ALK14	<i>HSPG2, LAMA2</i>	3.3	0.04
Adherens junction <sup>3</sup>	ALK6 ALK14	<i>CTNNA3, SMAD4</i>	3.0	0.05
Ras signaling pathway	ALK14	<i>FASLG, PTPN11, VEGFA</i>	2.7	0.07
Jak-STAT signaling pathway	ALK6 ALK14	<i>PTPN11, STAT3</i>	1.9	ns
PI3K-AKT signaling pathway	ALK14	<i>FASLG, LAMA2, VEGFA</i>	1.8	ns
Focal adhesion <sup>3</sup>	ALK14	<i>LAMA2, VEGFA</i>	1.6	ns
VEGF signaling pathway <sup>3</sup>	ALK14	<i>VEGFA</i>	1.3	ns
TGF-beta signaling pathway	ALK6	<i>SMAD4</i>	1.2	ns
MAPK signaling pathway	ALK14	<i>FASLG, MAP3K20</i>	1.1	ns
Wnt signaling pathway <sup>3</sup>	ALK6	<i>SMAD4</i>	1.0	ns
Regulation of actin cytoskeleton	ALK6	<i>SCIN</i>	0.5	ns

<sup>1</sup> *p*-value based on analysis with Partek; ns, not significant; <sup>2</sup> Borderline or significant in DAVID; <sup>3</sup> Borderline or significant in Gene Network; Bold type indicates additional genes not included in the four pathways significantly enriched in Partek.

In the one non-responding patient, ALK8, we observed somatic variants in three additional EMT-related genes (*ITGAM*, *CACNA1E*, and *RUVBL1*). These three genes are involved in cell adhesion molecules, MAPK signaling, Wnt signaling, and regulation of the actin cytoskeleton.

Immunohistochemistry (IHC) revealed changes in the E-cadherin and/or Vimentin staining intensity compatible with EMT for only three out of 11 cases (patients ALK20, ALK25, and ALK26). IHC was inconclusive for ALK14 and ALK16. This indicates that other resistant pathways may be operational, although technical issues, such as scarcity of tumor tissue and differences in tissue origin, hampered the reliability of our scoring.

Using a different analysis strategy, performing separate pathway analysis for genes mutated in each individual patient we identified the metabolism pathway as the only pathway that was shared by all five patients for whom we had paired primary and resistant tumor samples. In total, we found 15 variants in 13 genes from this pathway, of which four had CADD scores > 20 and one frameshift insertion. The genes are indicated in Supplementary Table S2. One gene, *CYP11B1*, was also mutated in the non-responding patient ALK8.

#### 2.4. Treatment-Related Copy Number Alterations

We compared WES-based copy number alteration (CNA) plots of the primary tumors to those of their paired resistant samples. Although we did identify several differences in CNAs between primary and resistant samples (Supplementary Figure S1), we did not see a copy number gain in any of the patients for the *ALK*, *MET*, or *KIT* loci known to be associated with resistance. Nor did we observe recurrent CNAs in other parts of the genome shared between these patients.

### 3. Discussion

In this study, we identified 175 variants in 156 genes that are enriched in crizotinib-resistant tumor samples as compared to matched pre-crizotinib samples. In the three patients without *ALK*-resistance-associated mutations, pathway analysis revealed a significant enrichment of nine genes in four EMT pathways. These four pathways were proteoglycans in cancer (*ANK2*, *FASLG*, *HSPG2*, *PTPN11*, *STAT3*, and *VEGFA*), HIF-1 signaling (*ARNT*, *STAT3*, and *VEGFA*), FoxO signaling (*FASLG*, *SMAD4*, and *STAT3*), and ECM-receptor interaction (*HSPG2* and *LAMA2*). No convincing differences in CNAs were observed between the primary and resistant tumor samples, making it

unlikely that gene copy number gains were associated with the resistance that was observed in the patients included in this study.

In our study, we have explored ALK-independent mechanisms of crizotinib resistance in advanced NSCLC patients by comparing variants observed in the crizotinib-resistant tumor to the variants present in the tumor before crizotinib treatment. Many mechanisms for crizotinib resistance have been proposed based on cell line studies and studies using resistant samples without comparison to the pre-treatment samples [25–27]. These studies revealed activation of several ALK-independent bypass mechanisms, including activation of EGFR, KRAS, SRC, and ERBB and MAPK signaling [25–27]. Although these studies generated valuable data, a direct comparison between pre- and post-crizotinib treatment samples is required to truly pinpoint the causal resistance-associated alterations.

