Introduction

Worldwide, lung cancer is the most common malignancy and by far the leading cause of cancer deaths, as most patients are diagnosed at an advanced, inoperable stage of disease (1). Only with the discovery of actionable oncogenic driver alterations, the outcome of advanced-stage non-small cell lung cancer (NSCLC) patients with biomarker-driven treatment has significantly improved. Particularly lung adenocarcinoma (LUAD) is not a single disease but a cluster of distinct molecular subtypes defined by a single oncogenic driver alteration, comprising gene mutations, rearrangements, and amplifications (2). Importantly, about 70% of these oncogenic driver alterations can be targeted by approved or investigational drugs. So far, EGFR, ALK,
ROS1, BRAF and in some countries NTRK inhibitors are approved by health care regulatory authorities in advanced-stage NSCLC harboring the respective oncogenic alteration (ESMO and NCCN guidelines) (3). Recurrent gene rearrangements act as strong oncogenic drivers, leading to a state of oncogene addiction and, therefore, their fusion proteins are ideal targets for anticancer drugs. With the availability of improved detection strategies, there has been an exponential discovery of gene fusions across various malignancies, which is paralleled by the development of new targeted drug compounds (4,5).

In this review, we will summarize the growing number of relevant fusion genes in NSCLC, highlight their underlining molecular mechanism, and discuss different methods for the detection of fusion genes. In addition, we will provide examples of how the detection of gene fusions can be integrated into a predictive testing algorithm to guide treatment selection of NSCLC patients. We present the following article in accordance with the NARRATIVE REVIEW reporting checklist (available at http://dx.doi.org/10.21037/tlcr-20-676).

**Tyrosine kinase (TK) fusions in lung cancer and molecular mechanisms**

In NSCLC several targetable gene fusions involving ALK, ROS1, NTRK, RET, NRG1, and FGFR have been discovered (3,6). Overall these are rare events with ALK rearrangements being the most prevalent in 3–5% of all advanced-stage LUAD, followed by ROS1 with a prevalence of 1–2% (7,8). NTRK fusions are extremely rare and only present in 0.2% of NSCLC and can involve all of the three NTRK genes (9). However, it should be emphasized that the detection of even these rare NSCLC patients is important as they show an excellent response to TK inhibitors (3). RET fusions are currently investigated in clinical trials and found in 2% of NSCLC. First studies with selective RET inhibitors show remarkable response and RET testing should be offered, to allow patients access to either clinical trials or off-label use of the drug (10). Several other rearrangements are known in lung cancer but have not yet proven to be of clinical benefit in terms of treatment. These include NRG1, especially in invasive mucinous LUAD, FGFR1/2/3 (mostly in squamous cell carcinomas) and PDGFR, MET and BRAF rearrangements (5,11).

All of these oncogenic gene fusions involve TK and result from structural rearrangements on the DNA level, such as interchromosomal translocations or intrachromosomal inversions, deletions and insertions (12). Oncogenic rearrangements put the TK domain of the involved gene under the promotor control of a partner gene, which leads to overexpression of chimeric fusion proteins. Otherwise, constitutive activation can be driven by oligomerization mediated by the fusion partner. Importantly, the fusion preserves the TK function with a conservative breakpoint region in the TK domain and numerous fusion partners, resulting in many gene fusion variants (5). For example, the breakpoint of ALK is highly conservative and located at exon 20, with various breakpoints in the amino-terminal part of EML4, leading to different EML4-ALK fusion variants. Over 20 different ALK fusion variants have been described in NSCLC, mostly comprising different EML4-ALK isoforms and rarely involving other fusion partners (13). The same principle applies also to ROS1 and NTRK rearrangements (14). Importantly all variants result in fusion proteins, which are optimized for constitutive ligand-independent kinase activation and downstream signaling and are therefore predictive for response to the respective TK inhibitor.

Therefore, screening methods should be able to detect all possible rearrangements irrespective of the fusion partner. In this review, we will discuss different methods for the detection of fusion genes and how they can be integrated into predictive testing algorithms to guide treatment selection of NSCLC patients for drugs targeting gene fusions.

