
ABSTRACT
The advances in understanding the inherited biological mechanisms of non-small cell lung cancer harbouring epidermal growth factor receptor (EGFR) mutations led to a significant improvement in the outcomes of patients treated with EGFR tyrosine kinase inhibitors. Despite these clinically impressive results, clinical results are not always uniform, suggesting the need for deepening the molecular heterogeneity of this molecularly defined subgroup of patients beyond the clinical and biological surface. The availability of tissue and blood-based tumour genotyping allows us to improve the understanding of molecular and genetic intratumor heterogeneity, driving the measurement of clonal evaluation in patients with lung cancer carrying EGFR mutations. Genetic diversification, clonal expansion and selection are highly variable patterns of genetic diversity, resulting in different biological entities, also a prerequisite for Darwinian selection and therapeutic failure. Such emerging pieces of evidence on the genetic diversity, including adaptive and immunomodulated aspects, provide further evidence for the role of the tumour microenvironment (TME) in drug-resistance and immune-mediated mechanisms. Matching in daily clinical practice, the detailed genomic profile of lung cancer disease and tracking the clonal evolution could be the way to individualise the further target treatments in EGFR-positive disease. Characterising the tumour and immune microenvironment during the time of the cancer evaluation could be the way forward for the qualitative leap needed from bench to bedside. Such a daring approach, aiming at personalising treatment selection in order to exploit the TME properties and weaken tumour adaptivity, should be integrated into clinical trial design to optimise patient outcome.

INTRODUCTION
Lung cancer is the main cause of death for cancer worldwide. In the last decades, many efforts have been spent in order to improve the overall survival (OS) and quality of life of patients with advanced-stage non-small cell lung cancer (NSCLC). In particular, significant results were obtained from several clinical trials that investigated the adoption of tyrosine kinase inhibitors (TKIs) instead of chemotherapy, improving significantly overall response rates (RRs) and progression-free survival (PFS) in different molecular subsets, including the EGFR-mutated subtype.2-6

The importance of the epidermal growth factor receptor (EGFR) gene molecular assessment for EGFR TKI administration was raised over a decade ago. As a matter of fact, the efficacy of these novel drugs is subordinated to the identification of sensitising mutations in the EGFR gene.7 From an epidemiological point of view, EGFR mutations range from 10% to 15% in Caucasian patients and up to 50% in East-Asian patients, and are identified more frequently in patients with adenocarcinoma, who are female and who have never smoked or are former smokers.8 9 For this reason, the College of American Pathologists, the International Association for the Study of Lung Cancer, the Association for Molecular Pathology, the National Comprehensive Cancer Network and the American Society of Clinical Oncology guidelines identified several genes to that necessarily require testing in patients with advanced NSCLC, including EGFR.10-12 Noteworthy, despite the clinical efficacy of EGFR TKIs in the vast majority of patients with advanced-stage NSCLC harbouring a sensitising EGFR mutation, is that a non-negligible percentage of patients displayed mixed responses or progressive disease.13 Furthermore, the intratumor heterogeneous distribution of EGFR mutations has been shown to be involved as a resistance mechanism during TKI treatments.14 As a consequence, only cells harbouring EGFR mutations will respond to TKI action, with the remaining tumour cells, insensitive to the treatment, responsible for disease persistence and ultimately progression15 (figure 1). In this setting, the EGFR sensitive mutant tumour cells may coexist with other subclonal tumour cells with either different EGFR genetic alterations or with other gene mutations.16 Interestingly, the presence of multiple nodules is related
to a higher rate of heterogeneity.\textsuperscript{13} Besides, mutational heterogeneity is present in a non-negligible percentage of patients with advanced-stage NSCLC between the primary tumour and its metastases.\textsuperscript{17} Driver mutations (eg, EGFR) are early events in tumour development and for this reason are homogenously existing in all tumour cells, whereas other alterations may arise during cancer progression and adaptation.\textsuperscript{17, 16} Of note, a wide range of mutations can affect EGFR with different responsiveness to the different target treatments.\textsuperscript{17} In addition, EGFR mutations may coexist with other alterations in the same gene or other different genes.\textsuperscript{20, 21}

Patients harbouring EGFR sensitive mutations, who receive EGFR TKIs in the metastatic setting, usually achieved a significant disease response with prolonged survival.