Despite the relatively large number of patients in our previous paper, the total number of patients with matched samples eligible for this study was limited. A general problem in the clinic is the lack of re-biopsies from the growing tumor upon development of resistance. In patients with re-biopsies, the tumor cell percentages are usually low and the biopsy is often fully consumed by the routine diagnostic molecular tests needed to guide further treatment. The limited DNA quality of formalin-fixed paraffin-embedded (FFPE) tumor samples resulted in a suboptimal coverage for part of the exons, which may have prevented the detection of mutations that are present in small subclones of the tumor. However, this should not have affected the identification of treatment-related driver mutations, because they should be present in the majority of the tumor cells in the resistant samples.

With regard to ALK-dependent mechanisms, we observed an *ALK* G1269A mutation in the resistant tumor of patient ALK4. This mutation was not detected in the primary tumor, as consistent with our previous study using RNAseq on the resistant tumor and ddPCR to analyze the primary tumor sample [14]. This *ALK* gatekeeper mutation has been reported previously in several other studies as a resistance mechanism to crizotinib treatment in *ALK*-positive NSCLC patients [21,28,29].

For the remaining three patients, no *ALK* gatekeeper mutations were found, indicating the activation of *ALK*-independent bypass mechanisms. Here, we focused on treatment-related variants specific for, or with increased mutant read frequencies (MRFs) in, post-treatment samples. In addition to crizotinib treatment, there are several other factors that may have contributed to what we define as treatment-related variants. These factors include treatment history, time interval between biopsies, and differences in anatomic location of the biopsies. Indeed, in patient ALK16, whose biopsies are taken from the same anatomic location, the number of putative treatment-related variants is low. However, development of crizotinib resistance is the common factor in these patients, and thus we can assume that at least a subset of the mutations is directly related to crizotinib resistance. Since there was no commonly mutated gene in these three patients, *ALK*-independent mechanisms appear to be more diverse, a finding that is consistent with the broad variations in resistance mechanisms proposed in the literature [15,28]. To find a possible common mode of action, we performed pathway analysis on the treatment-related resistance genes in these three patients. Four of the significantly enriched pathways were EMT related. Our subsequent analysis of additional EMT-related pathways revealed a total of 12 mutated genes in two of the three patients with *ALK*-independent resistance mechanisms. Several studies have indicated EMT as a mediator of resistance against *ALK* inhibitors [21–23]. Silencing of Vimentin restored the responsiveness of a crizotinib-resistant cell line to *ALK* inhibitors [22], whereas in another cell line, the responsiveness to NMS could not be restored [23]. Nevertheless, both of the studies support a role of EMT in crizotinib resistance. Five out of eleven *ALK*-positive NSCLC patients treated with the second-generation *ALK* inhibitor ceritinib showed EMT involvement based on immunostaining for E-cadherin and Vimentin [21]. In our study, seven out of the 12 mutated genes involved in EMT had variants that may potentially impact the protein, based on either a CADD score >20 or the presence of out-of-frame INDELS.

Among the seven mutated genes with high CADD scores, *SMAD4* is known to be a crucial regulator of the TGF-beta signaling pathway, which is one of the most important mechanisms leading to EMT [24]. The *SMAD4* W323\* mutation observed in ALK6 is in the same functional MH2-domain

as the commonly observed R361H and R361C hotspot mutations. This supports a potential causal role for the *SMAD4* W323\* mutation. *PTPN11* positively regulates TGF1-induced EMT in lung cancer [30], and the *PTPN11* E76K mutation that is detected in ALK14 was proven to be an activating mutation in lung cancer cells [31]. Overexpression of *VEGFA* in breast cell lines induced SNAIL protein levels, which further reduced E-cadherin expression levels [32]. This suggests that if the *VEGFA* G328R observed in ALK14 is indeed an activating mutation, it may induce loss of E-cadherin. The relevance of the other four genes with high CADD score mutations (*ARNT* F446L, *CTNNA3* E690K, and *MLTK* F157L) or a deleterious out-of-frame deletion (*LAMA2* P596Sfs\*13) remains unclear.