**Clinicopathological features of NSCLC with gene rearrangements**

Gene rearrangements in NSCLC have been associated with several clinicopathological features. ALK, ROSI and RET rearrangement are enriched in younger patients, female sex and never-smokers (15–17). Data on NTRK fusions in NSCLC are still limited due to low numbers, but NTRK fusions appear to occur across gender, age, and smoking history (18). Several histological features are associated with rearrangements. ALK, ROSI and RET rearrangements are more frequent in LUAD with a solid-predominant pattern, and solid signet ring or mucinous cribriform patterns. Additionally, club cell (Clara cell)-like cells were found to be typical for ALK rearranged NSCLC and nuclei with macronucleoli for ROS1 rearranged NSCLC, respectively (16,19–21). Psammoma bodies and lymphangitic spread have been described as characteristic for a subset of
RET rearranged LUAD (22). The uncommon \textit{NRG1} rearrangements are enriched in but not restricted to \textit{KRA4S} wild-type invasive mucinous LUAD (23,24). Despite of these associations, neither clinical nor histological features are reliable enough to predict or exclude predictive gene rearrangements in non-squamous NSCLC. Moreover, although these rearrangements are mostly found in LUAD or non-small cell cancer, not otherwise specified (NSCC-NOS), they have rarely been reported in lung squamous cell carcinoma (25).

**Diagnostic approaches**

Several methods are at disposal to demonstrate, directly or indirectly, the presence of a gene fusion in tumor samples. Fusions can be detected on the protein level by immunohistochemistry (IHC), on the DNA level by fluorescence in-situ hybridization (FISH) or DNA-based targeted next-generation sequencing (NGS) and on the RNA level by reverse transcriptase-polymerase chain reaction (RT-PCR), multiplex digital color-coded barcode technology or RNA-based targeted NGS (26). Each of these methods has its advantages and disadvantages, which should be taken into account when performing and evaluating an analysis for gene fusions. Notably, all techniques discussed below are not only applicable to histological but also to cytological specimens as long as the required quality parameters are fulfilled and continuous quality assurance is in place (27-29).

**IHC**

In LUAD or NSCC-NOS, predictive IHC for ALK and ROS1 is common practice and has recently been complemented by pan-Trk IHC for prescreening of NTRK rearrangements (30,31). ALK, ROS1 and pan-Trk expression detected by IHC is a surrogate for the respective rearrangements as protein levels of these genes are mostly below detection limit by IHC in their native form. IHC is cost-effective, has a fast turn-around time, is way simpler to evaluate compared to molecular methods, and can easily be integrated into a diagnostic laboratory. Moreover, IHC only needs a small amount of tumor cells on one unstained formalin-fixed and paraffin-embedded (FFPE) slide or one cytological specimen. Examples of ALK and ROS1 IHC are shown in Figure 1.

Because preanalytic procedures, antibody clones, and detection platforms vary between laboratories, a proper validation of IHC protocols is crucial to ensure accurate and reproducible IHC staining. General recommendations for analytic validation of predictive laboratory-developed tests (LDT) have been published (32). These guideline statements propose that for initial analytical validation of a new predictive LDT protocol, a minimum of 20 positive and 20 negative controls, fixed and processed in the same manner as the clinical cases, should be tested. The LDT result should ideally be compared with the result of a validated assay on the same validation set and achieve an overall concordance of at least 90%. Benign tissue controls (for example ganglion cells of the appendix for ALK and brain or testis for pan-Trk) or cell lines harboring the respective fusions can serve as positive controls. However, clinical tumor specimens with a known rearrangement are the optimal positive control. The high number of proposed positive controls is unrealistic in this setting, as the prevalence of oncogenic rearrangements in NSCLC is low (NTRK only 0.2%). Internal and external quality control measures are therefore crucial to ensure accurate IHC results.

**ALK**

For ALK IHC two highly sensitive antibody clones are available, 5A4 (Novocastra™, Leica Biosystems) and D5F3 (Cell Signaling, Ventana). Many studies have demonstrated that the performance of both antibodies using well-validated IHC protocols is very good compared to FISH. The pooled sensitivity and specificity for both antibodies are 97% and 100%, respectively (30). In addition to laboratory developed tests there is a commercially available and FDA approved, highly standardized, fully automated ALK kit assay for BenchMark immunostainers using D5F3 (Ventana). Discordant ALK IHC and FISH results are rare and have been reported in 1% of NSCLC (33). It has been suggested that these are mainly driven by false-positive FISH results, especially in borderline FISH positive NSCLC with ALK-positive cells ranging from 15% to 20% (34). Aberrant ALK staining has rarely been reported in high-grade neuroendocrine carcinomas (35).

