

## ARTICLE OPEN



# Early identification of disease progression in ALK-rearranged lung cancer using circulating tumor DNA analysis

Arlou Kristina Angeles<sup>1,2,12</sup>, Petros Christopoulos<sup>2,3,12</sup>, Zhao Yuan<sup>4</sup>, Simone Bauer<sup>1,2</sup>, Florian Janke<sup>1,2,5</sup>, Simon John Ogrodnik<sup>1,2</sup>, Martin Reck<sup>6</sup>, Matthias Schlesner<sup>1,2,4,10</sup>, Michael Meister<sup>2,7</sup>, Marc A. Schneider<sup>2,7</sup>, Steffen Dietz<sup>1,2,11</sup>, Albrecht Stenzinger<sup>8,9</sup>, Michael Thomas<sup>2,3</sup> and Holger Sültmann<sup>1,2</sup>✉

Targeted kinase inhibitors improve the prognosis of lung cancer patients with *ALK* alterations (ALK+). However, due to the emergence of acquired resistance and varied clinical trajectories, early detection of disease progression is warranted to guide patient management and therapy decisions. We utilized 343 longitudinal plasma DNA samples from 43 ALK+ NSCLC patients receiving ALK-directed therapies to determine molecular progression based on matched panel-based targeted next-generation sequencing (tNGS), and shallow whole-genome sequencing (sWGS). *ALK*-related alterations were detected in 22 out of 43 (51%) patients. Among 343 longitudinal plasma samples analyzed, 174 (51%) were ctDNA-positive. *ALK* variant and fusion kinetics generally reflected the disease course. Evidence for early molecular progression was observed in 19 patients (44%). Detection of ctDNA at therapy baseline indicated shorter times to progression compared to cases without mutations at baseline. In patients who succumbed to the disease, ctDNA levels were highly elevated towards the end of life. Our results demonstrate the potential utility of these NGS assays in the clinical management of ALK+ NSCLC.

*npj Precision Oncology* (2021)5:100; <https://doi.org/10.1038/s41698-021-00239-3>

## INTRODUCTION

About 3–7% of non-small-cell lung cancer (NSCLC) belong to a molecular subgroup defined by the presence of *ALK* (anaplastic lymphoma kinase) rearrangements<sup>1</sup>. Since its discovery, *ALK*-rearranged (ALK+) NSCLC has been a model disease for targeted therapies using tyrosine kinase inhibitors (TKIs) that potently attenuate the function of *ALK*<sup>2–5</sup>. While initial responses to selective TKIs are durable, ALK+ tumors eventually and inevitably develop resistance to targeted therapy. For example, the first FDA/EMA-approved *ALK* inhibitor crizotinib elicited improved objective response rates and progression-free survival (PFS) in randomized phase III trials compared to chemotherapy, but disease relapse was nonetheless observed within 1 year of treatment<sup>6</sup>. Consequently, significant insights into the molecular underpinnings of TKI therapy resistance have been gained. At the same time, increasingly potent and selective next-generation *ALK* TKIs have also been developed and approved for clinical use<sup>7–10</sup>. Although sequential therapy using next-generation *ALK* inhibitors improves PFS and mitigates brain metastases<sup>11</sup>, ALK+ tumors continue to adapt and develop alternative resistance mechanisms against these drugs.

The dynamic adaptability of ALK+ tumors against *ALK* inhibitors requires a personalized approach to disease monitoring for accurate and timely clinical patient management. Next-generation sequencing (NGS)-based analysis of tissue DNA is an established and a highly robust method for baseline genotypic evaluations<sup>12,13</sup>. However, while longitudinal tissue biopsies could

identify emerging resistance mutations, this strategy is limited by a number of factors: (a) single-site tumor biopsies would most likely fail to capture the full spectrum of genomic alterations of a molecularly heterogeneous tumor<sup>14</sup>; (b) multiple tumor re-biopsies pose procedural risks, and sometimes are not feasible<sup>15</sup>; and (c) tissue sample preparation such as formalin fixation could lead to false-positive results in molecular assays due to high levels of base transitions<sup>15,16</sup>. Therefore, liquid biopsy technologies are currently emerging as minimally invasive and easily accessible alternatives to tissue biopsies. Peripheral blood plasma naturally harbors molecular components that can be analyzed and attributed to certain pathologies. Among these, the circulating tumor DNA (ctDNA) compartment provides highly sensitive and insightful genetic information for evaluating tumor heterogeneity and clonal evolution. In addition to being a surrogate for localized tissue biopsy, ctDNA has the potential to capture the complete genomic profile of the primary tumor and metastases at expansive time points without spatial bias. Detection of such exhaustive genetic tumor profiles is particularly relevant for disease monitoring of patients undergoing treatment, where tumor subclones could emerge due to selective pressures introduced by various therapeutic challenges.

Previously, overall ctDNA levels have been utilized as independent biomarkers of disease progression or therapy response in multiple cancer entities including metastatic breast cancer<sup>17</sup>, melanoma<sup>18</sup>, metastatic gastrointestinal cancer<sup>19</sup>, metastatic bladder cancer<sup>20</sup>, and lung cancer<sup>21</sup>. ctDNA detection has also been explored for prediction of disease recurrence and minimal

<sup>1</sup>Division of Cancer Genome Research, German Cancer Research Center (DKFZ), German Cancer Consortium (DKTK), and National Center for Tumor Diseases (NCT), Heidelberg, Germany. <sup>2</sup>Translational Lung Research Center Heidelberg, German Center for Lung Research (DZL), Heidelberg, Germany. <sup>3</sup>Department of Oncology, Thoraxklinik and National Center for Tumor Disease (NCT) at Heidelberg University Hospital, Heidelberg, Germany. <sup>4</sup>Bioinformatics and Omics Data Analytics, German Cancer Research Center (DKFZ), Heidelberg, Germany. <sup>5</sup>Medical Faculty, Heidelberg University, Heidelberg, Germany. <sup>6</sup>Lung Clinic Grosshansdorf, Airway Research Center North, German Center for Lung Research, Grosshansdorf, Germany. <sup>7</sup>Translational Research Unit, Thoraxklinik at Heidelberg University Hospital, Heidelberg, Germany. <sup>8</sup>Institute of Pathology, Heidelberg University, Heidelberg, Germany. <sup>9</sup>German Cancer Consortium (DKTK), Heidelberg, Germany. <sup>10</sup>Present address: Biomedical Informatics, Data Mining and Data Analytics, Faculty for Applied Informatics, Augsburg University, Augsburg, Germany. <sup>11</sup>Present address: AstraZeneca GmbH, Wedel, Germany. <sup>12</sup>These authors contributed equally: Arlou Kristina Angeles, Petros Christopoulos. ✉email: [h.sueltmann@dkfz.de](mailto:h.sueltmann@dkfz.de)

residual disease<sup>22–24</sup>. In ALK+ NSCLC, liquid biopsies have successfully been utilized for ALK resistance mutation profiling, presenting important implications for therapy decisions<sup>21,25,26</sup>.

We first reported the potential clinical utility of a combination of cfDNA assays for longitudinal monitoring of ALK+ NSCLC under TKI therapy<sup>27</sup>. We performed matched capture-based high coverage targeted NGS (tNGS) using a commercially available panel comprised of genes optimized for longitudinal tumor burden monitoring in lung and colorectal cancers. Simultaneously, we applied the trimmed median absolute deviation from the copy number neutral state (t-MAD) score using shallow whole-genome sequencing (sWGS) to quantify global copy number changes derived from ctDNA<sup>28</sup>. The information derived from these ctDNA assays revealed the complex mutational kinetics of ALK+ tumors across therapy lines, as well as the increasing genomic instability of the tumor after sequential TKI treatments. Here, we extend the number of longitudinal plasma samples from our initial ALK+ NSCLC cohort and for the first time, using tNGS and sWGS data, identify threshold values for changes in variant allele frequencies (VAFs) and t-MAD scores that potentially indicate disease progression earlier than current conventional modalities. With additional sampling points, we show that a combined NGS assay approach can reveal molecular progression that precedes radiographic assessment, particularly in patients with detectable ctDNA. Furthermore, our study cohort demonstrates the potential of targeted sequencing in baseline risk stratification of patients, and corroborates previous data about the negative prognostic impact of *TP53* alterations detected during therapy.

## RESULTS

### Patient characteristics

In total, 43 metastatic ALK+ patients with corresponding 343 longitudinal plasma samples were considered for NGS analysis (Fig. 1a, see Methods). The median age of the study population was 57 years (range 39–80), with almost equal distribution of males and females (53% males). Seventy eight percent of patients were never/light-smokers (<10 pack-years, Table 1). Based on tissue biopsies, *EML4-ALK* variant 1 (V1, E13:A20) was identified in 35% (15/43), V2 (E20:A20) in 16% (7/43), V3 (E6:A20) in 30% (13/43), and other variants in 9% (4/43) of tumors. Other cases of ALK fusion were detected with tissue RNA-NGS at initial diagnosis. *TP53* mutation were detectable in 11/39 (28%) patients with available tissue DNA at diagnosis. The therapy regimens that patients underwent at each sampling time point are shown in Fig. 1b. The first-generation TKI crizotinib was administered to 25/43 patients. Second-generation TKIs (ceritinib, alectinib, brigatinib) were administered to 36/43 patients, and third-generation TKI lorlatinib was given to 4/43 patients.