Immunostaining for EMT-related protein expression was inconclusive in the subset of available paired pre- and post-treatment samples. This was due to difficulties in scoring caused by differences in tissue origin and scarcity of tumor cells. Moreover, the four EMT-related pathways that are discussed above are also involved in various other biological processes, and these processes might also be related to resistance independent of their association with EMT. Additional cases need to be studied to fully explore the resistance mechanisms that are related to crizotinib resistance. Thus, our data cannot exclude that additional pathways may be involved in the resistance. The metabolism pathway may be a good candidate: it is the only pathway implicated in all five patients and dysregulation of the metabolism pathway has been linked to drug resistance in various types of cancer [33,34]. *PDE4D* (Ser14Phefs\*11), which was mutated in ALK16, was shown to impair cAMP generation through cAMP hydrolysis in lung cancer cell lines. Inhibitors of PDE4 decrease cell proliferation and angiogenesis [35]. Thus, inhibition of the metabolism pathway may provide additional mechanisms for overcoming crizotinib resistance.

When crizotinib resistance occurs, second-generation ALK inhibitors (ceritinib, alectinib, brigatinib), and, only subsequently, later-generation inhibitors (lorlatinib) are advised as next treatment options, even when the *ALK*-gatekeeper mutation status is unknown [36]. A more optimal approach would be to determine the presence of *ALK*-gatekeeper mutations in progressing tumors in order to justify the choice of the next ALK inhibitor. It is unclear whether tumors that are driven by ALK-independent resistance mechanisms also show tumor response to these drugs. In our study, one of the three patients with an ALK-independent resistance mechanism had a good response to alectinib, suggesting that this drug, in addition to being ALK-specific, may target ALK-independent resistance mechanisms [37,38].

## 4. Materials and Methods

### 4.1. Patient Inclusion

In a previous study we described 29 *ALK*-positive non-squamous NSCLC patients who were treated with crizotinib [12]. Their *ALK* status was assessed by immunohistochemistry. Patients involved in our earlier study gave informed consent for the use of their biopsies for further analysis. For four of these 29 patients, we had sufficient FFPE or frozen pre- and post-treatment tumor material to carry out WES. The four patients are designated ALK4, ALK6, ALK14, and ALK16 in concordance with our previous study [12]. In addition to these samples, we also had FFPE tumor material from one patient, ALK8, who had no response to crizotinib. We had previously found an *ALK* G1269A gatekeeper mutation in ALK4 by RNA-sequencing of the resistant tumor sample of this patient [14]. For each patient, white blood cells were isolated from peripheral venous blood using standard protocols. Tumor samples from patient ALK14 before crizotinib treatment and from patients ALK4 and ALK6 upon developing resistance to crizotinib treatment were fresh frozen. All the other tumor specimens were FFPE. Macrodissection was applied to the tumor samples to achieve tumor cell content greater than 60%. Patients provided informed consent for procurement of extra tumor tissue and analysis of tissue specimen and was approved by the Medical Ethical Committee.

#### 4.2. DNA Isolation

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

#### 4.3. Whole Exome Sequencing

WES was carried out on 0.6–2 µg genomic DNA of normal and tumor-derived DNA samples (BGI Tech Solutions Co. Ltd., Hong Kong, China). Target enrichment was done using the Agilent SureSelect All Exon V5 kit (Agilent technologies, Santa Clara, CA, USA). Paired-end sequencing with a read length of  $2 \times 90$  nt was performed on Illumina HiSeq2000.

As a part of the validation procedure, we performed WES a second time on the crizotinib-resistant samples from patients ALK4 and ALK6, following our previously published protocol and data analysis pipeline [14]. RNA-sequencing data from ALK4 and ALK6 resistant samples were re-analyzed to confirm the presence of mutations [14].

#### 4.4. Immunohistochemistry

Eleven paired formalin-fixed tissue samples from our previous study (ALK1, ALK4, ALK5, ALK6, ALK14, ALK16, ALK17, ALK20, ALK22, ALK25, and ALK26) were used for IHC staining of two EMT biomarkers: E-cadherin (using mouse monoclonal E-cadherin 1:100, clone 36/E-cadherin; BD Biosciences, Breda, The Netherlands) and Vimentin (using mouse monoclonal Vimentin 1:100, clone sc-6260; Santa Cruz Biotechnology; Bioconnect, Huissen, The Netherlands).