Although the effectiveness of TKIs is confirmed by a multitude of preclinical and clinical studies, not all patients equally benefit from these target treatments,\textsuperscript{22} mainly confirming three clusters within the treated EGFR population, characterised by different prognosis subgroups: poor prognosis, characterised with limited survival and fast progression (OS<12 months); good prognosis, overlapping survival of pivotal randomized clinical trials (RCTs) (OS 24–30 months); excellent prognosis, characterised by doubled or tripled survival, compared with standards (OS more than 36 months). This survival heterogeneity appears closely linked to the tumour heterogeneity, through the increasing role of co-occurring mutations and immune-microenvironment, identifying new potential prognostic and predictive biomarkers in EGFR-positive disease. In this review, we overview the tumour heterogeneity of NSCLC harbouring EGFR mutations, from tumour clonality to co-occurring mutations, through the complex role of the tumour microenvironment (TME).

\textbf{EGFR TUMOUR CLONALITY}

The concept of clonal evolution of tumour cell population was reported for the first time as early as 1976.\textsuperscript{23} In this theory, neoplastic cells take origin from a single progenitor that subsequently acquires, under selective pressure, different genomic alterations.\textsuperscript{23} For this reason, these genomic alterations, which occur in the early phases of cancer development and give an advantage in cancer growth, are identified in all neoplastic cells.\textsuperscript{23, 24}

To date, lung adenocarcinoma seems to originate from a multistep progression, from atypical adenomatous hyperplasia to adenocarcinoma in situ, and finally invasive adenocarcinoma.\textsuperscript{25} In this evolution, EGFR driver alterations are acquired in the early step of cancer progression and can be identified in the vast majority of neoplastic cancer cells. Driver mutations, such as those in the EGFR, have been shown to be significantly more often truncal events compared with mutations in non-driver genes that are usually branch mutations.\textsuperscript{26} As a consequence, the heterogeneous distribution of EGFR mutations in lung adenocarcinomas is extremely rare, as demonstrated in a seminal study of Yatabe \textit{et al}.\textsuperscript{15} Of note, similar results were obtained by Sun \textit{et al}.\textsuperscript{27} In this study, identical EGFR mutations were identified in different areas of tumours featuring mixed histology.\textsuperscript{27} In the experience of Mattsson \textit{et al}, even if three different histological areas were selected for molecular analysis, the same EGFR molecular status emerged.\textsuperscript{28} In order to evaluate the spatial distribution of EGFR and Kirsten Rat Sarcoma Viral Oncogene Homolog mutations, Dietz \textit{et al} analysed central tumour sections (5 mm×5 mm segments) of lung mutated
adecarcinomas. Of note, driver mutations were identified in 462 (98.9%) out of 467 tumour segments with different allelic frequencies (range: 0.04–19.36). Sun et al evidenced the higher presence of EGFR mutations in cancer cells, performing a cytological fine needle aspiration (FNA) approach. A higher concordance (91.7%) was reached when the FNA molecular approach was compared with the histological one. Despite these pieces of evidence, it is currently known that tumours are characterised by distinct subclones with several genomic alterations. Intratumorl heterogeneity was reported in different cancer types, including NSCLC, as a consequence of genetic and epigenetic alterations derived from genomic and chromosomal instability and different patterns of clonal evolution over space and time. In fact, despite driver mutations (e.g., EGFR) occurring early in tumour growth and development and consequently homogenously distributing within the tumour, other alterations may arise during cancer progression and adaptation. Single-cell analysis may provide insight into the occurrence of intratumoral heterogeneity of EGFR mutations. In cases in which both EGFR-mutated and non-mutated neoplastic cells are present, response to TKIs may be of low intensity. When different neoplastic cells are discovered within the same lesion, only tumour cells harbouring EGFR sensitising mutations display responsiveness to TKI treatments. Conversely, the remaining non-mutated neoplastic cells, which are not sensitive to the target treatment, without the selective pressure may replace the decaying cells. In this setting, subclonal tumour cells without EGFR sensitising mutations may coexist with EGFR sensitive mutant tumour cells. Interestingly, the higher rate was evidenced when multiple nodules affect the same patient. As far as mutational heterogeneity is concerned, metastases feature different genomic alterations with respect to the primary site in a non-negligible percentage of patients with advanced-stage NSCLC. Liquid biopsy may overcome the limitation of spatial heterogeneity in lung cancer. Each single tumour cell actively or passively sheds nucleic acids into the bloodstream. This evidence may be relevant, in particular when resistance mutations arise.

