Identifying Mechanisms of Drug Tolerance in EGFR Mutant Lung Cancer

Bomiao Hu
Yale University Graduate School of Arts and Sciences, hubomiao2008@gmail.com

Follow this and additional works at: https://elischolar.library.yale.edu/gsas_dissertations

Recommended Citation
https://elischolar.library.yale.edu/gsas_dissertations/62

This Dissertation is brought to you for free and open access by EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Yale Graduate School of Arts and Sciences Dissertations by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.
Abstract

Identifying Mechanisms of Drug Tolerance in EGFR Mutant Lung Cancer

Bomiao Hu

2021

The majority of EGFR mutant lung adenocarcinomas respond well to EGFR tyrosine kinase inhibitors (TKIs). However, most of these responses are partial with drug tolerant residual disease remaining even at the time of maximal response. This residual disease serves as a reservoir for the emergence of acquired resistance and tumor relapse, which inevitably occurs in patients treated with TKIs. It is thus critical to understand the biology of residual tumor cells and find out the mechanisms that underlie drug tolerance. Knowledge of such mechanisms could lead to the identification of potential strategies to forestall the emergence of drug resistance.

Studies of residual disease have been hampered by the difficulty of studying these persister cells in patient specimens and most studies to date have relied on analysis of established cell lines in culture. To investigate the cellular and molecular properties of residual tumor cells in vivo, we leveraged patient-derived models of EGFR mutant lung cancer. Three EGFR mutant PDXs were treated with the third-generation TKI osimertinib. Tumors regressed in all cases but measurable residual tumor remained in 2 out of the 3 PDXs even after 6 weeks of osimertinib treatment. Whole-exome sequencing (WES) of the untreated PDXs compared to the residual tumors revealed an unchanged mutational landscape between the samples, indicating that genetic mechanisms did not account for drug tolerance. Bulk RNA-sequencing, however, demonstrated extensive transcriptional changes between the untreated PDX and residual disease. In one of the PDXs, we identified upregulation of the neuroendocrine lineage transcription factor ASCL1 in residual disease compared to untreated tumors. Using single-cell
RNA-sequencing we found a pre-existing ASCL1 hi tumor cell population in untreated tumors, suggesting that these cells which possessed drug tolerant properties were selected for during drug treatment. Depending on the cell line examined, expression of ASCL1 in human mutant EGFR lung cancer cell lines gave rise to persister clones following osimertinib treatment. This result demonstrated functionally that ASCL1 could lead to TKI tolerance and that whether it did this depended on the cellular context. Further gene expression profiling of ASCL1-transfected cell lines identified an ASCL1-induced epidermal-to-mesenchymal transition (EMT) signature that potentially resulted in tolerance to osimertinib. Our studies provide insights into the role of the neuroendocrine factor ASCL1 as a potential driver of drug tolerance in mutant EGFR lung cancer and ongoing work is focused on identifying the mechanisms underlying the cellular context specificity of ASCL1-mediated EMT.

In addition to PDX models, we also performed complementary studies on TKI tolerance in a genetically-engineered mouse model (GEMM) of mutant EGFR-driven lung adenocarcinoma. To identify and isolate drug tolerant persisters, we developed a transgenic mouse model in which mutant EGFR-expressing lung epithelial cells were labelled with a fluorescent marker, mKate. Using this model, we found that tumor-bearing mice responded almost completely to osimertinib and the number of mKate+ cells was decreased in response to TKI treatment and plateaued after 4 weeks of treatment. Targeted deep sequencing of the mutant EGFR transgene did not show enrichment of any resistance conferring mutations in mKate+ cells following TKI treatment, indicating that on-target EGFR mutations did not contribute to TKI tolerance in the GEMM. In contrast, RNA sequencing of DTPs sorted from this model revealed deregulation of metabolic and developmental pathways that could play a role in drug tolerance.