#### 4.5. Data Analysis

Raw reads were processed using our in-house pipeline, as described previously (24). In brief, our variant calling pipeline is based on the Genome Analysis Toolkit (GATK) workflow and uses Molgenis Compute as workflow management software [39]. Read alignment was done using Burrows-Wheeler Aligner and GATK, using the human genome reference build GRCh37 with decoys from the GATK bundle [39,40]. Picard Tools was used for format conversion and marking duplicates. HaplotypeCaller was used for integrated calling of the variants for all of the samples from the same patient/cohort. Variants were annotated using SnpEff/SnpSift with the Ensembl release 75 gene annotations and the dbNSFP2.7 database [41]. GATK was used to identify variants. Variants were annotated with dbsnp 138, Cosmic v72, 1000 genomes phase 3 and the ExAC 0.3 databases [42–44]. The data were filtered for quality metrics similarly to GATK recommendations using custom filters for population frequency and variant effect. Based on the presence of variants in the normal sample, germline variants were excluded from further analysis. Additionally variants were excluded if present in 100 genomes phase 3 with a minor allele frequency  $> 0.02$ .

CADD scores were used to predict pathogenicity of the identified variants. Variants with a CADD score  $\geq 20$  are defined as deleterious, those between 10 and 20 possibly deleterious, and those  $< 10$  non-deleterious [45].

To identify somatic mutations we first excluded variants for which the total number of reads in the normal sample was  $< 10$ . Next, we excluded all of the variants for which one or more mutant reads were present in the normal sample. We then excluded variants with total reads  $< 10$  in either the primary or the resistant samples. The remaining variants, with two or more mutant reads in either the pre-treatment tumor or post-treatment tumor samples, were considered somatic. Variants detected in the resistant samples with MRFs  $\geq 0.20$ , and for which the MRF is at least twice the MRF in the paired primary sample, are indicated as "treatment-related" variants.



Pathways overrepresented in the set of genes mutated specifically in the crizotinib-resistance tumor samples were identified with the Partek Genomics Suite 6.6 software package (Partek Inc., St. Louis, MO, USA). In addition, we used Gene Network [46] and DAVID v6.8 [47] to validate our findings by Partek.

#### 4.6. WES-Based Analysis of CNAs

Pseudo probe data were generated with SAMtools VarScan2 and DNACopy [48–50]. Briefly, when frozen tumor tissue was used for the isolation of high quality DNA samples, pseudo-probe-derived GC-normalized log<sub>2</sub> copy number ratios were generated for each sample using the corresponding normal sample. For FFPE tumor samples, we used WES data of a merged pool of normal FFPE samples as the reference. All of the alignments with a mapping quality > 40, in combination with a minimal segment size of 2 kb and a maximal segment size of 5 kb, with a mean base-wise coverage of at least 1×, were used to calculate the ratios. DNA copy number variant calls were compared between the resistant and the primary tumor sample based on assigned ploidity and by visual inspection of the copy number changes.

## 5. Conclusions

In conclusion, we have identified mutations in genes involved in EMT-related and metabolism pathways in crizotinib-resistant tumors without *ALK*-gatekeeper mutations. From a clinical perspective, the mutational status of patients may provide therapeutic guidance for clinical management of their NSCLC by targeting loss of epithelial differentiation and/or metabolism pathways.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2072-6694/10/1/10/s1>. Figure S1: Copy number variant plots across four patients with *ALK* rearranged advanced NSCLC with matched primary and resistant tumor samples. Table S1: Whole exome sequencing data in five patient samples; Table S2: Treatment-related variants in all tumor samples in *ALK4*, *ALK6*, *ALK14*, and *ALK16*; Table S3: Significantly enriched pathways harboring treatment-related variants in patients *ALK6*, *ALK14* and *ALK16*.

**Acknowledgments:** This work was supported by a grant from the University Medical Centre Groningen. We would like to acknowledge Kate McIntyre for revising the manuscript, and Tineke van der Sluis for technical assistance with the experimental work. We thank the UMCG Genomics Coordination center, the UG Center for Information Technology and their sponsors BBMRI-NL and TarGet for storage and compute infrastructure. We would like to thank the Exome Aggregation Consortium and the groups that provided exome variant data for comparison. A full list of contributing groups can be found at <http://exac.broadinstitute.org/about>.