Biodiversity of EGFR Mutations: Driver, Passenger and Co-occurring Mutations

As far as EGFR mutations are concerned, the vast majority is represented by in-frame deletions involving exon 19 (about 45%) and exon 21 p.L858R (about 40%). Of note, these mutations lie in the tyrosine kinase domain of EGFR protein and are targetable by TKIs. As early as 2004, Lynch et al and Paez et al discovered for the first time the driver role of EGFR mutations in patients with NSCLC. The authors reported for the first time that mutations involving the tyrosine kinase domain of EGFR protein might predict responsiveness to the first generation TKI gefitinib. The remaining (10%–15%) ‘uncommon’ EGFR mutations are still under investigation for their ability to predict response or resistance to specific EGFR TKIs. In this setting, a broad range of different alterations, covering exons 18–21, should be correctly classified in order to administrate the best treatment choice. Exon 18 mutations rarely occur (about 3% and 4%) in patients with advanced-stage NSCLC and limited literature focused the attention on their function. However, most frequently, the alterations within exon 18 lie in codons 719 and 709. Collectively, these mutations seem to be more sensitive to second-generation EGFR TKIs followed by third-generation, and then first-generation inhibitors (primary resistance or low responsiveness). In addition to the frequent classical deletions (comprising up to 30 alterations) investigated in the different clinical trials, exon 19 harbours many other less investigated deletions. Of note, exon 19 deletions may interest the entire exon (codons 746–761) and, in a non-negligible percentage of cases (>50%), may be associated with additional insertions (indels). Despite a high RR to all TKIs, it would be preferred to administrate the third-generation TKIs. Exon 20 harbours a heterogeneous group of mutations (point mutations, duplications, insertions). The resistance mutation p.T790M represents the most common EGFR exon 20 point mutation. This latter occurs in 50%–60% of patients with advanced-stage NSCLC with acquired resistance to first-generation or second-generation TKIs, but it is sensitive to the third-generation TKI osimertinib. Noteworthy, the prevalence of this alteration in treatment naive patients is quite low with (about 2%), and has been associated with inherited susceptibility to lung cancer. Exon 20 is also involved in other resistance mechanisms. Thress et al reported for the first time the EGFR exon 20 p.C797S resistance point mutation after treatment with osimertinib. Conversely, EGFR exon 20 insertions represent 4%–12% of all EGFR mutations, whereas less frequent point mutations are identified in codon 708 (p.S768I; about 1%). Nevertheless, in both cases, responsiveness to afatinib or osimertinib was reported. As far as exon 21 is concerned, the second most frequent mutated codon is 861 (p.L861Q; 1% and 2%). The spectrum of the response of this alteration is similar to that seen in exon 20 insertions and p.S768I. When considering EGFR mutations, it pivotal to distinguish between mutations able to confer an advantage in tumour growth and development (so-called ‘driver mutations’) and the other mutations that can arise in cancer cells without pathogenic features (so-called ‘passenger mutations’). In patients with NSCLC, tobacco habits may induce the highest rate of mutations. Of note, a high percentage of these alterations are passenger mutations, useful to identify a mutational signature characteristic of tobacco smoking. In particular, it was evidenced a higher percentage of signature 4 (C>A mutations) and 5 (C>T and T>C mutations) in lung cancer associated with smoking history. To this end, several efforts have been spent in order to classify these alterations correctly. In this setting, computational analysis may be helpful to define driver and passenger...
mutations. In particular, the analysis of amino acid residues through the protein in both wild-type and mutant proteins and the analysis of tridimensional structure should be taking into account. Anoosha et al. underlined that leucine and glycine substitutions in helix and strand are more frequently associated with driver mutations, whereas charged residues arginine and glutamic acid are more frequently associated with coil-buried and coil-exposed mutants, respectively. Finally, EGFR mutations can be associated with each other or with other genetic alterations which may be present. EGFR multiple mutations account for about 25% of patients with EGFR mutations. In the vast majority of cases, classical sensitising mutations are associated with additional rare alterations. In these cases, second-generation and third-generation TKIs may play a pivotal therapeutic role. Interestingly, EGFR mutations may be associated with other gene alterations, in particular in tumor protein P53 (TP53). Yu et al. identified in pretreated EGFR-mutated samples co-occurring mutations in TP53, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha (PIK3CA), catenin beta 1 and retinoblastoma 1 (RB1). The authors underlined a shorter time to progression on EGFR TKI when a TP53 mutation was evidenced. In another study, Chen et al. identified co-occurring mutations in TP53, RB1, PIK3CA, FA tumour suppressor homolog 1 (FAT1), or ATP-binding cassette subfamily B member 1 (ABCB1), mitogen-activated protein kinase kinase 2. Another gene that may occur in association with EGFR is represented by AT-rich interaction domain IA. Despite its clinical role is not completely explain, the authors hypothesised that this association might limit targeted therapy response. Recent findings showed that co-mutations can occur in several druggable genes, as well, including MET deregulation, BRAF mutations, HER2 amplification and gene fusions. Despite the encouraging results from dual TKI inhibition at the occurrence of EGFR TKI resistance, much less is known about the role of targetable co-mutations in EGFR TKI-naïve patients, and future investigation is needed in this setting.