In summary we have established and characterized state-of-the-art in vivo models to study drug tolerance and unveiled new potential mechanisms of TKI tolerance which serve as the foundation for future studies into this critical problem in cancer therapeutics.
Identifying Mechanisms of Drug Tolerance in *EGFR* Mutant Lung Cancer

A Dissertation

Presented to the Faculty of the Graduate School

Of

Yale University

In Candidacy for the Degree of

Doctor of Philosophy

By

Bomiao Hu

Dissertation Director: Katerina Politi, PhD

June 2021
© 2021 by Bomiao Hu
All rights reserved.
# TABLE OF CONTENTS

ABSTRACT ........................................................................................................................................ 1  

ACKNOWLEDGMENTS ................................................................................................................. viii  

CHAPTER 1 LITERATURE REVIEW ................................................................................................. 1  

Lung Cancer .................................................................................................................................. 1  

Statistics ......................................................................................................................................... 1  

Histological Classifications ............................................................................................................ 1  

Driver Mutations in Lung Adenocarcinoma .................................................................................... 2  

Treatments for Non-Small Cell Lung Cancer .................................................................................... 3  

Surgery ........................................................................................................................................... 3  

Chemotherapy ................................................................................................................................ 4  

Immunotherapy ............................................................................................................................... 4  

Targeted Therapy ............................................................................................................................ 5  

EGFR TKI Therapies in Mutant EGFR Lung Cancer ........................................................................ 7  

Tolerance to EGFR TKIs in Mutant EGFR Lung Adenocarcinoma .................................................. 9  

Incomplete Response to EGFR TKIs ............................................................................................... 9  

Current Knowledge about Drug Tolerance in Mutant EGFR Lung Cancer .................................... 10  

Pre-clinical Models of Mutant EGFR Lung Adenocarcinoma .......................................................... 11  

Cell Line ....................................................................................................................................... 11  

Patient-Derived Xenograft .............................................................................................................. 11  

Genetically-Engineered Mouse Model ............................................................................................ 12  

CHAPTER 2: MODELING RESIDUAL DISEASE IN PDX MODELS OF MUTANT EGFR LUNG CANCER ................................................................................................................................. 13  

SUMMARY ..................................................................................................................................... 14  

INTRODUCTION ............................................................................................................................. 14  

RESULTS ....................................................................................................................................... 16
Identification of drug-tolerant residual disease in PDXs of EGFR-driven lung cancer treated with the EGFR TKI osimertinib.

The clonal and mutational landscape of osimertinib-tolerant disease was similar to that of untreated tumors.

Persistent tumor cells exhibited dramatic changes in transcriptional profiles compared to untreated tumor cells.

Figure 2.1 PDXs of mutant EGFR lung cancer showed incomplete responses to osimertinib treatment.

Figure 2.2 Drug tolerant residual tumors exhibited low level of proliferation and apoptosis.

Figure 2.3 Drug tolerant residual disease showed minimal changes in mutational landscape but dramatic changes in gene expression profile compared to untreated tumors.

DISCUSSION

EXPERIMENTAL PROCEDURES

Figure S2.1 Mutant EGFR lung cancer PDXs were resistant to erlotinib.

Figure S2.2 Tumor cells were still viable after development of drug tolerance.

Figure S2.3 The changes in VAF after osimertinib treatment remained statistically insignificant for the majority of the non-synonymous mutations identified by WES.

Figure S2.4 Human tumor cells were major components of the PDX tumors.

Figure S2.5 Transcriptional features of drug tolerant residual cells from PDX models.

Table S2.1 Changes in the expression of selected lung alveolar signature genes in osimertinib-tolerant cells vs untreated cells from both PDXs.

CHAPTER 3: ASCL1 DRIVES TOLERANCE TO OSIMERTINIB IN EGFR MUTANT LUNG CANCER IN PERMISSIVE CELLULAR CONTEXTS.
An ASCL1-mediated transcriptional program was present in drug tolerant residual disease.

Rare ASCL1-high tumor cells pre-existing in untreated tumors were selected for upon osimertinib treatment.

ASCL1 triggered epithelial-to-mesenchymal transition (EMT) which was the likely mediator of drug tolerance.

ASCL1 was expressed and drove neuroendocrine phenotype in a subset of TKI-resistant EGFR mutant lung cancer with SCLC transformation.

Figure 3.1 An ASCL1 gene expression program was specifically activated in YLR102 residual tumor cells.

Figure 3.2 ASCL1\textsuperscript{hi} persister-like cells pre-existed in the YLR102 untreated tumor.

Figure 3.3 ASCL1 overexpression caused TKI tolerance in HCC827 cell line.

Figure 3.4 ASCL1 expression led to EMT phenotype in permissive cellular contexts.