**Author Contributions:** Klaas Kok, Anke van den Berg, Harry J. M. Groen, Wim Timens, Ed Schuurung designed the study. Anthonie J. van der Wekken, T. Jeroen N. Hiltermann collected clinical samples and patient information. Ali Saber and Anke van den Berg were involved DNA isolation and sequencing. Miente M. Terpstra performed bioinformatics analysis and Jiacong Wei and Klaas Kok were involved in the interpretation of data. Jiacong Wei drafted the manuscript. All authors were involved in the revision of the manuscript and approved the final version of the submitted manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

## References

1. Malik, P.S.; Raina, V. Lung cancer: Prevalent trends & emerging concepts. *Indian J. Med. Res.* **2015**, *141*, 5–7. [PubMed]
2. Mountain, C.F. Revisions in the international system for staging lung cancer. *Chest* **1997**, *111*, 1710–1717. [CrossRef] [PubMed]
3. El-Telbany, A.; Ma, P.C. Cancer genes in lung cancer: Racial disparities: Are there any? *Genes Cancer* **2012**, *3*, 467–480. [CrossRef] [PubMed]
4. Morris, S.W.; Kirstein, M.N.; Valentine, M.B.; Dittmer, K.G.; Shapiro, D.N.; Saltman, D.L.; Look, A.T. Fusion of a kinase gene, *ALK*, to a nucleolar protein gene, *NPM*, in non-Hodgkin's lymphoma. *Science* **1994**, *263*, 1281–1284. [CrossRef] [PubMed]