Based on published evidence and according to current predictive testing guidelines, ALK IHC, using 5A4 or D5F3, is an equivalent alternative to FISH, which used to be the gold standard for ALK testing (30). However, evaluation of ALK IHC is not standardized and criteria for an ALK-positive result vary significantly across different studies (36). ALK staining in ALK-rearranged NSCLC is usually cytoplasmic and diffuse across the tumor with moderate to strong intensity (Figure 1A,B). In NSCLC with
moderate to strong and diffuse ALK staining, confirmation by a molecular method can be omitted. Nonetheless, the threshold to perform confirmation by a molecular method should be low and should be performed in case of heterogeneous, focal, or weak staining. As a matter of fact, we still confirm every ALK IHC positive case by FISH for quality assurance purposes at our institute.

**ROS1**

For ROS1 IHC there are two commercially available antibody clones, the D4D6 (Cell Signaling Technology) and the recently introduced SP384 (Ventana). In contrast to ALK, no commercial ROS1 assay is available, and ROS1 IHC has to rely on LDTs. Additionally, compared to ALK, significantly fewer studies have investigated the performance of ROS1 IHC. However, these studies demonstrated a high sensitivity and specificity using D4D6 ranging from 94–100% and 87–100%, respectively (30). The first study investigating SP384 showed a similar performance (37). ROS1 rearranged NSCLC typically reveals finely granular cytoplasmic IHC staining, though the EZR-ROS1 fusion can result in a membranous staining (8). The staining can be more heterogeneous compared to ALK-positive carcinomas with variable staining intensities within the same tumor (8). Although ROS1 protein is essentially absent in normal human lung tissue, a non-specific IHC staining may be observed in reactive alveolar type II pneumocytes and macrophages and should not be misinterpreted as a positive result. According to current predictive testing guidelines, ROS1 IHC may be used as a screening test, but positive
ROS1 IHC results should be confirmed by a molecular method (30).

**NTRK**

NTRK includes three genes, NTRK1, 2, and 3, which encode transmembrane receptor TK TrkA, B, and C, respectively (38). These three proteins have a high level of homology between the kinase domains and can all be detected by pan-Trk antibodies. Commercially available pan-Trk clones are EPR17341 (Abcam) and A7H6R (Cell Signaling Technologies). Trk expression by IHC can be variable in intensity and subcellular localization (cytoplasmic, nuclear or membranous), which might depend on the 5’ fusion gene partner (31). The largest study on the performance of pan-Trk IHC across different tumor types, including lung cancer, has been performed using the EPR17341 antibody (9). The sensitivity for detecting NTRK fusions seems higher for NTRK1 and 2 (96% and 100%, respectively) and lower for NTRK3 (79%). Additionally, specificity differs between different tumor types. Lung cancer showed a sensitivity of 88% and a specificity of 100%. Again, there is no standardized pan-Trk IHC interpretation. Any positive staining, cytoplasmic, nuclear or membranous, even if only present in 1% of tumor cells, should be evaluated by a molecular method (9,31). Pan-Trk IHC can be used for screening in NSCLC with confirmation of positive results by a molecular method, preferentially by RNA-sequencing (31). However, in the absence of other driver alterations (EGFR, KRAS, ALK, and ROS1), molecular confirmation of a negative pan-Trk staining might also be considered as pan-Trk IHC can miss up to 20% of NTRK3 rearrangements.

In summary, since ALK, ROS1 and NTRK fusions are rare, IHC greatly simplifies large-scale screening of NSCLC and reduces costs, since IHC-negative NSCLC do not normally require molecular testing for fusions already screened by IHC.

**Other fusion proteins**

No established antibody exists for the detection of RET fusions (39). Similarly, there is no reliable antibody to detect NRG1 fusions. However, studies have reported that phospho-HER3 (p-ERBB3) IHC has a high sensitivity of 100% and a specificity of 97.5% for the detection of NRG1 fusions in LUAD (23,40). Although several studies have included IHC for FGFR1, 2 or 3, none of these antibodies has proven reliable sensitivity and specificity, in particular for FGFR3 fusions, the most frequent fusion (41,42). For other rare fusions involving genes such as MET, BRAF, EGFR, and PIK3CA, there are also no antibodies available for accurate detection of the respective fusion proteins (39).