Figure 3.5 TKI resistant EGFR-driven LUAD with SCLC transformation exhibited heterogeneous expression of subtype-defining transcriptional regulators of SCLC.
Figure S3.2 Human tumor cells were enriched for single-cell RNA-sequencing by FACS. ................................................................. 63
Figure S3.3 DTRC-like tumor cells pre-existed in untreated YLR074 PDX. ............ 64
Figure S3.4 ASCL1 overexpression did not affect osimertinib sensitivity in PC9 or H1975 cell lines. .......................................................................................... 65
Figure S3.5 ASCL1 expression led to EMT phenotype in permissive cellular contexts. .................................................................................................................. 66
Table S3.1 Top 5 upregulated pathways in HCC827 and PC9 cells after ASCL1 overexpression and osimertinib treatment as predicted by MetaCore software........ 68
Table S3.2 Expression level of NE markers and subtype-defining transcriptional regulators of SCLC in pre- and post-TKI EGFR-driven lung cancer. .................. 69

CHAPTER 4: MODELING RESIDUAL DISEASE IN GENETICALLY-ENGINEERED MOUSE MODEL OF MUTANT EGFR LUNG CANCER .......................................................... 70
SUMMARY ........................................................................................................... 71
INTRODUCTION .................................................................................................... 71
RESULTS ............................................................................................................... 73
Identification and isolation of drug tolerant residual cells in a GEMM of EGFR mutant lung cancer. ........................................................................................................... 73
Transcriptional profiling of the tumor cells isolated from the pre- and post-treatment GEMM revealed pathways deregulated in DTRCs. ................................................................ 74
Figure 4.1 Tumor-bearing transgenic mice of mutant EGFR LUAD showed complete responses to osimertinib. ................................................................. 77
Figure 4.2 Isolation of mKate+ cells from the GEMM. .............................................. 78
Figure 4.3 Transcriptomic profiling of mKate+ cells isolated from the GEMM. ....... 79
DISCUSSION ......................................................................................................... 80
EXPERIMENTAL PROCEDURES ........................................................................ 82
Figure S4.1 Tumor-bearing transgenic mice of mutant \textit{EGFR} LUAD showed partial responses to erlotinib. ..................................................................................................................85

Figure S4.2 On-target resistance conferring \textit{EGFR} mutations were not enriched in mKate+ cells isolated from TKI-treated GEMM. ........................................................................................................86

Table S4.1 Expression of lung epithelial markers in mKate+ cells isolated from healthy, untreated and osimertinib-treated mice............................................................................................................87

CLOSING REMARKS ........................................................................................................................................................................88

BIBLIOGRAPHY .........................................................................................................................................................................................93
Acknowledgments

During the entire period of my graduate studies, I have received endless assistance and support from people within the lab, within the Pathology Department and within the Yale community, as well as from my family members. First of all, I want to thank my thesis advisor, Dr. Katie Politi, for her excellent mentorship. For my almost six years in the lab as a PhD student, she has provided me with the best-in-class opportunity for scientific training and guided me through a challenging thesis project. She always encourages me to think independently and to explore all the possibilities to find out the solutions to difficult scientific questions, which elevates both my experimental and analytic skills. I also appreciate that she could spend time with me in our regular individual meetings, which gave me the chance to keep my research progress up-to-date and address the issues that I encountered in a timely manner. Besides research, she cares about my connection with the scientific community in general and encourages me to participate and present at conferences. She has been the role model for me as a scientist and I feel very fortunate to be her student and learn from her how to conduct research.

I am also very grateful to my thesis committee members: Dr. Rong Fan, Dr. David Stern and Dr. Qin Yan. Their professional expertise in different research fields has broadened my knowledge base and added more possibilities to the directions of my research. They are willing to share the resources in their labs, from which I learned a variety of techniques and got to know people with different backgrounds, promoting a friendly and collaborative environment. In my committee meetings and seminars where I gave talks, they were always ready to point out the mistakes and statements that lacked sufficient thinking. Their attitudes to science fostered my ability to think more critically and perform experiments more carefully.

Next, I want to thank my colleagues in Politi Lab. We have spent so much time together collaborating and troubleshooting. We are willing to share our knowledge and techniques when other lab members are in need. I want to show special thanks to our lab manager Mary Ann
Melnick for her pilot experiments to start this whole project and her crucial role in keeping the lab organized; Anna Wurtz for coordinating tissue collection and archiving; Dr. Fernando de Miguel from whom I learned different ways to analyze gene expression data; and Camila Robles-Oteiza who taught me how to stain tumor cell suspensions for FACS. I would like to thank the postgrads in the lab for maintaining our mouse colonies. It is a great pleasure to work with all the current and previous members of the Politi Lab.