5. Soda, M.; Choi, Y.L.; Enomoto, M.; Takada, S.; Yamashita, Y.; Ishikawa, S.; Fujiwara, S.-I.; Watanabe, H.; Kurashina, K.; Hatanaka, H.; et al. Identification of the transforming EML4–ALK fusion gene in non-small-cell lung cancer. *Nature* **2007**, *448*, 561–566. [[CrossRef](#)] [[PubMed](#)]
6. Solomon, B.; Varella-Garcia, M.; Camidge, D.R. ALK gene rearrangements a new therapeutic target in a molecularly defined subset of non-small cell lung cancer. *J. Thorac. Oncol.* **2009**, *4*, 1450–1454. [[CrossRef](#)] [[PubMed](#)]
7. Soria, J.-C.; Tan, D.S.W.; Chiari, R.; Wu, Y.-L.; Paz-Ares, L.; Wolf, J.; Geater, S.L.; Orlov, S.; Cortinovis, D.; Yu, C.-J.; et al. First-line ceritinib versus platinum-based chemotherapy in advanced ALK-rearranged non-small-cell lung cancer (ascend-4): A randomised, open-label, phase 3 study. *Lancet* **2017**, *389*, 917–929. [[CrossRef](#)]
8. Shaw, A.T.; Kim, D.-W.; Nakagawa, K.; Seto, T.; Crinó, L.; Ahn, M.-J.; De Pas, T.; Besse, B.; Solomon, B.J.; Blackhall, F. Crizotinib versus chemotherapy in advanced ALK-positive lung cancer. *N. Engl. J. Med.* **2013**, *368*, 2385–2394. [[CrossRef](#)] [[PubMed](#)]
9. Shaw, A.T.; Gandhi, L.; Gadgeel, S.; Riely, G.J.; Cetnar, J.; West, H.; Camidge, D.R.; Socinski, M.A.; Chiappori, A.; Mekhail, T. Alectinib in ALK-positive, crizotinib-resistant, non-small-cell lung cancer: A single-group, multicentre, phase 2 trial. *Lancet Oncol.* **2016**, *17*, 234–242. [[CrossRef](#)]
10. Gettinger, S.N.; Bazhenova, L.A.; Langer, C.J.; Salgia, R.; Gold, K.A.; Rosell, R.; Shaw, A.T.; Weiss, G.J.; Tugnait, M.; Narasimhan, N.I.; et al. Activity and safety of brigatinib in ALK-rearranged non-small-cell lung cancer and other malignancies: A single-arm, open-label, phase 1/2 trial. *Lancet Oncol.* **2016**, *17*, 1683–1696. [[CrossRef](#)]
11. Solomon, B.J.; Mok, T.; Kim, D.-W.; Wu, Y.-L.; Nakagawa, K.; Mekhail, T.; Felip, E.; Cappuzzo, F.; Paolini, J.; Usari, T.; et al. First-line crizotinib versus chemotherapy in ALK-positive lung cancer. *N. Engl. J. Med.* **2014**, *371*, 2167–2177. [[CrossRef](#)] [[PubMed](#)]
12. Van der Wekken, A.; Pelgrim, R.; 't Hart, N.; Werner, N.; Mastik, M.; Hendriks, L.; van der Heijden, E.H.; Looijen-Salamon, M.; de Langen, A.J.; Staal-van den Brekel, J.; et al. Dichotomous ALK-IHC is a better predictor for ALK inhibition outcome than traditional ALK-FISH in advanced non-small cell lung cancer. *Clin. Cancer Res.* **2017**, *23*, 4251–4258. [[CrossRef](#)] [[PubMed](#)]
13. Camidge, D.R.; Doebele, R.C. Treating ALK-positive lung cancer—early successes and future challenges. *Nat. Rev. Clin. Oncol.* **2012**, *9*, 268–277. [[CrossRef](#)] [[PubMed](#)]
14. Saber, A.; van der Wekken, A.J.; Kok, K.; Terpstra, M.M.; Bosman, L.J.; Mastik, M.F.; Timens, W.; Schuurings, E.; Hiltermann, T.J.; Groen, H.J.; et al. Genomic aberrations in crizotinib resistant lung adenocarcinoma samples identified by transcriptome sequencing. *PLoS ONE* **2016**, *11*. [[CrossRef](#)] [[PubMed](#)]
15. Van der Wekken, A.J.; Saber, A.; Hiltermann, T.J.; Kok, K.; van den Berg, A.; Groen, H.J. Resistance mechanisms after tyrosine kinase inhibitors afatinib and crizotinib in non-small cell lung cancer, a review of the literature. *Crit. Rev. Oncol. Hematol.* **2016**, *100*, 107–116. [[CrossRef](#)] [[PubMed](#)]
16. Katayama, R.; Shaw, A.T.; Khan, T.M.; Mino-Kenudson, M.; Solomon, B.J.; Halmos, B.; Jessop, N.A.; Wain, J.C.; Yeo, A.T.; Benes, C. Mechanisms of acquired crizotinib resistance in ALK-rearranged lung cancers. *Sci. Transl. Med.* **2012**, *4*, 120ra117. [[CrossRef](#)] [[PubMed](#)]
17. Gouji, T.; Takashi, S.; Mitsuhira, T.; Yukito, I. Crizotinib can overcome acquired resistance to ch5424802: Is amplification of the MET gene a key factor? *J. Thorac. Oncol.* **2014**, *9*, e27–e28. [[CrossRef](#)] [[PubMed](#)]
18. Katayama, R.; Khan, T.M.; Benes, C.; Lifshits, E.; Ebi, H.; Rivera, V.M.; Shakespeare, W.C.; Iafrate, A.J.; Engelman, J.A.; Shaw, A.T. Therapeutic strategies to overcome crizotinib resistance in non-small cell lung cancers harboring the fusion oncogene EML4-ALK. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 7535–7540. [[CrossRef](#)] [[PubMed](#)]
19. Dagogo-Jack, I.; Shaw, A.T. Crizotinib resistance: Implications for therapeutic strategies. *Ann. Oncol.* **2016**, *27*, iii42–iii50. [[CrossRef](#)] [[PubMed](#)]
20. Toyokawa, G.; Seto, T. Updated evidence on the mechanisms of resistance to ALK inhibitors and strategies to overcome such resistance: Clinical and preclinical data. *Oncol. Res. Treat.* **2015**, *38*, 291–298. [[CrossRef](#)] [[PubMed](#)]
21. Gainor, J.F.; Dardaei, L.; Yoda, S.; Friboulet, L.; Leshchiner, I.; Katayama, R.; Dagogo-Jack, I.; Gadgeel, S.; Schultz, K.; Singh, M. Molecular mechanisms of resistance to first- and second-generation ALK inhibitors in ALK-rearranged lung cancer. *Cancer Discov.* **2016**, *6*, 1118–1133. [[CrossRef](#)] [[PubMed](#)]