### Genomic alteration landscape of ALK+NSCLC ctDNA

Among the patients under study, 22 (51%) had detectable ALK alterations (mutation or rearrangement) in at least one plasma sample based on tNGS profiling (Fig. 1a and Supplementary Table 1). A total of 36 ALK mutational events were detected, 34 of which were known or probable resistance mutations, while two were silent mutations. Of the remaining 21 patients, 11 (52%) had either intracranial progression ( $n = 6$ ) or stable disease at all sampling points ( $n = 5$ ), potentially explaining the low amount of ctDNA shedding and unmeasurable ALK alterations in plasma DNA (Fig. 1c). ctDNA was detected in a total of 174/343 (51%) plasma samples. Most of these were missense, synonymous, and truncating mutations, disruptive inframe insertions, as well as fusions and copy number changes in ALK and *TP53* with 56% and 52% detection rate in ctDNA (+) plasma samples, respectively. Other genes prominently altered were *MET*, *EGFR*, *KRAS*, *ERBB2*,

*PDGFRA*, *KIT*, *RET*, *NRAS*, *BRAF*, *APC*, and *BRCA1* (Fig. 1d). Copy number gains were also detected in *MET*, *EGFR*, and *ERBB2*.

### Dynamics of ALK rearrangement and resistance mutations during TKI therapy

We interrogated the concordance between ALK mutation or fusion abundance in ALK ctDNA+ ( $n = 22$ ) patients and disease status. Emergence of at least one ALK mutation coincided with progressive disease in eight of these patients (Supplementary Table 2), while seven patients (ALK\_06, 13, 18, 20, 25, 44, 62) showed increasing *EML4-ALK* fusion abundance upon progression on independent therapy lines (Supplementary Fig. 1). Five patients developed multiple secondary mutations (ALK\_12, 13, 28, 44, 62), which included known TKI therapy resistance variants I1171N, I1171T, F1174L, F1174V, F1174C, L1196M, G1202R, and G1269A<sup>29</sup>. The median VAF of these ALK mutations was 0.18%. During the follow-up period, four patients (ALK\_03, 13, 18, and 28) underwent subsequent treatment with a TKI that targeted the ALK mutation (Supplementary Fig. 1). In all cases, a reduction of the VAF of the mutation was observed after the change of therapy until the onset of further disease progression.

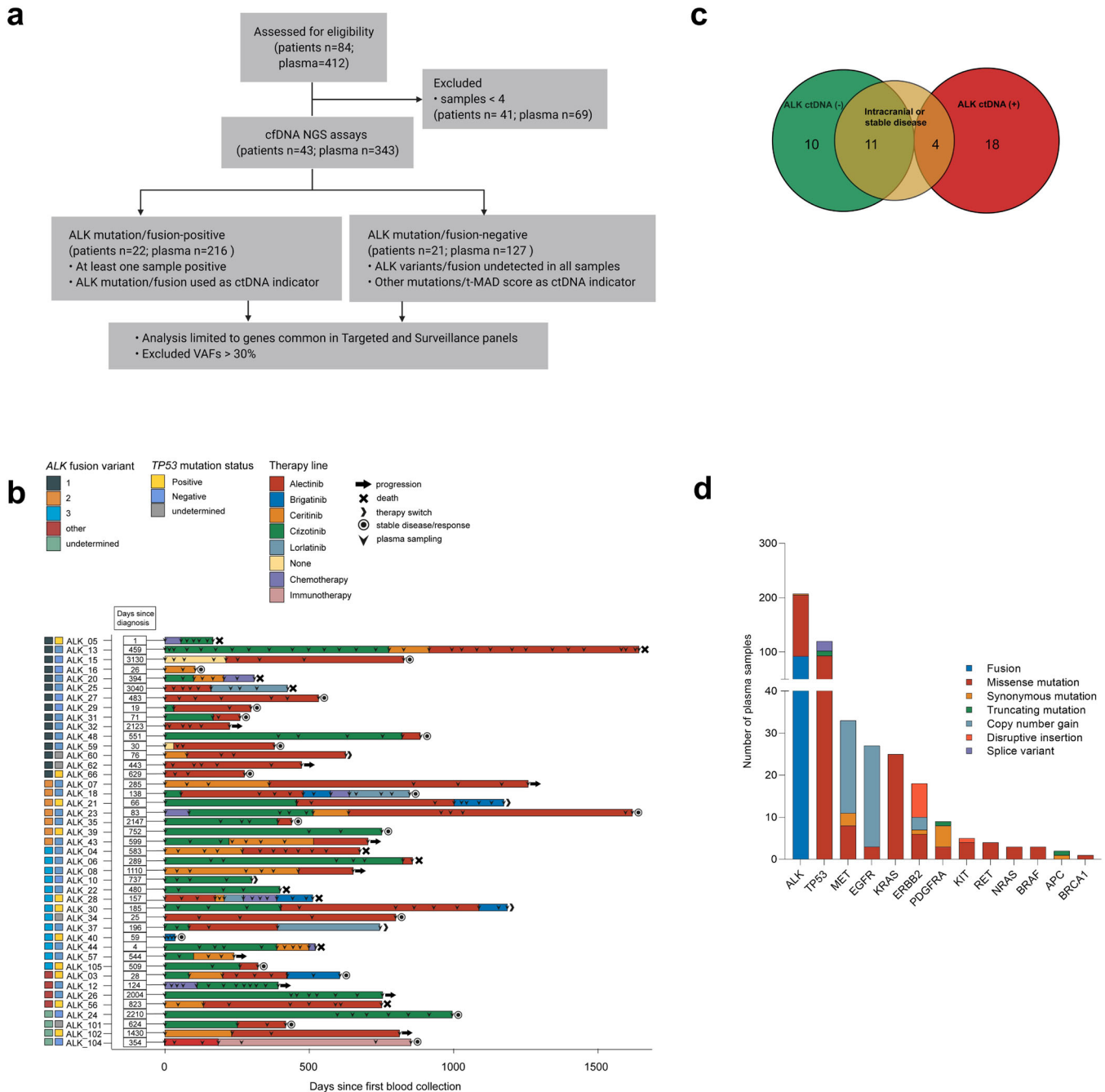
### Prognostic utility of ctDNA

We analyzed the relevance of detecting variants at treatment initiation in relation to subsequent therapy durability, indicated by the time to progression from treatment baseline. We determined that upon independent evaluation of each therapy instance, detection of VAF (i.e.,  $VAF_{\text{mean}} > 0$ ) at therapy baseline reflected significantly shorter time to progression compared to cases without detectable variants (median: 97 vs. 224 days) ( $P < 0.0495$ ; Fig. 2a). Additionally, we observed elevated ctDNA levels as measured by  $VAF_{\text{mean}}$  towards the end of life in 14 out of 16 patients who eventually succumbed to the illness. In six representative patients, the measured ctDNA levels were among the highest in samples prior to death (Fig. 2b).

### Evaluating the potential utility of changes in genome-wide copy number and $VAF_{\text{mean}}$ as indicators of disease progression

We asked whether the t-MAD score, as a measure of global copy number changes derived from ctDNA, could be an indicator of clinical progression. Since the range of absolute t-MAD scores for each patient varied due to the heterogeneity of clinical status and applied therapy lines, we utilized a relative measurement of t-MAD based on percentage of change ( $\% \Delta t\text{-MAD}$ ) to determine whether this score of genome-wide copy number aberration could indicate disease progression. We found that the relative change in t-MAD score at successive sampling points was associated with extracranial progressive disease coincident with therapy change (AUC: 0.8036 [0.7333–0.8740],  $P < 0.0001$ , Fig. 2c). The cutoff with maximum sensitivity (100%) and specificity (51.68%) was a t-MAD score increase of 0.5%.

Similarly, ctDNA levels based on tNGS were applied to assess and predict therapy response<sup>17,30–33</sup>. Here, we used the difference between the detectable  $VAF_{\text{mean}}$  ( $\Delta VAF_{\text{mean}}$ ) of consecutive sampling points to quantify the change in ctDNA abundance<sup>31</sup>. The  $\Delta VAF_{\text{mean}}$  also correlated significantly with clinical status (i.e., extracranial progressive disease coincident with therapy change), but with a lower AUC than  $\% \Delta t\text{-MAD}$  (AUC: 0.6499 [0.5277–0.7720],  $P = 0.0080$ , Fig. 3b). The optimal  $\Delta VAF_{\text{mean}}$  cutoff was 0.0850, which showed a sensitivity of 57.58% and specificity of 75.97% (Fig. 2c).



**Fig. 1 Overview of the ALK+ NSCLC cohort.** **a** Untargeted sWGS and targeted NGS assays were used to assess ctDNA in 43 ALK+ NSCLC patients corresponding to 343 longitudinal plasma samples. Targeted NGS analysis only included genes common to both the Avenio Targeted and Surveillance panels. In total, 51% (22 of 43) of patients had an ALK alteration in at least one liquid biopsy. **b** Timeline of plasma collection and therapy administration in the patient cohort. ALK fusion variants and TP53 status from tissue analysis at diagnosis are indicated by colored boxes. **c** Distribution of patients based on detected ALK alterations—including mutations and rearrangements. The group without detectable ALK alterations (ALK-) is enriched for patients with intracranial progression or stable disease. **d** Molecular alterations identified in 174 ctDNA (+) plasma samples based on tNGS.