**IMMUNE HETEROGENEITY IN EGFR-POSITIVE NSCLC**

EGFR-mutant NSCLCs represent a big challenge for the immunotherapy treatments that have dramatically changed survival in NSCLC but still have not found a role in treating patients with EGFR-positive NSCLC. Though in a phase II trial, patients with pretreated EGFR/ALK-positive aNSCLC showed similar RRs to durvalumab when compared with wild-type population, the main subgroup analysis and meta-analysis of randomised controlled trials with immunotherapy are concordant in demonstrating no improvements in OS compared with standard chemotherapy. The only exception to this evidence, to date, is represented by the first-line combination treatment of chemo-immunotherapy with an antiangiogenic drug (IMpower150 trial, evaluating the addition of atezolizumab to carboplatin/paclitaxel and bevacizumab), where the benefit in PFS and OS is present regardless of the EGFR/ALK status. However, these results should be interpreted with caution given the limited number of patients included and the heterogeneity of patients with EGFR mutations enrolled (patients with sensitising and resistant mutations, TKI-naïve and TKI-pretreated patients). Further prospective studies are required, and some ongoing clinical trials are exploring this question.

Those controversial evidences well reflect the immune heterogeneity of EGFR-mutant NSCLC, where the efficacy of treatments with immune checkpoint inhibitors is dependent on the strong interplay among tyrosine kinase pathway mediators, programmed death-ligand 1 (PD-L1), TME factors such as vascular endothelial growth factor (VEGF) and interferon-gamma (IFNg).

It is well established that EGFR-positive NSCLC is associated with low mutational burden, consistently with the evidence that this kind of molecular alteration is more common in no/light smoker patients. In contrast, tumour mutation burden (TMB) is strongly related to smoking history. As TMB is related to response to monotherapy with immune checkpoint inhibitors, the low TMB in EGFR-mutant NSCLCs is consistent with the lack of benefit from such treatments in these patients. In contrast with this aspect, TMB was found to be a negative prognostic factor for EGFR mutant NSCLC treated with EGFR TKIs. Interestingly, a recent work by Hastings et al. showed that EGFR L858R and G719 tumours have higher TMB compared with EGFR del19 tumours, consistently with the finding of worse outcome with ICIs of EGFR del19 tumours compared with EGFR WT.

Conversely, PD-L1 expression shows an opposite behaviour than TMB in EGFR-mutant NSCLC, as it is frequently highly expressed in oncogene-addicted tumours, both at preclinical and clinical level. This finding is apparently in contrast with the data showing an increase in response to immunotherapy with the increasing levels of PD-L1. Indeed, PD-L1 expression in EGFR-mutant cells is the result of signalling pathways that are activated downstream of the receptor tyrosine kinase (RTK). Phosphoinositide 3-kinase/AKT pathway, as well as mitogen-activated protein kinase kinase and signal transducer and activator of transcription 3 (STAT3) through Src and Src-homology region 2 domain-containing phosphatase-2, can induce upregulation of PD-L1. EGFR inhibition with EGFR TKIs decreases PD-L1 levels, which are restored at the occurrence of TKI resistance. Since EGFR TKI resistance is commonly mediated by the activation of other RTKs and downstream mediators, a profound role of cross-talking pathways and signaling molecules as immune modulators are emerging and attempts to combine TKIs and immunotherapy are currently ongoing.

The immune features of EGFR-mutant cells are also strongly dependent on the TME. EGFR-mutant tumours
show a complex interaction with the TME, leading to an increase in T regulatory cells (T regs), decrease in tumour infiltrating lymphocytes and downregulation of major histocompatibility complex.83

IFNg, secreted by immune infiltrating cells, modulates STAT3/STAT1 balance through Janus kinase and, consequently, mediates PD-L1 expression. Moreover, through the activation of cyclin-dependent kinase 5, it also inhibits the activity of PD-L1 repressors.84 85 In EGFR-mutant cells, where the activity of STAT3 is crucial for survival, the ability of IFNg of modulating STAT3 plays an essential role in immune modulation.