I also want to thank the bioinformatic team: Dr. Jungmin Choi, Dr. Zongzhi Liu, Dr. Francesc Lopez-Giraldez and Dr. Dejian Zhao. They have been of great help to me with processing and analyzing WES, RNA-seq and scRNA-seq data. Their expertise is critical for most of the discoveries described in this project.

I feel very grateful to my parents for their understanding and willingness to let me choose the major that I am interested in, study abroad and pursue my own career development. Family is always the safe haven for me since I left home eleven years ago for college. The support from my family members is always a booster in my long journey of research and study. I wish that I could spend more time with them when I have the chance.

Last but not least, I would like to extend my sincere gratitude to all the patients who are willing to share their tumor specimens and experimental animals which have been dedicated to disease modeling and pre-clinical drug testing. Without their contribution, our knowledge about cancer biology would never go as far as today.

My graduate studies would never be possible without the generous support from Yale GSAS, Department of Pathology, PDMC grant awarded by NCI (5U01CA235747-02) and The Lo Graduate Fellowship for Excellence in Stem Cell Research awarded by Yale Stem Cell Center.
CHAPTER 1: LITERATURE REVIEW

Lung Cancer

Statistics

Lung cancer is currently one of the most commonly diagnosed cancers and the leading cause of cancer-related death in the US [1]. In 2021, there are estimated to be more than 230K new cases and 130K deaths from lung cancer in the US. The incidence of lung cancer has been steadily reduced by 2% per year since mid-2000s and the death rates have been decreased by 54% in men and 30% in women compared to their respective historical highest rates, which is partially due to public efforts to decrease cigarette smoking and improved screening and treatment options including targeted therapy [1, 2]. However, despite all these improvements, the overall 5-year survival rate for lung cancer is only 21%, much below the average rates for all cancer types taken together [1]. Most of the lung cancer cases (83%) are diagnosed at an advanced stage, with regional or distant metastasis, which negatively impacts treatment efficacy and prognosis.

Histological Classifications

Lung cancer is collection of multiple diseases. Based on the histological features lung cancer is categorized into small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC) and the latter can be further divided into three groups: lung adenocarcinoma (LUAD), lung squamous cell carcinoma and large cell lung carcinoma [3]. The incidence rates for each subtype are illustrated in Figure 1.1. Lung adenocarcinoma is by far the most common histological subtype of lung cancer, accounting for 40% of all lung cancer cases, usually arising in the lung periphery close to terminal bronchioles and alveoli and shows features of glandular differentiation. On the other hand, squamous carcinoma is localized more centrally in the lungs near the bronchus and characterized by keratinization [3]. Small cell lung cancer is distinct from
the other cancer types in that it is thought to arise from neuroendocrine (NE) cells of the lungs and show positivity for NE marker staining such as chromogranin A and synaptophysin.

![Figure 1.1 Histological classification of lung cancer, adapted from [3]](image)

**Driver Mutations in Lung Adenocarcinoma**

Studies of the genetics and biology of LUAD have brought tremendous insights into the understanding of this disease. Various driver mutations have been identified and the functions of these mutated oncogenes have been elucidated and translated into improvements of clinical practice [4]. The most commonly mutated oncogene in LUAD is KRAS, with mutations occurring in 30% of all LUAD cases (Figure 1.2). These mutations lock KRAS in a GTP-bound active state and prevent GTP hydrolysis [5]. Therefore, the mutant form of KRAS is able to transduce constitutive downstream growth signals, leading to uncontrollable cell proliferation. The second most common oncogenic mutations occur in EGFR and the mutated forms are found in 10-15% of LUADs in Caucasians and in even higher incident rates in East Asians (Figure 1.2). The majority of these LUAD-associated EGFR mutations fall into two different locations in exons encoding the kinase domain of EGFR: one is single nucleotide substitution in exon 21, leading to a leucine-to-arginine transition at codon 858 (L858R); the other is a set of short in-frame deletions in exon 19, resulting in a loss of 4-6 amino acids in the N-lobe of EGFR kinase domain (del19).
Structural analyses of these EGFR mutants have determined the biochemical consequences of these alterations. The L858R mutation resides in the activation loop of EGFR kinase domain and this amino acid substitution destabilizes its inactive conformation [6] whereas exon 19 deletion mutations are in proximity to the αC helix and in the truncated mutants αC helix is pulled towards the catalytic site, thus stabilizing the active conformation [7]. Both forms of the EGFR mutations result in constitutive activation of the EGFR kinase and uncontrolled transduction of growth signals, which is similar to the consequences of KRAS mutations identified in LUAD.