22. Kim, H.R.; Kim, W.S.; Choi, Y.J.; Choi, C.M.; Rho, J.K.; Lee, J.C. Epithelial-mesenchymal transition leads to crizotinib resistance in H-2228 lung cancer cells with EML4-ALK translocation. *Mol. Oncol.* **2013**, *7*, 1093–1102. [[CrossRef](#)] [[PubMed](#)]
23. Gower, A.; Hsu, W.H.; Hsu, S.T.; Wang, Y.; Giaccone, G. EMT is associated with, but does not drive resistance to ALK inhibitors among EML4-ALK non-small cell lung cancer. *Mol Oncol.* **2016**, *10*, 601–609. [[CrossRef](#)] [[PubMed](#)]
24. Lamouille, S.; Xu, J.; Derynck, R. Molecular mechanisms of epithelial–mesenchymal transition. *Nat. Rev. Mol. Cell Biol.* **2014**, *15*, 178–196. [[CrossRef](#)] [[PubMed](#)]
25. Kimura, M.; Endo, H.; Inoue, T.; Nishino, K.; Uchida, J.; Kumagai, T.; Kukita, Y.; Kato, K.; Imamura, F.; Inoue, M. Analysis of ERBB ligand-induced resistance mechanism to crizotinib by primary culture of lung adenocarcinoma with EML4-ALK fusion gene. *J. Thorac. Oncol.* **2015**, *10*, 527–530. [[CrossRef](#)] [[PubMed](#)]
26. Azzato, E.M.; Deshpande, C.; Aikawa, V.; Aggarwal, C.; Alley, E.; Jacobs, B.; Morrisette, J.; Daber, R. Rare complex mutational profile in an ALK inhibitor-resistant non-small cell lung cancer. *Anticancer Res.* **2015**, *35*, 3007–3012. [[PubMed](#)]
27. Crystal, A.S.; Shaw, A.T.; Sequist, L.V.; Friboulet, L.; Niederst, M.J.; Lockerman, E.L.; Frias, R.L.; Gainor, J.F.; Amzallag, A.; Greninger, P. Patient-derived models of acquired resistance can identify effective drug combinations for cancer. *Science* **2014**, *346*, 1480–1486. [[CrossRef](#)] [[PubMed](#)]
28. Doebele, R.C.; Pilling, A.B.; Aisner, D.L.; Kutateladze, T.G.; Le, A.T.; Weickhardt, A.J.; Kondo, K.L.; Linderman, D.J.; Heasley, L.E.; Franklin, W.A.; et al. Mechanisms of resistance to crizotinib in patients with ALK gene rearranged non-small cell lung cancer. *Clin. Cancer Res.* **2012**, *18*, 1472–1482. [[CrossRef](#)] [[PubMed](#)]
29. Kim, S.; Kim, T.M.; Kim, D.W.; Go, H.; Keam, B.; Lee, S.H.; Ku, J.L.; Chung, D.H.; Heo, D.S. Heterogeneity of genetic changes associated with acquired crizotinib resistance in ALK-rearranged lung cancer. *J. Thorac. Oncol.* **2013**, *8*, 415–422. [[CrossRef](#)] [[PubMed](#)]
30. Li, S.; Wang, L.; Zhao, Q.; Liu, Y.; He, L.; Xu, Q.; Sun, X.; Teng, L.; Cheng, H.; Ke, Y. SHP2 positively regulates TGF $\beta$ 1-induced epithelial-mesenchymal transition modulated by its novel interacting protein Hook1. *J. Biol. Chem.* **2014**, *289*, 34152–34160. [[CrossRef](#)] [[PubMed](#)]
31. Yu, B.; Liu, W.; Yu, W.-M.; Loh, M.L.; Alter, S.; Guvench, O.; MacKerell, A.D.; Tang, L.-D.; Qu, C.-K. Targeting protein tyrosine phosphatase SHP2 for the treatment of PTPN11-associated malignancies. *Mol. Cancer Ther.* **2013**, *12*, 1738–1748. [[CrossRef](#)] [[PubMed](#)]
32. Wanami, L.S.; Chen, H.-Y.; Peiró, S.; de Herreros, A.G.; Bachelder, R.E. Vascular endothelial growth factor—A stimulates Snail expression in breast tumor cells: Implications for tumor progression. *Exp. Cell Res.* **2008**, *314*, 2448–2453. [[CrossRef](#)] [[PubMed](#)]
33. Rahman, M.; Hasan, M.R. Cancer metabolism and drug resistance. *Metabolites* **2015**, *5*, 571–600. [[CrossRef](#)] [[PubMed](#)]
34. Zhao, Y.; Butler, E.B.; Tan, M. Targeting cellular metabolism to improve cancer therapeutics. *Cell Death Dis.* **2013**, *4*, e532. [[CrossRef](#)] [[PubMed](#)]
35. Pullamsetti, S.; Banat, G.; Schmall, A.; Szibor, M.; Pomagruk, D.; Hänze, J.; Kolosionek, E.; Wilhelm, J.; Braun, T.; Grimminger, F. Phosphodiesterase-4 promotes proliferation and angiogenesis of lung cancer by crosstalk with HIF. *Oncogene* **2013**, *32*, 1121–1134. [[CrossRef](#)] [[PubMed](#)]
36. Wu, J.; Savooji, J.; Liu, D. Second-and third-generation ALK inhibitors for non-small cell lung cancer. *J. Hematol. Oncol.* **2016**, *9*, 19. [[CrossRef](#)] [[PubMed](#)]
37. Awad, M.M.; Shaw, A.T. ALK inhibitors in non-small cell lung cancer: Crizotinib and beyond. *Clin. Adv. Hematol. Oncol.* **2014**, *12*, 429–439. [[PubMed](#)]
38. Ou, S.-H.I.; Ahn, J.S.; Petris, L.D.; Govindan, R.; Yang, J.C.-H.; Hughes, B.; Lena, H.; Moro-Sibilot, D.; Bearz, A.; Ramirez, S.V.; et al. Alectinib in crizotinib-refractory ALK-rearranged non-small-cell lung cancer: A phase II global study. *J. Clin. Oncol.* **2016**, *34*, 661–668. [[CrossRef](#)] [[PubMed](#)]
39. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytzky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M. The genome analysis toolkit: A mapreduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **2010**, *20*, 1297–1303. [[CrossRef](#)] [[PubMed](#)]
40. Li, H.; Durbin, R. Fast and accurate long-read alignment with burrows–wheeler transform. *Bioinformatics* **2010**, *26*, 589–595. [[CrossRef](#)] [[PubMed](#)]
41. Liu, X.; Jian, X.; Boerwinkle, E. dbNSFP v2.0: A database of human non-synonymous SNVs and their functional predictions and annotations. *Hum. Mutat.* **2013**, *34*, E2393–E2402. [[CrossRef](#)] [[PubMed](#)]