Also, VEGF is essential in EGFR-mutant NSCLC, not only because of its well-established role in EGFR-VEGF cross-talk, alteration affecting peritumoral and intratumoral vascularisation and consequently drug delivery impairment and EGFR TKI resistance.86 Indeed, VEGF is also an important immune modulator. In the presence of VEGF, myeloid-derived suppressor cells are stimulated to migrate and accumulate within tumour size, where they are responsible for the increase in T regs and decrease in T cytotoxic cells’ activity.87 88 This mechanism is probably responsible for the synergism observed with immune checkpoint inhibitors and antiangiogenic drugs in patients with EGFR-mutant NSCLC.89

The complex mechanisms of EGFR-mediated immune modulation are thus at the basis of the dynamic immune heterogeneity in EGFR-mutant NSCLC, which reflects the influence of EGFR-activating status in different moments of the lung cancer disease, EGFR TKI-naïve, TKI treatment, TKI resistance and subsequent treatments.

CONCLUSIONS
Several distinct features contribute to heterogeneity in EGFR-positive lung tumours. Tumour clonality affects intratumoral heterogeneity, whereas the biodiversity of EGFR mutations and the presence of co-mutations have an impact also on immune heterogeneity and clinical heterogeneity. Indeed, specific subtypes of EGFR mutations determine different patterns of response to EGFR TKIs, ranging from high and prolonged sensitivity to minimal or no benefit. The presence of co-occurring mutations can reduce EGFR TKI activity driving earlier resistance to EGFR inhibition sustained by the selection of resistant cell clones. On the other hand, the presence of co-mutations may increase TMB, therefore influencing tumour immunogenicity and subsequent potential efficacy of immune-modulating drugs.

In parallel, TME is dynamically influenced by EGFR signalling pathways. Consequently, it may substantially differ at different tumour sites, not only due to intrinsic organ-specific features but also as a reflection of EGFR tumour clonality across metastatic sites.

Current clinical standard of care in EGFR-mutant lung cancer is barely able to face this complex biological and clinical scenario. The described mechanisms responsible for intratumoral, clonal, immune and clinical heterogeneity are not easy to assess in a comprehensive and dynamic study. In our view, a multilevel diagnostic approach based on both tissue and blood next-generation sequencing should always be considered, when available, to obtain a more comprehensive snapshot of EGFR-mutant disease. The application of such a systematic and dynamic approach including repeated tissue and liquid biopsies at disease progression could be a highly effective bench-to-bedside method, with the potentiality to better select treatments according to specific features and correlates of EGFR heterogeneity.

Twitter Antonio Passaro @apasaromd, Umberto Malapelle @UmbertoMalapelle, Marzia Del Re @Marzia_Del_Re, Pasquale Pisapia @PasqualePisapia and Christian Rolfo @ChristianRollo

Contributors AP, UM, MDR, IA and CR developed the first draft of the manuscript. All the authors revised the manuscript by refining the contents with feedback and comments. All the authors approved the final draft before the submission.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests AP has received honoraria for consulting, advisory role or lectures from AstraZeneca, Agilent/Dako, Boehringer Ingelheim, Bristol-Myers Squibb, Eli Lilly, Merck Sharp & Dohme, Pfizer and Roche Genentech. UM has received personal fees from Boehringer Ingelheim, Roche, MSD, Amgen, Merck and AstraZeneca. FDM has served in a consultant/advisory role for AstraZeneca, Boehringer Ingelheim, Bristol-Myers Squibb, Celgene, Merck Sharp & Dohme, Novartis, Roche Genentech, Takeda and Pfizer. CR has received speakers’ bureau from AstraZeneca and MSD, a research grant from the Lung Cancer Research Foundation–Pfizer, and research support from Guardant Health and Biomark; has an advisory board role with ARCHER, Inivata and Merck Serono; and has consulted for Mylan and Oncopass.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, any changes made are indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID IDs
Antonio Passaro http://orcid.org/0000-0002-7575-3870
Marzia Del Re http://orcid.org/0000-0001-7343-6161
Pasquale Pisapia http://orcid.org/0000-0002-6429-0620
Romano Danesi http://orcid.org/0000-0002-4414-8934
Christian Rolfo http://orcid.org/0000-0002-5109-0267

REFERENCES