**Molecular progression precedes clinical progression in a subset of ctDNA (+) patients**

Next, we investigated whether ctDNA detection, as indicated by molecular progression, could predate and thus predict eventual clinical progression. Here, we used the  $\Delta VAF_{mean}$  and  $\% \Delta t-MAD$  cutoff values defined from the ROC curve analysis of progressive disease points with therapy change as described above (i.e.,  $\Delta VAF_{mean} = 0.0850$ ,  $\% \Delta t-MAD = 0.500\%$ ). Due to the

heterogeneous therapy regimens across patients, we evaluated the time difference in days between early molecular and clinical progression (i.e., lead times) independently for each therapy line per patient. To identify early molecular progression, we examined sampling time points under treatment with stable disease according to RECIST v1.1. We identified 30 and 49 instances of such molecular progression using  $\Delta VAF_{mean}$  and  $\% \Delta t-MAD$ , respectively, with 14 instances meeting both

**Table 1.** Patient characteristics.

ALK+ NSCLC patients analyzed in this study ( <i>n</i> = 43)	
Age, median (39–80)	57 (10)
Sex, % male	53%
Smoking status (% never smokers) <sup>a</sup>	78%
ECOG PS (%) at baseline	
0	26
1	14
2	1
No data	2
Histology <sup>b</sup>	
adenocarcinoma	42/43
ALK fusion variant <sup>c</sup>	
<i>EML4-ALK</i> V3	13
<i>EML4-ALK</i> V1/V2	22
Other	4
No data	4
<i>TP53</i> status at baseline, mutated <sup>d</sup>	11/39
ALK TKI, patient number	
Crizotinib	25
Ceritinib/alectinib/brigatinib	36
Lorlatinib	4
Chemotherapy	7
Follow-up in months (median, [Q3–Q1])	37 (52–27)
Number of samples analyzed per patient (mean [range])	8 (4–31)
Percentage of cases with treatment-naïve samples	21%
Number of samples at disease progression per patient, mean	2.8
Number of TKI lines covered with liquid biopsy per patient, mean	1.9

*SD* standard deviation, *EML4-ALK* echinoderm microtubule-associated protein-like 4 (*EML4*) and anaplastic lymphoma kinase (*ALK*) fusion, *PS* performance status, *TKI* tyrosine kinase inhibitor.

<sup>a</sup>Data available for 41/43 cases.

<sup>b</sup>One patient had an ALK+ large-cell neuroendocrine lung carcinoma responsive to ALK inhibitors.

<sup>c</sup>Data available for 39/43 cases; one case with E18A20, one with E9A20, one with K9A20 (*KCL1*), and one with K24A20 (*KIF5B*).

<sup>d</sup>Data available for 39 cases by NGS of tissue biopsies at diagnosis of stage IV disease.

thresholds (Fig. 3a). Subsequently, we assigned a lead time if the initial increase in ctDNA was sustained at subsequent sampling points and eventually led to a progressive disease state. This second criterion validated 6 out of the 16 (37.5%) points based on  $\Delta\text{VAF}_{\text{mean}}$  alone and 12 out of 35 (34.2%) based on  $\% \Delta t\text{-MAD}$  alone as true lead times. In the 14 cases where both NGS assay cutoffs were surpassed, 10 (71.4%) points were validated. These values indicated the higher sensitivity of  $\% \Delta t\text{-MAD}$  compared to  $\Delta\text{VAF}_{\text{mean}}$ , while both showed similar specificity. Moreover, these results emphasized the potential of the combined NGS assays in increasing prediction confidence of early molecular progression. In total, we observed lead times in 19 out of 43 patients (44%), corresponding to 21 instances during TKI treatment, and one instance during chemotherapy. The median lead time to clinical progression was 88 days (range: 20–638 days). Fourteen of these 19 patients belonged to the group with detectable *ALK* alterations by tNGS (*n* = 22) as described above (Fig. 3b),

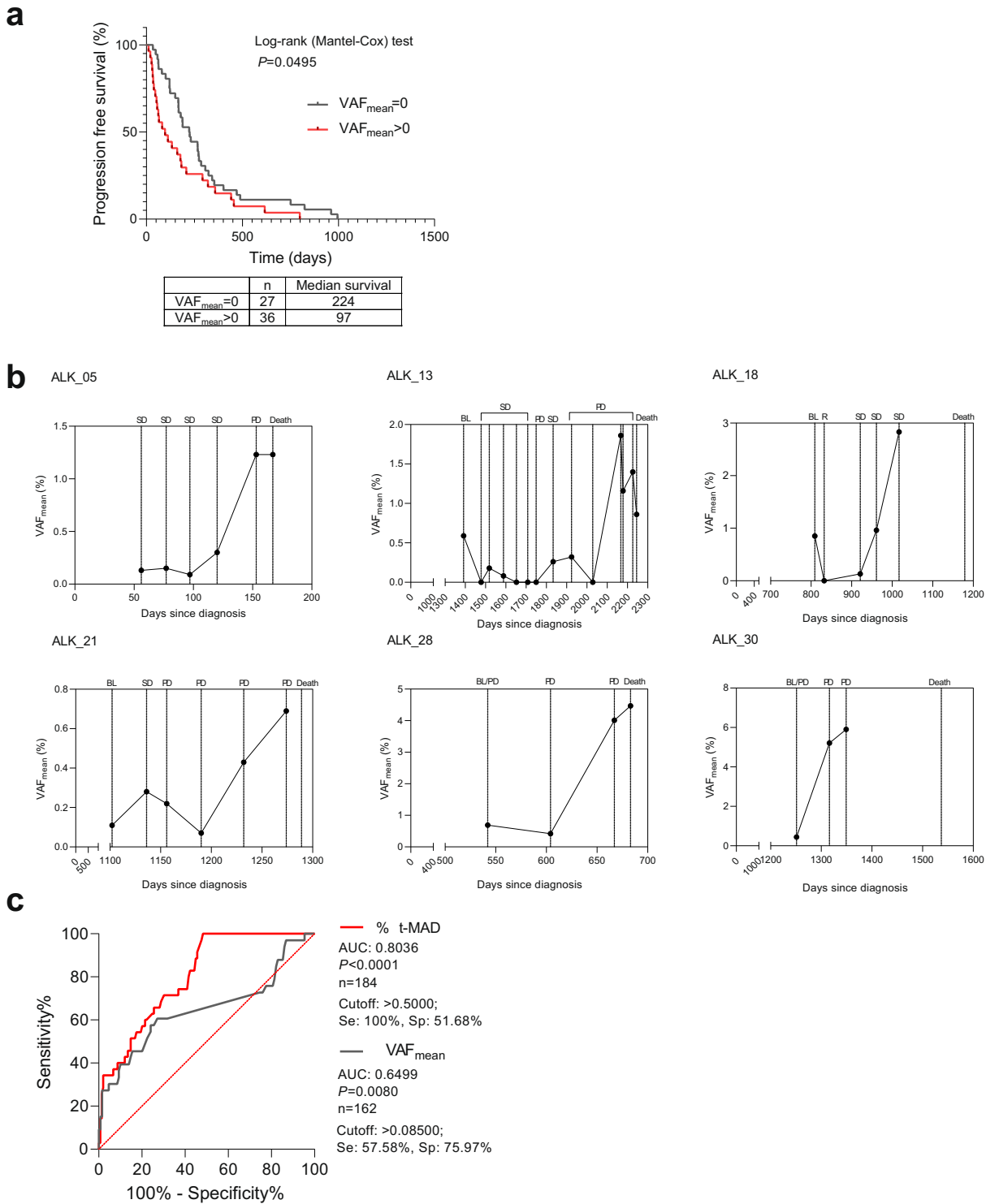
indicating that calling lead times would be most relevant in patients or samples with high ctDNA content. Through this analysis, we also demonstrated the agreement and complementarity of tNGS (indicated by  $\Delta\text{VAF}_{\text{mean}}$ ) and sWGS (indicated by  $\% \Delta t\text{-MAD}$ ) data in indicating molecular progression prior to standard radiographic assessment. Combined tNGS and sWGS data predicted lead times in 7 cases out of 22 (32%), tNGS data alone predicted lead times in five (23%) cases, and sWGS data alone was predictive in ten (45%) cases (Fig. 3c). No significant differences were found for the length of lead time predicted by either NGS assay, both independently and in combination (Fig. 3d). Lead time determination was also independent of therapy line (Fig. 3e). Time to molecular progression was significantly shorter in cases with lead time compared with clinical progression based on radiographic assessment (Fig. 3f). Lead times were longer for patients and/or therapy lines with longer duration of response (Fig. 3g). Out of 12 early molecular progression cases predicted by tNGS, 8 included *TP53* mutations (*TP53<sup>mut</sup>*) in the variants monitored longitudinally. Upon lead time comparison of *TP53<sup>wt</sup>* versus *TP53<sup>mut</sup>*, we observed significantly shorter lead times in cases with *TP53* mutation (Fig. 3h).