42. Forbes, S.A.; Beare, D.; Gunasekaran, P.; Leung, K.; Bindal, N.; Boutselakis, H.; Ding, M.; Bamford, S.; Cole, C.; Ward, S. Cosmic: Exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res.* **2015**, *43*, D805–D811. [[CrossRef](#)] [[PubMed](#)]
43. Consortium, G.P. An integrated map of genetic variation from 1,092 human genomes. *Nature* **2012**, *491*, 56–65.
44. Lek, M.; Karczewski, K.; Minikel, E.; Samocha, K.; Banks, E.; Fennell, T.; O'Donnell-Luria, A.; Ware, J.; Hill, A.; Cummings, B. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **2016**, *536*, 285–291. [[CrossRef](#)] [[PubMed](#)]
45. Kircher, M.; Witten, D.M.; Jain, P.; O'Roak, B.J.; Cooper, G.M.; Shendure, J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **2014**, *46*, 310. [[CrossRef](#)] [[PubMed](#)]
46. Fehrmann, R.S.N.; Karjalainen, J.M.; Krajewska, M.; Westra, H.-J.; Maloney, D.; Simeonov, A.; Pers, T.H.; Hirschhorn, J.N.; Jansen, R.C.; Schultes, E.A.; et al. Gene expression analysis identifies global gene dosage sensitivity in cancer. *Nat. Genet.* **2015**, *47*, 115–125. [[CrossRef](#)] [[PubMed](#)]
47. Huang, D.W.; Sherman, B.T.; Lempicki, R.A. Systematic and integrative analysis of large gene lists using david bioinformatics resources. *Nat. Protoc.* **2009**, *4*, 44–57. [[CrossRef](#)] [[PubMed](#)]
48. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997
49. Koboldt, D.C.; Zhang, Q.; Larson, D.E.; Shen, D.; McLellan, M.D.; Lin, L.; Miller, C.A.; Mardis, E.R.; Ding, L.; Wilson, R.K. varScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **2012**, *22*, 568–576. [[CrossRef](#)] [[PubMed](#)]
50. Venkatraman, E.S.; Olshen, A.B. A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics* **2007**, *23*, 657–663. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).