**FISH**

FISH can detect large structural variants at the DNA level and is widely used in clinical laboratories to test for oncogenic fusions (Figure 2). Its advantage is the little amount of tissue needed and that fusions can be detected within the cells of interest. However, besides the expertise required and its labor-intensive nature, FISH analysis also has some technical shortcomings one has to keep in mind. When using fusion probes, the fusion partner has to be known and only one partner at a time can be detected. Otherwise, break-apart probes are necessary, which cover all possible partners, with the downside that the fusion partner remains unknown. In theory, break-apart probes can detect large structural variants with sufficient sensitivity and specificity, however, short inversions and intrachromosomal translocations may be missed by FISH due to an insufficient splitting signal as has been shown, e.g., for ALK (43). This notion may also be relevant for NTRK1 fusions, which are often intrachromosomal events (44). Likewise, translocations can be complicated by deletions or atypical fusion signals, leading to false-negative results (45). Moreover, while a break-apart probe with a split positive signal might show a structural variant involving the gene probed, it cannot be determined whether the abnormal signal actually results in the generation of a fusion transcript or protein (46). Indeed, it has been shown that patients whose tumors were ALK-negative by IHC but positive for FISH, did not respond to ALK inhibitor therapy, except when IHC was negative due to poor sample quality (47,48). Conversely, ALK IHC positive but FISH negative NSCLC have been associated with impaired survival after targeted treatment (49). Finally, FISH testing for NTRK1/2/3 fusions is laborious and expensive in routine practice since it requires three different break-apart probes for coverage.

**RT-PCR**

Extracted RNA can be reverse-transcribed and RT-PCR can be performed to either qualitatively or quantitatively detect the presence of a single oncogenic fusion for which both fusion partners are known. The advantage of this method is the low cost per assay and the high sensitivity and specificity. However, because of the large number
of different fusion partners and break-points involved, the utility of RT-PCR for individual fusion transcripts is limited, because each fusion would need to be detected separately. However, differences in expression of the 5' versus the 3' end of a gene can provide indirect evidence of a fusion, a phenomenon which is also being used in NGS technology using the ratio of imbalance (50).

**Multiplex digital color-coded barcode technology**

Among the non-NGS fusion detection methods, multiplex digital color code technology has recently gained attention, in particular the nCounter platform (NanoString, Seattle, WA). This method enables the detection of fusions by direct counting of specific mRNA molecules without any retro-transcription of amplification steps. Thus, it is particularly suitable for degraded RNA and samples with low amount of RNA (51). Because it detects known fusion transcripts but also measures the 3' and 5' gene region imbalance, this technique can detect previously unknown fusions. Moreover, it provides the possibility to test several fusions at the same time such as ALK, ROS1, RET and NTRK1. nCounter fusion gene assays have been successfully used in both histological and cytological samples with high concordance to FISH results and represents a valuable technique (52-54).

**NGS**

NGS provides a precise method to detect fusions in lung cancer. One main advantage is that multiple fusions and their corresponding partners can be tested at the same...
time from one single tumor sample. In contrast, turnaround time and costs are higher in comparison to IHC and FISH and more tumor material is necessary. Several NGS approaches to detect fusions exist, which can be applied on either DNA or RNA level.

**DNA based NGS**

The main application of NGS analysis on tumor samples is the assessment of the somatic mutational status. This is performed in most diagnostic laboratories by a targeted NGS panel, which covers the most frequently and clinically relevant mutated genes for a given entity. Besides, some of the respective panels also allow identifying a large series of different fusions (6). One of the main challenges in detecting fusions by DNA-based NGS is that most genomic breakpoints leading to fusion genes occur in introns. These introns cannot always be fully covered by targeted NGS panels because they are either too long or contain repetitive elements (6,55,56). Consequently, inadequate coverage and difficulty in assessing highly repetitive regions can lead to false-negative results. Besides, and analogous to FISH, some fusions discovered by DNA-based NGS panels might not have a functional consequence, representing a non-functional event.

**RNA based NGS**

In the case of RNA-based NGS, the mature mRNA is sequenced, which has several advantages. Mature mRNA is devoid of long introns, which facilitates sequencing and data analysis. The detection of fusions at the RNA level also provides direct evidence that these fusions are indeed transcribed, which increases confidence in the results (46). Fusion transcripts can also be detected at low tumor cell content because they are often highly expressed and unique in the tissue. Finally, RNA-based NGS can test for multiple fusions simultaneously, which justifies to some extent its higher costs and longer turnaround time. For these reasons, an RNA-based NGS would be the method of choice. However, the most important disadvantage is that RNA is less stable than DNA and the quality of the RNA does not always meet the requirements of sequencing. Especially with RNA from FFPE samples this can lead to a higher dropout rate and thus possibly to false negative results.