### Relevance of *ALK* resistance mutations in early molecular progression events

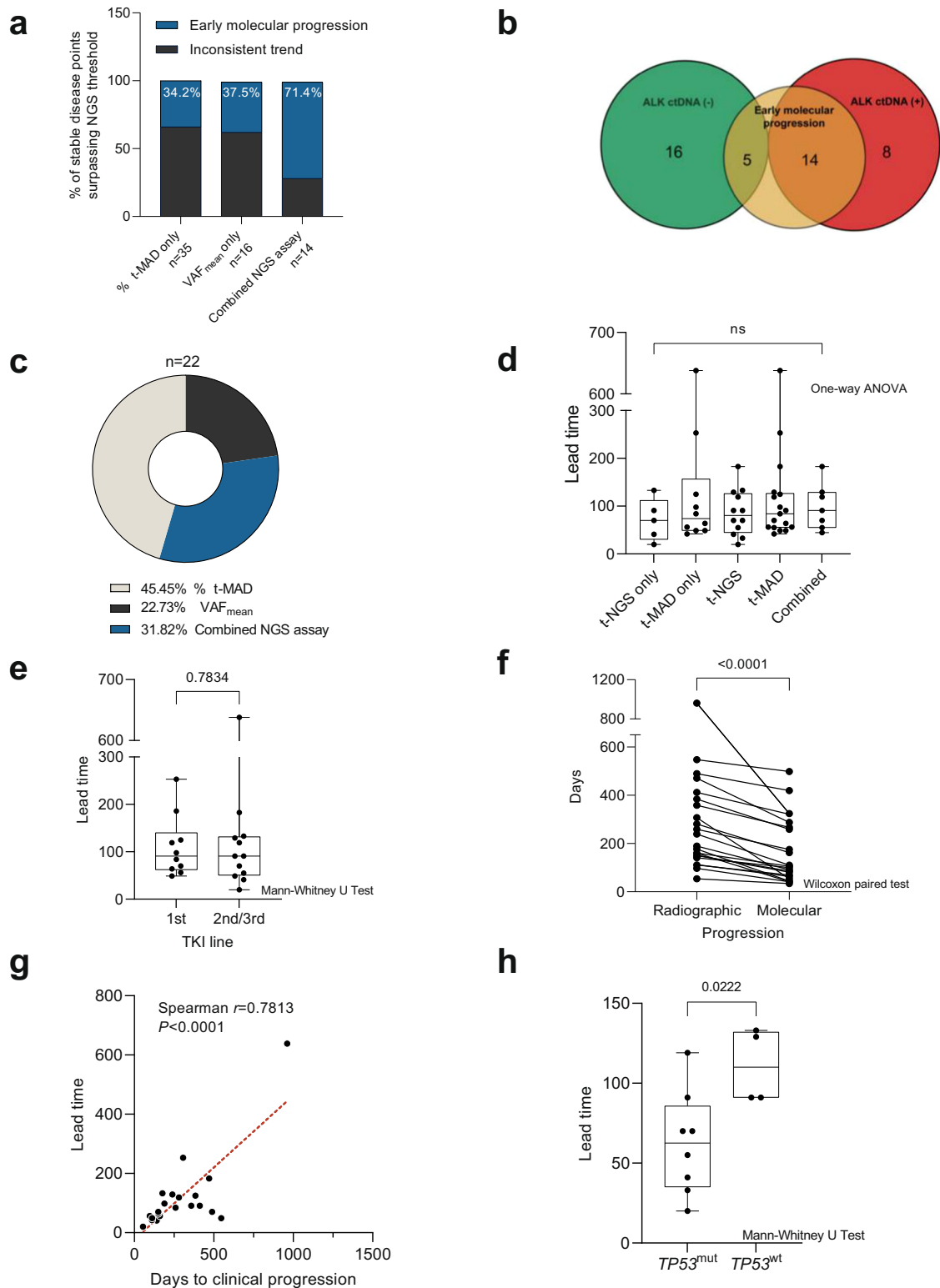
We observed that early molecular progression identified through increased  $\Delta\text{VAF}_{\text{mean}}$  were, in most cases, due to the cumulative contribution of known or probable *ALK* resistance mutations. Out of a total of 36 *ALK* mutations identified in our study, 20 overlapped with called lead times due to consistent and threshold-surpassing  $\Delta\text{VAF}_{\text{mean}}$ . From the remaining 16 mutations, 9 were detected at the time of radiologic progression (i.e., were not informative of lead time). The remaining seven instances represented mutations detected at the first sample analyzed for each patient (*n* = 3), which are likely clonal mutations present prior to therapy administration, and mutations arising under treatment (*n* = 4) whose associated  $\Delta\text{VAF}_{\text{mean}}$  did not reach the threshold. In addition, the VAFs of these mutations declined before the next instance of radiologic progression. For emergent non-*ALK* mutations (*n* = 23), which included alterations in *BRAF*, *ERBB2*, *KRAS*, *MET*, *NRAS*, *RET*, and *TP53*, 13 mutation events were detected at the time of radiologic progression. From the remaining 10 mutations, two contributed to early molecular progression detection (i.e., their emergence contributed to  $\Delta\text{VAF}_{\text{mean}}$  which exceeded the defined threshold), and 8 mutations emerged under treatment whose  $\Delta\text{VAF}_{\text{mean}}$  did not reach the threshold, and subsequently the VAFs of these mutations declined before the next instance of radiologic progression. Altogether, our data indicate that resistance mutations, particularly known *ALK* resistance mutations, can indicate early molecular progression in many cases. However, the detection of lead time is more reliable in instances where such novel mutations arise in conjunction with increased  $\Delta\text{VAF}_{\text{mean}}$ , as applied in this work. Otherwise, detection of isolated and low allele frequency mutations pose the risk of identifying false-positive pseudo-progression events<sup>34</sup>.

### Longitudinal monitoring in representative ALK+ NSCLC patients

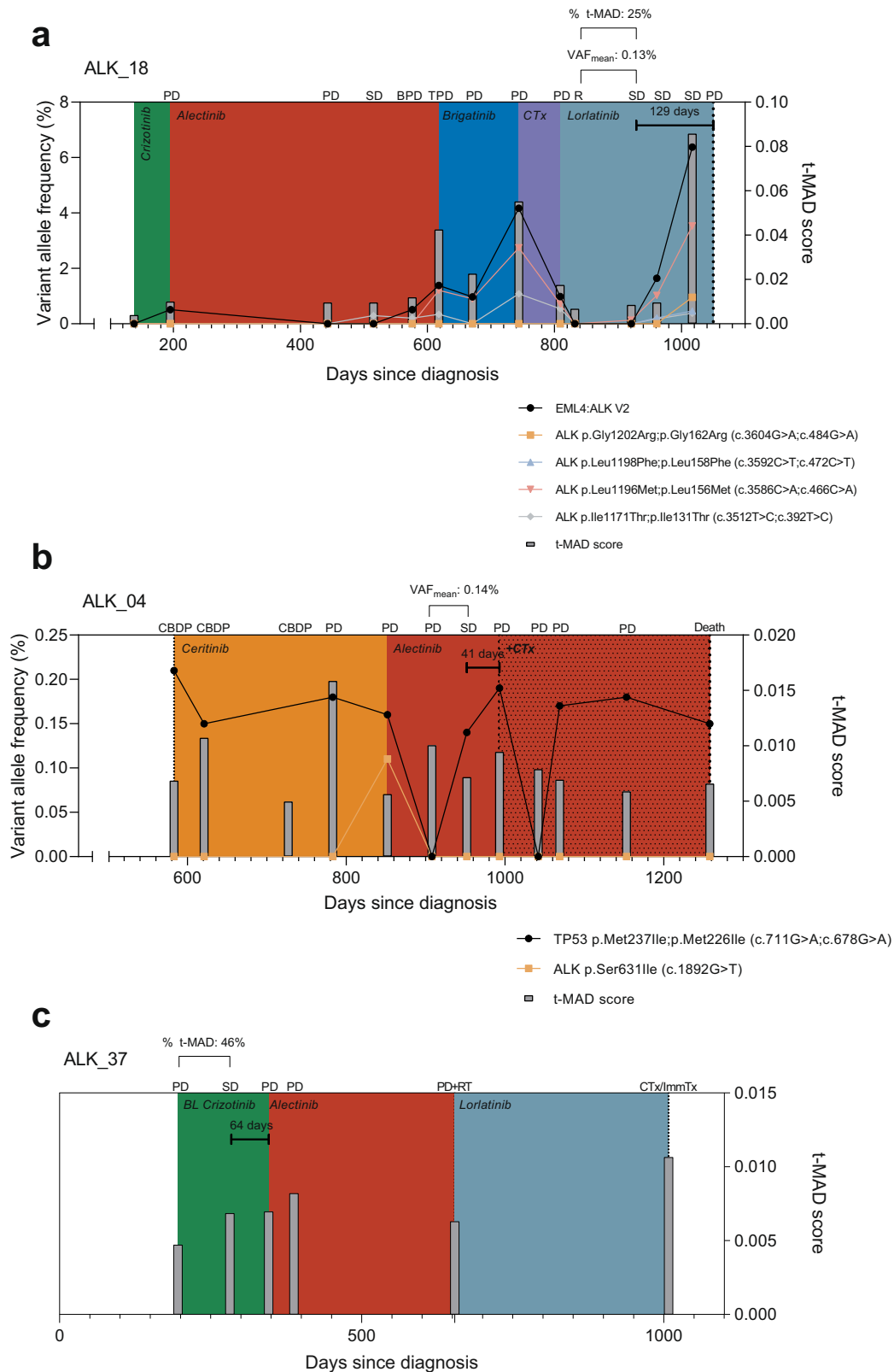
Representative cases of early molecular progression are illustrated in Fig. 4. Patient ALK\_18 (Fig. 4a) depicts a case wherein both  $\Delta\text{VAF}_{\text{mean}}$  and  $\% \Delta t\text{-MAD}$  were in agreement in detecting early molecular progression at a stable disease point. During lorlatinib therapy, ALK\_18 initially exhibited response as reflected by decreasing VAF and t-MAD score 65 days post initiation of treatment. Early molecular progression was detected 112 days post therapy initiation while the disease was evaluated as stable,



**Fig. 2 Prognostic utility of ctDNA NGS assays. a** Mean variant allele frequency at baseline indicates therapy durability. Progression-free survival plot showing significantly longer therapy durability in instances when ctDNA is undetected ( $VAF_{mean} = 0$ ) compared to detection ( $VAF_{mean} > 0$ ) at treatment initiation. **b** VAF kinetics in representative patients showing elevated  $VAF_{mean}$  towards end of life. SD stable disease, PD progressive disease, BL therapy baseline, R response. **c** Comparison of receiver operating characteristic curves (ROC) using  $\Delta VAF_{mean}$  (gray) or  $\% \Delta t-MAD$  (red) in discriminating progressive and non-progressive disease states.



**Fig. 3 Evaluation of the utility of  $\Delta$ VAF<sub>mean</sub> and % $\Delta$ t-MAD in identifying early molecular progression.** **a** Number of stable disease points that surpassed the numerical cutoffs for individual NGS assay (% $\Delta$ t-MAD or  $\Delta$ VAF<sub>mean</sub>) and in combination. Gray stacks indicate the percentage of points with inconclusive NGS assay values leading to clinical progression. Blue stacks represent points that showed sustained % $\Delta$ t-MAD and/or  $\Delta$ VAF<sub>mean</sub> increase leading to clinical progression, indicating early molecular progression. **b** Patients identified with early molecular progression were enriched in the ALK ctDNA (+) group. **c** Breakdown of 22 patients with early molecular progression based on NGS assays. **d** Comparison of called lead times based on NGS assays. **e** Comparison of called lead times based on therapy lines. **f** Comparison of cases with called lead time showing significant difference in days to radiographic progression versus molecular progression. **g** Length of lead time per therapy line was significantly correlated with duration of progression-free response to the respective treatment. **h** Lead times in cases with wild-type *TP53* (*TP53*<sup>wt</sup>) were significantly longer compared to cases with mutated *TP53* (*TP53*<sup>mut</sup>) detected by tNGS. The box plot show the the median, upper quartile, and lower quartile values. The whiskers indicate minimum and maximum values.



**Fig. 4 Representative cases illustrating early molecular progression as indicated by elevated  $\Delta\text{VAF}_{\text{mean}}$  and  $\% \Delta\text{t-MAD}$ .** **a** Patient ALK\_18 presented a case—during lorlatinib therapy—where both  $\Delta\text{VAF}_{\text{mean}}$  and  $\% \Delta\text{t-MAD}$  were informative of the lead time. Early molecular progression was apparent 129 days prior to clinical progression, as indicated by  $\Delta\text{VAF}_{\text{mean}}$  of 0.13% and  $\% \Delta\text{t-MAD}$  of 25%. Both values surpass the NGS metric thresholds identified in this study. **b** Patient ALK\_04 depicts a case wherein  $\Delta\text{VAF}_{\text{mean}}$  indicated early molecular progression during alectinib therapy. A  $\Delta\text{VAF}_{\text{mean}}$  of 0.14% was calculated 41 days prior to clinical progression. **c** Patient ALK\_37 emphasizes the relevance of untargeted NGS in ctDNA monitoring particularly in cases where genetic alterations remained undetected by tNGS. Here,  $\% \Delta\text{t-MAD}$  of 46% was already apparent at a stable disease time point, 64 days prior to clinical progression. PD progressive disease, SD stable disease, BPD brain PD, TPD thoracic PD, R response, CDBP treatment continuation beyond disease progression, RT radiotherapy, CTx chemotherapy, ImmTx immunotherapy.

and two subsequent sampling points showed continued increase of NGS metrics (i.e.,  $\Delta\text{VAF}_{\text{mean}} = 0.13\%$ ;  $\% \Delta\text{t-MAD} = 25\%$ ) with stable disease until disease progression was observed. This point came 129 days after the onset of molecular progression. Patient ALK\_04 (Fig. 4b) represents a case where  $\Delta\text{VAF}_{\text{mean}}$  was instrumental in identifying a lead time during alectinib therapy. After disease progression on ceritinib (852 days since diagnosis), the therapy was switched to alectinib in which stable disease was observed 100 days post administration. However, molecular analysis at this point already revealed  $\Delta\text{VAF}_{\text{mean}}$  of 0.14%, surpassing the threshold. Clinical progression was eventually observed 41 days later, supporting the result of ctDNA evaluation. Finally, patient ALK\_37 (Fig. 4c) shows the utility of t-MAD monitoring particularly in the absence of detectable genetic alterations by tNGS. During crizotinib therapy, a  $\% \Delta\text{t-MAD}$  increase of 46% was already apparent 86 days post treatment initiation, at which point the disease was still clinically stable. Disease progression occurred 64 days later, which led to a TKI therapy switch to alectinib. These case examples represent the 19 instances in our cohort where disease progression lead times were indicated by elevated levels of ctDNA through quantification of  $\Delta\text{VAF}_{\text{mean}}$  and  $\% \Delta\text{t-MAD}$ .

## DISCUSSION

Due to the relatively long duration of treatment and emergence of actionable acquired resistance, therapy monitoring is especially important for TKI-treated NSCLC patients<sup>35,36</sup>. Early detection of disease progression could facilitate significant improvement in patient outcomes, as approximately 25–30% of NSCLC patients failing ALK or EGFR TKI miss subsequent therapy or chemotherapy due to rapid clinical deterioration<sup>37,38</sup>. Previously, shallow whole-genome sequencing has been applied to estimate global copy number changes from ctDNA through computation of the t-MAD score<sup>28,39</sup>. We further showed that this metric was associated with shorter overall survival in a subset of our ALK+ NSCLC cohort<sup>27</sup>. Here, we systematically examine whether combined tNGS and sWGS could also enable the detection of acquired resistance and disease progression earlier than radiological imaging.

Previously, an important challenge was the lack of a general consensus on how to measure ctDNA levels and response cutoffs. To date, multiple studies have used various ctDNA quantification methods including the maximum variant allele frequency<sup>40</sup>, the mean variant allele frequency<sup>31,41</sup>, and the total ctDNA concentration<sup>32</sup>. Similarly, response thresholds were typically established using a single time point after therapy baseline with arbitrary cutoff values<sup>18,32,42</sup>. Here, we employed  $\Delta\text{VAF}_{\text{mean}}$  and the  $\% \Delta\text{t-MAD}$  scores because these metrics can capture ctDNA changes regardless of how many SNV and/or CNV are present in the baseline sample. The use of ultra-deep targeted sequencing with optimized gene panels enables efficient surveillance of alteration hotspots of high clinical impact. These somatic changes are then used for early screening, treatment monitoring, and detection of residual disease. In parallel, untargeted sWGS approaches to liquid biopsy have been successful in other cancer entities<sup>39,43,44</sup> and have been shown to reflect tumor status, even in cases without detectable SNVs<sup>27</sup>. Our data show that  $\Delta\text{VAF}_{\text{mean}}$  offers better specificity than  $\% \Delta\text{t-MAD}$  in discriminating progressive versus non-progressive disease. In contrast,  $\% \Delta\text{t-MAD}$  showed increased sensitivity. These results suggest that  $\Delta\text{VAF}_{\text{mean}}$  could be a more relevant ctDNA metric for clinical utility in the advanced setting where specificity is a critical performance assay criterion and greater ctDNA shedding occurs<sup>44</sup>, whereas  $\% \Delta\text{t-MAD}$  could be applied at early disease stages. Nonetheless, residual false-positive and false-negative calls of disease status based on ctDNA evaluations could still take place due to tumors that shed minimal amounts of ctDNA; due to cases with apparent disagreement between molecular progression and imaging evaluations; and due

to clearance of tumors leading to ctDNA shedding during earliest cycles of treatment.

In our study population,  $\text{VAF}_{\text{mean}} > 0$  at therapy baseline predicted shorter therapy duration compared to cases without detected variants ( $\text{VAF}_{\text{mean}} = 0$ ). This parallels findings demonstrating that ctDNA at baseline indicates poor prognosis in various cancers<sup>18,27,45–47</sup>. Pretreatment VAF levels have also been shown to be prognostic of treatment response with immune checkpoint inhibitors in patients with advanced solid tumors<sup>31,48</sup>. Pretreatment  $\text{VAF}_{\text{mean}}$  and maximum VAF were significantly inversely correlated with overall survival in NSCLC, urothelial cancer, microsatellite instable cancers, gastroesophageal cancer, and ovarian cancer<sup>31</sup>. Our results extend the potential of ctDNA detectability at therapy baseline as biomarkers for therapy response across various treatment lines in ALK+ NSCLC. Nonetheless, we recognize the marginal significance ( $P < 0.045$ ) of our survival data. Variations in therapy sequence, therapy duration, and interpatient variabilities could be sources of confounding factors in the analysis. Each therapy line per patient was also treated as an independent case, discounting the potential effects of prior therapy.

A potential confounding factor in longitudinal monitoring lies in the assumption that visual evaluation of tumor radiographs accurately reflects tumor growth or progression<sup>49</sup>. In fact, there were sampling points attributed with clinically stable disease that showed consistently increasing levels of  $\text{VAF}_{\text{mean}}$  and/or t-MAD, which could indicate onset of disease progression below the limit of detection of radiographic modalities. By specifying numerical cutoffs for both NGS assays through ROC analysis, we were able to systematically identify such early molecular progression events in 44% of patients, across all therapy lines. This suggests that in this subset of our study cohort, ctDNA analysis exhibits higher sensitivity in detecting tumor progression than visual staging. It is also worthwhile to note that in majority of cases, early emergence of known or probable ALK mutations accounted for increased  $\Delta\text{VAF}_{\text{mean}}$  that resulted in lead time identification. This emphasizes the sensitivity of ctDNA analysis in identifying potentially actionable mutations early in the disease course. From the tumor biology standpoint, such cases reflect the dynamic clonal evolution within the tumor that could potentially be missed by tissue re-biopsies. Nonetheless, there were also mutations—both ALK and non-ALK—that were inconsistent or sporadic throughout the longitudinal course. Using  $\Delta\text{VAF}_{\text{mean}}$  as a cumulative metric of ctDNA levels minimizes identification of such false-positive pseudo-molecular progression events that could have arisen from stochastic sampling issues<sup>50</sup>.