Many NGS platforms allow for the detection of fusions on RNA based NGS, either using an amplicon-based or hybrid capture methodology. Amplicon based methods enrich for target genes by PCR amplification of a distinct set of genes but can only detect fusion partners that are already known and included in the panel. Alternatively, fusion detection by an imbalance of 5’ to 3’ gene expression can be used to detect fusions with unknown partners, however, this needs further confirmation. Capture based approaches for RNA fusion analysis work by transcribing RNA first in cDNA and subsequent sequencing similar to DNA based sequencing. With this method, only one fusion partner needs to be known. Finally, an anchored multiplex PCR (AMP), for example with the Archer FusionPlex platform, can be used which also allows for the detection of novel fusion partners. Thanks to the initial adapter ligation step that facilitates priming without a priori knowledge of the gene fusion partner, the AMP method has been shown to have high technical sensitivity and specificity even in FFPE-derived RNA samples (31).

Several studies have therefore used the AMP technology to select for patients with, e.g., NTRK fusions followed by FISH for confirmation (57).

These different methods and techniques need different amounts of RNA input. Therefore, when choosing RNA-based sequencing technology, the amount of available tissue must also be considered. Finally, because of the reduced RNA quality, methods have to be established to interrogate RNA quality such as measuring RNA fragment size distribution and examining amplification of a housekeeping gene in a quantitative PCR based assay (58). This will allow greater confidence in the fusion results obtained.

**Testing algorithms considerations**

Despite the scarcity of most targeted rearrangements in NSCLC, they should eventually be tested in every patient, at least in the case of wild-type status of other established driver genes (esp., EGFR and KRAS). While the importance of identifying patients that could benefit from targeted therapy is unquestioned, feasibility in terms of tumor cell quantity, locally available testing methods, access to drugs, and economic aspects have to be taken into consideration when creating testing algorithms and guidelines. Methods might need to be combined and algorithms flexibly adapted. Simultaneous frontline DNA- and RNA-based NGS for mutation and rearrangement testing appears to be the most comprehensive and preferable approach. In contrast, sequential testing by restricting RNA-based rearrangement testing to the 50% of patients without exclusive driver mutations is more economical but can lead to treatment delay. For NTRK, which is particularly rare and challenging,
several guidelines have been published (26,31,46). In our practice, we initiate IHC to prescreen for ALK, ROS1, and NTRK rearrangements simultaneously with mutational testing by NGS in all qualifying NSCLC (Figure 3). This allows the rapid identification of the rare 5–7% of patients with predictive rearrangements within one day. Positive or ambiguous results are further evaluated using an additional method such as FISH or RNA-based NGS. RNA-based NGS with a panel that covers fusions for RET, NRG1, FGFR2/3, and MET exon 14 skipping mutation is the preferred method for NSCLC which are wild-type for EGFR, KRAS, ALK, ROS1 and pan-Trk negative. RET FISH is a valid final analysis if the material is insufficient for RNA-NGS panel tests.

Acknowledgments

Funding: None.
Footnote

Provenance and Peer Review: This article was commissioned by the Guest Editors (Silvia Novello, Francesco Passiglia) for the series “Looking for Chimeras in NSCLC: Widen Therapeutic Options Targeting Oncogenic Fusions” published in Translational Lung Cancer Research. The article has undergone external peer review.

Reporting Checklist: The authors have completed the NARRATIVE REVIEW reporting checklist. Available at http://dx.doi.org/10.21037/tlcr-20-676

Peer Review File: Available at http://dx.doi.org/10.21037/tlcr-20-676

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/tlcr-20-676). The series “Looking for Chimeras in NSCLC: Widen Therapeutic Options Targeting Oncogenic Fusions” was commissioned by the editorial office without any funding or sponsorship. Dr. MSM reports personal fees from Thermo Fisher, outside the submitted work. Dr. SS reports personal fees from MSD, personal fees from Astra Zeneca, personal fees from Boehringer Ingelheim, personal fees from Roche, personal fees from Pfizer, personal fees from Thermo Fisher Scientific, outside the submitted work. Dr. LB reports grants from Sanofi, grants and personal fees from Roche, grants and personal fees from MSD, personal fees from Astra Zeneca, personal fees from Bayer, personal fees from BMS, personal fees from Boehringer Ingelheim, personal fees from Pfizer, personal fees from Takeda, outside the submitted work. The other author has no other conflicts of interest to declare.

Ethical Statement: the authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References