While both tNGS and sWGS were able to independently call early molecular progression events, we also present evidence that a combined evaluation increases the confidence in predicting lead times by approximately 50%, corroborating our previous observation<sup>27</sup> that untargeted sWGS complements tNGS. Moreover, there was a positive correlation between the length of predicted lead time and durability of response to therapy. Therefore, molecular monitoring for earlier detection of disease progression could be even more useful for patients treated with newer ALK inhibitors, namely alectinib, brigatinib, or lorlatinib, whose progression-free survival extends well beyond two years, in contrast to 10–12 months under crizotinib<sup>8–10</sup>. This approach also limits the risk due to higher cumulative radiation exposure from regular CTs<sup>51</sup>. Interestingly, we further observed that lead times involving mutations in *TP53* were shorter compared to *TP53* wild types. This implies that the emergence of *TP53* alterations during therapy course portends a swifter progression of disease, as previously noted in ALK+ NSCLC<sup>52–54</sup>. Ideally, early detection of molecular progression could prompt clinicians to halt ineffective treatments, avoiding potential side effects and added financial burden, and also facilitate an earlier therapy switch, while the patients are still capable of receiving next-line therapies<sup>38</sup>. Such a strategy could



be particularly useful for cases with higher risk, such as those with *TP53* mutations and/or the *EML4-ALK* variant 3 (refs. 53,55). This approach is complementary to monitoring of early ctDNA changes after therapy initiation, which have been shown to predict therapy durability<sup>44</sup>.

The main limitations of our study are related to its retrospective design, the relatively small patient number, and the lack of a separate validation cohort. Our findings should ideally be confirmed in a larger, prospective study with strictly defined sampling and imaging intervals. The variants by tNGS were also limited to the genes included in the panels utilized, thus other genetic alterations in ctDNA were not evaluated.

Our work demonstrates the applicability of a combined targeted and untargeted NGS analysis of cfDNA for longitudinal monitoring and identification of early molecular progression in ALK+ NSCLC. We established  $\Delta\text{VAF}_{\text{mean}}$  and  $\% \Delta\text{t-MAD}$  as quantifiable metrics for ctDNA levels and determined thresholds to enable prediction of disease progression and identification of early molecular progression. Currently, the development of robust criteria assessing ctDNA dynamics that correlate with clinical status such as disease progression, therapy response, and therapy resistance remains challenging partly due to stochastic sampling issues<sup>50</sup>, and the spectrum of ctDNA shedding tendencies of different malignancies. Aside from ctDNA, it should also be noted that other cell-free markers have the capacity to detect progression and predict disease recurrence. In NSCLC, for example, miRNA signatures in the plasma have been shown to associate with higher risk for progression<sup>56,57</sup>, while a separate panel could predict survival in squamous cell carcinoma patients<sup>56</sup>. Serum carcinoembryonic antigen (CEA) is also associated with distinct cancer progression profiles and carries predictive information of risk recurrence in NSCLC<sup>58,59</sup>. Oncogenic mRNA markers in plasma have also shown clinical utility in NSCLC<sup>60,61</sup>. Thus, in addition to genomic alterations in ctDNA, future investigations of other plasma solutes including proteins, miRNAs, mRNAs, and the epigenomic properties of ctDNA (e.g., methylation, nucleosome positioning) will advance the potential of liquid biopsy in complementing imaging technologies in improving personalized patient management and efficient therapy decisions.

## METHODS

### Patients and sample collection

The study was approved by the ethics committees of Heidelberg University (S-270/2001, S-296/2016) and Lübeck University (AZ 12-238). Written informed consent was obtained from all study participants. Newly diagnosed cases were screened for the presence of an *ALK* alteration in tissues by fluorescence in situ hybridization (FISH, ZytoLight SPEC *ALK* probe, ZytoVision GmbH, Bremerhaven, Germany) and reverse-transcription polymerase chain reaction until 2015, or by immunohistochemistry (D5F3 clone, Roche, Mannheim, Germany) and RNA-based next-generation sequencing (NGS, Thermo Fisher Lung Cancer Fusion Panel, Waltham, MA, USA) thereafter, as previously described<sup>12</sup>. Peripheral blood was prospectively drawn through venipuncture at each outpatient visit from 84 ALK+ patients with metastatic NSCLC under TKI therapy at Thoraxklinik Heidelberg and Lungenclinik Grosshansdorf, Germany<sup>62</sup>. In total, 412 longitudinal blood samples were collected and plasma was isolated within 1 h of blood draw, using the double spin method as previously described<sup>27</sup>. Samples were stored at  $-80^{\circ}\text{C}$  in the Lung Biobank Heidelberg/BMBH until further processing.

Since our study was focused on longitudinal therapy monitoring, we limited the scope of our analysis only on patients with at least four evaluable plasma samples throughout their treatment course. Using these criteria, we excluded 41 patients (69 plasma samples). Ultimately, 43 metastatic ALK+ patients with corresponding 343 plasma samples were considered for NGS analysis (Fig. 1a). The characteristics of these patients are summarized in Table 1. Clinical data and the results of routine radiographic assessments (every 8–12 weeks) using chest/abdominal CT and brain MRI were collected through a review of patient records with a cutoff on September 15, 2020.

### cfDNA isolation

The AVENIO cfDNA Isolation Kit (Roche Diagnostics) was used to isolate cfDNA from 2 mL of patient plasma, following the manufacturer's protocol. The characteristic mononucleosomal molecular weight profile (160–200 bp) of the purified cfDNA was assessed using the Bioanalyzer 2100 High Sensitivity DNA Kit (Agilent Technologies). The cfDNA was quantified using the Qubit dsDNA High Sensitivity Kit (Thermo Fisher).

### cfDNA NGS assays

Sequencing libraries for capture-based targeted sequencing were generated using the AVENIO ctDNA Library Preparation Kit with either the Targeted or Surveillance Panel (Roche Diagnostics), as previously described<sup>27</sup>. Equimolar 16-plex enriched library pools were sequenced on the Illumina NextSeq 550 platform with the High-Output Kit V2 ( $2 \times 150$  bp). The resulting raw BCL files were processed using the AVENIO ctDNA analysis software (Roche Diagnostics, version 2.0.0). The proprietary analysis pipeline applied by the software is adapted from the CAPP-Seq workflow with integrated digital suppression<sup>63,64</sup>. Called variants with variant allele frequencies (VAFs)  $\geq 30\%$  were classified as germline mutations and were excluded from subsequent analyses. Only genes common to both Targeted and Surveillance panels (Supplementary Table 3) were compared for all longitudinal assessments, with a VAF calling threshold of 0.01%, facilitated by an average of 4100 $\times$  unique target coverage as previously reported<sup>27</sup>. For SWGS, sequencing libraries were prepared using the KAPA HyperPrep Kit with KAPA Dual-Indexed Adapters for Illumina platforms, as previously described<sup>27</sup> using 1–2.5 ng cfDNA as input. Pooled multiplexes of 48–67 equimolar libraries were sequenced on the Illumina HiSeq 4000 platform ( $2 \times 100$  bp). To determine genome-wide copy number variations (CNVs) and calculate the t-MAD scores, raw SWGS sequencing data were processed and analyzed as previously described<sup>27</sup>. Briefly, automated sequence quality control and alignment were performed using the One Touch Pipeline<sup>65</sup>. Genome-wide copy number profiles and tumor fractions were estimated using ichorCNA implemented in R (version 3.3.1). T-MAD scores were computed using a 1-Mb bin size and without fragment size selection<sup>27</sup>.

### Genomic alteration and t-MAD score determination

Somatic SNVs and indels detected by the AVENIO ctDNA analysis software were used to calculate the mean variant allele frequency ( $\text{VAF}_{\text{mean}}$ ) per sampling point. This value is the sum of all VAFs divided by the number of all mutations at each time point to account for polyclonality<sup>17</sup>. The change in mean VAF ( $\Delta\text{VAF}_{\text{mean}}$ ) was the difference in  $\text{VAF}_{\text{mean}}$  between two successive time points as shown by Eqs. (1) and (2).

$$\Delta\text{VAF}_{\text{mean}} = \overline{\text{VAF}}_2 - \overline{\text{VAF}}_1 \quad (1)$$

where

$$\overline{\text{VAF}} = \frac{\sum_{i=1}^n \text{VAF}_i}{n} \quad (2)$$

Here,  $n$  represents the number of detected variants. If  $n = 0$ , or no variant was detected,  $\text{VAF}_{\text{mean}}$  was set to 0. Positive  $\Delta\text{VAF}_{\text{mean}}$  values indicate increasing  $\text{VAF}_{\text{mean}}$ . In parallel, the change in t-MAD ( $\% \Delta\text{t-MAD}$ ) was calculated, in percentage, as the difference between two successive t-MAD scores, divided by the t-MAD measured at the previous time point as shown by Eq. (3).

$$\% \Delta\text{t-MAD} = \left( \frac{\text{t-MAD}_2 - \text{t-MAD}_1}{\text{t-MAD}_1} \right) \times 100\% \quad (3)$$

$\Delta\text{VAF}_{\text{mean}}$  and  $\% \Delta\text{t-MAD}$  thresholds for disease progression were derived using the respective receiver operating characteristic (ROC) curve across measurement time points with progressive disease accompanied by therapy change ( $\text{PD}_{\text{therapy change}}$ ) as the event of interest. The annotation of  $\text{PD}_{\text{therapy change}}$  was verified manually based on radiological imaging according to RECIST criteria v1.1 (ref. 66), as well as information about patient treatment from the medical records. Area under the curve (AUC) values were generated using GraphPad Prism 9 (GraphPad software, La Jolla California USA), while cutoff values were identified using the Youden Index, which maximizes the sensitivity and specificity of the NGS metric (i.e.,  $\Delta\text{VAF}_{\text{mean}}$  and  $\% \Delta\text{t-MAD}$ )<sup>67</sup>. Seven patients (ALK\_22, 37, 57, 59, 60, 66, 105) were excluded from  $\Delta\text{VAF}_{\text{mean}}$  measurements due to the absence of detectable variants in all plasma DNA samples. In total, 162 measurements of  $\Delta\text{VAF}_{\text{mean}}$  and 184 measurements of  $\% \Delta\text{t-MAD}$  were used for ROC curve analyses (Supplementary Tables 4 and 5).

We defined early molecular progression as a change in either NGS metric (i.e.,  $\Delta VAF_{\text{mean}}$  or  $\% \Delta t\text{-MAD}$ ) surpassing the threshold value defined from the ROC curve analysis, from a point where the tumor was radiographically stable according to RECIST to a later progression point. The lead time was defined as the number of days between the first observation of molecular progression until the point of an initial clinical progression as per radiographic assessment based on RECIST.

### Statistical analysis and data visualization

Survival data were analyzed using the log-rank test. Lead time comparisons were performed using the Mann–Whitney *U* test or one-way ANOVA, as labeled in the graphs. Pairwise comparison of days to progression was performed using the Wilcoxon's paired test. Since the primary aim of this study was to establish the relationship between ctDNA changes and disease progression, samples from different progression time points of the same patient were analyzed as independent, in order to account for the fact that several factors can influence liquid biopsy positivity differently in each sample of the same patient (e.g. progression site, progression rate, tumor volume, preceding treatment), and that these factors act similarly across different patients. Statistical analyses, including ROC curves, and relevant graphs were performed and generated, respectively, using GraphPad Prism 9.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

### DATA AVAILABILITY

All sequencing data supporting the findings in this study are deposited in the European Genome-phenome Archive (EGA) under accession number EGAS00001005327.

Received: 2 June 2021; Accepted: 22 October 2021;

Published online: 07 December 2021

### REFERENCES

- Skoulidis, F. & Heymach, J. V. Co-occurring genomic alterations in non-small-cell lung cancer biology and therapy. *Nat. Rev. Cancer* **19**, 495–509 (2019).
- Solomon, B. J. et al. First-line crizotinib versus chemotherapy in ALK-positive lung cancer. *N. Engl. J. Med.* **371**, 2167–2177 (2014).
- Kim, D. W. et al. Activity and safety of ceritinib in patients with ALK-rearranged non-small-cell lung cancer (ASCEND-1): updated results from the multicentre, open-label, phase 1 trial. *Lancet Oncol.* **17**, 452–463 (2016).
- Shaw, A. T. et al. Alectinib in ALK-positive, crizotinib-resistant, non-small-cell lung cancer: a single-group, multicentre, phase 2 trial. *Lancet Oncol.* **17**, 234–242 (2016).
- Kim, D. W. et al. Brigatinib in patients with crizotinib-refractory anaplastic lymphoma kinase-positive non-small-cell lung cancer: a randomized, multicenter phase II trial. *J. Clin. Oncol.* **35**, 2490–2498 (2017).
- Chuang, J. C. & Neal, J. W. Crizotinib as first line therapy for advanced ALK-positive non-small cell lung cancers. *Transl. Lung Cancer Res.* **4**, 639–641 (2015).
- Gainor, J. F. et al. Molecular mechanisms of resistance to first- and second-generation ALK inhibitors in ALK-rearranged lung cancer. *Cancer Discov.* **6**, 1118–1133 (2016).
- Peters, S. et al. Alectinib versus crizotinib in untreated ALK-positive non-small-cell lung cancer. *N. Engl. J. Med.* **377**, 829–838 (2017).
- Camidge, D. R. et al. Brigatinib versus crizotinib in ALK-positive non-small-cell lung cancer. *N. Engl. J. Med.* **379**, 2027–2039 (2018).
- Shaw, A. T. et al. First-line lorlatinib or crizotinib in advanced ALK-positive lung cancer. *N. Engl. J. Med.* **383**, 2018–2029 (2020).
- Wang, W., Sun, X. & Hui, Z. Treatment optimization for brain metastasis from anaplastic lymphoma kinase rearrangement non-small-cell lung cancer. *Oncol. Res. Treat.* **42**, 599–606 (2019).
- Volckmar, A. L. et al. Combined targeted DNA and RNA sequencing of advanced NSCLC in routine molecular diagnostics: analysis of the first 3,000 Heidelberg cases. *Int. J. Cancer* **145**, 649–661 (2019).
- Benayed, R. et al. High yield of RNA sequencing for targetable kinase fusions in lung adenocarcinomas with no mitogenic driver alteration detected by DNA sequencing and low tumor mutation burden. *Clin. Cancer Res.* **25**, 4712–4722 (2019).
- Gerlinger, M. et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366**, 883–892 (2012).
- Ilie, M. & Hofman, P. Pros: can tissue biopsy be replaced by liquid biopsy? *Transl. Lung Cancer Res.* **5**, 420–423 (2016).
- De Rubis, G., Rajeev Krishnan, S. & Bewaby, M. Liquid biopsies in cancer diagnosis, monitoring, and prognosis. *Trends Pharmacol. Sci.* **40**, 172–186 (2019).
- Velimirovic, M. et al. Rising circulating tumor DNA as a molecular biomarker of early disease progression in metastatic breast cancer. *JCO Precis. Oncol.*, 1246–1262 (2020).
- Lee, J. H. et al. Longitudinal monitoring of ctDNA in patients with melanoma and brain metastases treated with immune checkpoint inhibitors. *Clin. Cancer Res.* **26**, 4064–4071 (2020).
- Parikh, A. R. et al. Serial ctDNA monitoring to predict response to systemic therapy in metastatic gastrointestinal cancers. *Clin. Cancer Res.* **26**, 1877–1885 (2020).
- Vandekerhove, G. et al. Plasma ctDNA is a tumor tissue surrogate and enables clinical-genomic stratification of metastatic bladder cancer. *Nat. Commun.* **12**, 184 (2021).
- Zhang, E. W. et al. Association between circulating tumor DNA burden and disease burden in patients with ALK-positive lung cancer. *Cancer* **126**, 4473–4484 (2020).
- Tie, J. et al. Circulating tumor DNA analysis detects minimal residual disease and predicts recurrence in patients with stage II colon cancer. *Sci. Transl. Med.* **8**, 346ra392 (2016).
- Garcia-Murillas, I. et al. Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. *Sci. Transl. Med.* **7**, 302ra133 (2015).
- Chaudhuri, A. A. et al. Early detection of molecular residual disease in localized lung cancer by circulating tumor DNA profiling. *Cancer Discov.* **7**, 1394–1403 (2017).
- Dagogo-Jack, I. et al. Tracking the evolution of resistance to ALK tyrosine kinase inhibitors through longitudinal analysis of circulating tumor DNA. *JCO Precis. Oncol.* 2018, <https://doi.org/10.1200/PO.17.00160> (2018).
- Noe, J. et al. ALK mutation status before and after alectinib treatment in locally advanced or metastatic ALK-positive NSCLC: pooled analysis of two prospective trials. *J. Thorac. Oncol.* **15**, 601–608 (2020).
- Dietz, S. et al. Longitudinal therapy monitoring of ALK-positive lung cancer by combined copy number and targeted mutation profiling of cell-free DNA. *EBio-Medicine* **62**, 103103 (2020).
- Mouliere, F. et al. Detection of cell-free DNA fragmentation and copy number alterations in cerebrospinal fluid from glioma patients. *EMBO Mol. Med.* **10**, <https://doi.org/10.15252/emmm.201809323> (2018).
- Lin, J. J., Riely, G. J. & Shaw, A. T. Targeting ALK: precision medicine takes on drug resistance. *Cancer Discov.* **7**, 137–155 (2017).
- Jia, N. et al. Serial monitoring of circulating tumor DNA in patients with metastatic colorectal cancer to predict the therapeutic response. *Front. Genet.* **10**, 470 (2019).
- Zhang, Q. et al. Prognostic and predictive impact of circulating tumor DNA in patients with advanced cancers treated with immune checkpoint blockade. *Cancer Discov.* **10**, 1842–1853 (2020).
- Nabet, B. Y. et al. Noninvasive early identification of therapeutic benefit from immune checkpoint inhibition. *Cell* **183**, 363–376 e313 (2020).
- Moding, E. J. et al. Circulating tumor DNA dynamics predict benefit from consolidation immunotherapy in locally advanced non-small-cell lung cancer. *Nat. Cancer* **1**, 176–183 (2020).
- Ricciuti, B. et al. Early plasma circulating tumor DNA (ctDNA) changes predict response to first-line pembrolizumab-based therapy in non-small cell lung cancer (NSCLC). *J. Immunother. Cancer* **9**, <https://doi.org/10.1136/jitc-2020-001504> (2021).
- Schrank, Z. et al. Current molecular-targeted therapies in NSCLC and their mechanism of resistance. *Cancers (Basel)* **10**, <https://doi.org/10.3390/cancers10070224> (2018).
- Gristina, V. et al. The emerging therapeutic landscape of ALK inhibitors in non-small cell lung cancer. *Pharmaceuticals (Basel)* **13**, <https://doi.org/10.3390/ph13120474> (2020).
- Magios, N. et al. Real-world implementation of sequential targeted therapies for EGFR-mutated lung cancer. *Ther. Adv. Med. Oncol.* **13**, 1758835921996509, <https://doi.org/10.1177/1758835921996509> (2021).
- Elsayed, M. et al. Feasibility and challenges for sequential treatments in ALK-rearranged non-small-cell lung cancer. *Front. Oncol.* **11**, <https://doi.org/10.3389/fonc.2021.670483> (2021).
- Smith, C. G. et al. Comprehensive characterization of cell-free tumor DNA in plasma and urine of patients with renal tumors. *Genome Med.* **12**, 23 (2020).

40. Goldberg, S. B. et al. Early assessment of lung cancer immunotherapy response via circulating tumor DNA. *Clin. Cancer Res.* **24**, 1872–1880 (2018).
41. Ococks, E. et al. Longitudinal tracking of 97 esophageal adenocarcinomas using liquid biopsy sampling. *Ann. Oncol.*, <https://doi.org/10.1016/j.annonc.2020.12.010> (2021).
42. Bratman, S. V. et al. Personalized circulating tumor DNA analysis as a predictive biomarker in solid tumor patients treated with pembrolizumab. *Nat. Cancer* **1**, 873–881 (2020).
43. Mouliere, F. et al. Enhanced detection of circulating tumor DNA by fragment size analysis. *Sci. Transl. Med.* **10**, <https://doi.org/10.1126/scitranslmed.aat4921> (2018).
44. Davis, A. A. et al. Early assessment of molecular progression and response by whole-genome circulating tumor DNA in advanced solid tumors. *Mol. Cancer Ther.* **19**, 1486–1496 (2020).
45. Jacob, S. et al. The use of serial circulating tumor DNA to detect resistance alterations in progressive metastatic breast cancer. *Clin. Cancer Res.*, <https://doi.org/10.1158/1078-0432.CCR-20-1566> (2020).
46. Cavallone, L. et al. Prognostic and predictive value of circulating tumor DNA during neoadjuvant chemotherapy for triple negative breast cancer. *Sci. Rep.* **10**, 14704 (2020).
47. Christopoulos, P. et al. Earlier extracranial progression and shorter survival in ALK-rearranged lung cancer with positive liquid rebiopsies. *Transl Lung Cancer Res* **10**, 2118–2131 (2021).
48. Marsavela, G. et al. Circulating tumor DNA predicts outcome from first-, but not second-line treatment and identifies melanoma patients who may benefit from combination immunotherapy. *Clin. Cancer Res.* **26**, 5926–5933 (2020).
49. Keup, C. et al. Longitudinal multi-parametric liquid biopsy approach identifies unique features of circulating tumor cell, extracellular vesicle, and cell-free DNA characterization for disease monitoring in metastatic breast cancer patients. *Cells* **10**, <https://doi.org/10.3390/cells10020212> (2021).
50. Hrebien, S. et al. Early ctDNA dynamics as a surrogate for progression-free survival in advanced breast cancer in the BEECH trial. *Ann. Oncol.* **30**, 945–952 (2019).
51. Volckmar, A. L. et al. A field guide for cancer diagnostics using cell-free DNA: from principles to practice and clinical applications. *Gene Chromosome Cancer* **57**, 123–139 (2018).
52. VanderLaan, P. A. et al. Mutations in TP53, PIK3CA, PTEN and other genes in EGFR mutated. *Lung Cancers* **106**, 17–21 (2017).
53. Christopoulos, P. et al. Detection of TP53 mutations in tissue or liquid rebiopsies at progression identifies ALK(+) lung cancer patients with poor survival. *Cancers* **11**, <https://doi.org/10.3390/cancers11010124> (2019).
54. Christopoulos, P. et al. Identification of a highly lethal V3(+)/TP53(+) subset in ALK(+) lung adenocarcinoma. *Int. J. Cancer* **144**, 190–199 (2019).
55. Christopoulos, P., Kirchner, M., Endris, V., Stenzinger, A. & Thomas, M. EML4-ALK V3, treatment resistance, and survival: refining the diagnosis of ALK(+) NSCLC. *J. Thorac. Dis.* **10**, S1989–S1991 (2018).
56. Sanfiorenzo, C. et al. Two panels of plasma microRNAs as non-invasive biomarkers for prediction of recurrence in resectable NSCLC. *PLoS ONE* **8**, e54596 (2013).
57. Shen, Y. et al. microRNA expression profiles associated with survival, disease progression, and response to gefitinib in completely resected non-small-cell lung cancer with EGFR mutation. *Med. Oncol.* **30**, 750 (2013).
58. Grunnet, M. & Sorensen, J. B. Carcinoembryonic antigen (CEA) as tumor marker in lung cancer. *Lung Cancer* **76**, 138–143 (2012).
59. Kuo, Y. S. et al. Association of divergent carcinoembryonic antigen patterns and lung cancer progression. *Sci. Rep.* **10**, 2066 (2020).
60. Castellanos-Rizaldos, E. et al. Exosome-based detection of EGFR T790M in plasma from non-small cell lung cancer patients. *Clin. Cancer Res.* **24**, 2944–2950 (2018).
61. Krug, A. K. et al. Improved EGFR mutation detection using combined exosomal RNA and circulating tumor DNA in NSCLC patient plasma. *Ann. Oncol.* **29**, 2143 (2018).
62. Wessels, S. et al. Comprehensive serial biobanking in advanced NSCLC: feasibility, challenges and perspectives. *Transl. Lung Cancer Res.* **9**, 1000–1014 (2020).
63. Newman, A. M. et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat. Med.* **20**, 548–554 (2014).
64. Newman, A. M. et al. Integrated digital error suppression for improved detection of circulating tumor DNA. *Nat. Biotechnol.* **34**, 547–555 (2016).
65. Reisinger, E. et al. OTP: an automatized system for managing and processing NGS data. *J. Biotechnol.* **261**, 53–62 (2017).
66. Eisenhauer, E. A. et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur. J. Cancer* **45**, 228–247 (2009).
67. Ruopp, M. D., Perkins, N. J., Whitcomb, B. W. & Schisterman, E. F. Youden Index and optimal cut-point estimated from observations affected by a lower limit of detection. *Biom. J.* **50**, 419–430 (2008).

## ACKNOWLEDGEMENTS

This work was supported by the German Center for Lung Research (DZL), by the German Cancer Consortium (DKTK), by the Heidelberg Center for Personalized Oncology at the German Cancer Research Center (DKFZ-HIPO), and by Roche Sequencing Solutions (Pleasanton, CA, USA). The authors would like to thank Ingrid Heinzmann-Groth, Karin Schnorr-Teichert, and Saskia Oestlinger of the Lung Biobank Heidelberg for assistance with patient sample collection. We also acknowledge the Genomics and Proteomics Core Facility at the German Cancer Research Center (DKFZ) for sample sequencing, and the Omics IT and Data Management Core Facility at the DKFZ for data management and processing. The authors thank Thomas Hielscher from the Division of Biostatistics at the DKFZ for support.

## AUTHOR CONTRIBUTIONS

A.K.A.: conceptualization, methodology, data acquisition, formal analysis, writing—original draft, writing—review and editing, visualization; P.C.: conceptualization, methodology, data acquisition, formal analysis, writing—original draft, writing—review and editing, visualization, supervision; Z.Y.: methodology, formal analysis, visualization, writing—review and editing; S.B. and F.J.: methodology, writing—original draft, writing—review and editing; S.J.O.: data acquisition, writing—review and editing; M.R.: conceptualization, data acquisition, methodology, writing—review and editing; M.S.: methodology, data acquisition, formal analysis, writing—review and editing; M.M.: methodology, data acquisition, writing—review and editing; M.A.S.: data acquisition, methodology, writing—review and editing; S.D.: conceptualization, methodology, data acquisition, formal analysis, writing—review and editing; A.S.: methodology, data acquisition, supervision, writing—review and editing; M.T.: data acquisition, formal analysis, writing—review and editing; H.S.: conceptualization, methodology, writing—original draft, writing—review and editing, supervision, project administration, funding acquisition.

## FUNDING

Open Access funding enabled and organized by Projekt DEAL.

## COMPETING INTERESTS

P.C. reports research funding from AstraZeneca, Novartis, Roche, and Takeda as well as advisory board and/or lecture fees from Boehringer Ingelheim, Chugai, Lilly, Pfizer, and Takeda, outside the submitted work. M.A.S. reports personal fees from the German Center for Lung Research (DZL), outside the submitted work. S.D. reports speaker's honoraria from Roche and research funding from Takeda, outside the submitted work. M.R. reports personal fees from Amgen, AstraZeneca, BMS, Boehringer Ingelheim, Lilly, Merck, MSD, Novartis, Pfizer, Roche, and Samsung, outside the submitted work. M.T. reports advisory board honoraria from Novartis, Lilly, BMS, MSD, Roche, Celgene, Takeda, AbbVie, Boehringer, speaker's honoraria from Lilly, MSD, Takeda, research funding from AstraZeneca, BMS, Celgene, Novartis, Roche and travel grants from BMS, MSD, Novartis, Boehringer, outside the submitted work. A.S. reports advisory board honoraria from BMS, Bayer, AstraZeneca, Thermo Fisher, Novartis, Seattle Genomics speaker's honoraria from BMS, Bayer, Illumina, AstraZeneca, Novartis, Thermo Fisher, MSD, Roche, as well as research funding from Chugai, Bayer and BMS, outside the submitted work. H.S. reports grants from Roche Sequencing Solutions, during the conduct of the study, and personal fees from Roche, outside the submitted work. The remaining authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41698-021-00239-3>.

**Correspondence** and requests for materials should be addressed to Holger Sultmann.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the

article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021