Beside KRAS and EGFR mutations, other less common oncogenic driver mutations have been gradually identified in recent years to fill in the mutational spectrum of LUAD (Figure 1.2), including BRAF V600E mutation, ALK translocation, ROS1 rearrangement etc. It is worth noting that all of these mutated oncogenes are centered at RTK-MAPK signaling pathways, highlighting the importance of these pathways in the growth and survival of LUAD. Many of these oncogenic drivers are druggable targets and targeted therapies have been approved or in clinical trials [8-11].

![Figure 1.2 Driver mutations in LUAD, adapted from [4]](image)

Treatments for Non-Small Cell Lung Cancer

*Surgery*
Surgery can be a good option for patients with overall good health and early-stage NSCLC [12]. Since the lesions are usually local and cancer cells are less aggressive, removing the tumor area all together provides the best chance for cure. For diseases in advanced stages, lesions cannot be completely removed through surgery and therapeutic agents are thus necessary for controlling the disease progression. In rare cases where tumor cells form single distant metastasis (for example, in the brain) and tumors in the lungs are manageable, surgery can be used to resect metastatic lesions, followed by treatment of cancers in primary sites. Patients usually receive adjuvant chemotherapy to eliminate any tumor cells left over and to reduce recurrence. Neoadjuvant therapy is also used under certain scenario to reduce tumor burden and improve the outcome of subsequent surgical procedures.

**Chemotherapy**

Chemotherapy is a common therapeutic option for NSCLC in general and can be used either alone or in combination with other types of therapies. It historically has been one of the major treatment options for patients where no driver mutations are identified or targeted therapies are not available (for example, cancers driven by mutant \( KRAS \)) and now is frequently used in combination with immunotherapies [12].

Chemotherapy drugs like platinum-based drugs and taxanes are cytotoxic reagents that interfere with DNA synthesis or microtubule functions, both of which are critical biological processes for rapidly-proliferating cells including cancer cells. Due to the lack of specificity against malignant cells, chemotherapy agents can cause various side effects on multiple organs and systems that require constant cell proliferation and renewal, such as the digestive and hematopoietic systems, limiting the dose and duration of such treatments.

**Immunotherapy**
In recent years, immune checkpoint inhibitors (ICIs) have become one of the main classes of drugs used to treat NSCLC. Tumor microenvironment is usually immunosuppressive and the normal cytotoxic functions of T cells are silenced by immune checkpoint proteins mostly expressed on surface of tumor cells. The idea of this drug family is to activate immune system to target cancer cells by blocking the inhibitory interactions between checkpoint receptors and ligands expressed in immune cells and tumor cells respectively. The major checkpoint pathways involved in the context of cancer are mediated by PD1 and CTLA4, which are the drug targets for immune checkpoint inhibitors [13].

Three classes of ICIs targeting PD1, PD-L1 or CTLA4 have been approved to treat patients with advanced NSCLC. Despite the fact that a proportion of patients exhibit durable responses to ICI treatment and improved overall survival as compared with chemotherapy, a large number of patients do not show objective response at all and the determinants of response to immunotherapy in NSCLC are still under investigation. Some retrospective studies suggest that patients with higher tumor mutational burden might benefit more from immunotherapy as nonsynonymous somatic mutations tend to produce more neoantigens that facilitate recognition of tumor cells by immune system [14, 15].

Similar to other therapies for NSCLC, patient develop acquired resistance to ICIs as well, which may arise through different mechanisms [13]. One prominent mechanism of resistance to immunotherapy is impaired antigen presentation in tumor cells, mainly through loss of functional β2-microglobulin, which results in defective MHC I molecules [16, 17]. Besides antigen presentation deficiency, other tumor cell intrinsic signaling pathways and dysfunction of T-cells are also implicated in driving ICI resistance [18].

**Targeted Therapy**

During the past few decades, genetic profiling of patient tumors has identified subgroups of patients with genetic abnormalities that create vulnerabilities to drugs targeting certain gene
products [8-11, 19, 20]. These genetic alterations occur in certain oncogenes and cancers driven by these mutants become dependent on the activity of the oncogenes. These oncogenic drivers thus serve as drug targets. Unlike chemotherapy, the action of the targeted therapy is usually more restricted to cancer cells. Therefore, targeted therapy usually shows fewer side effects and higher efficacy compared to traditional chemotherapy when given systemically to patients whose tumors harbor the relevant drug targets [20].

Many oncogenic drivers identified in LUAD, the most common subtype of NSCLC, are druggable (Figure 1.2). Besides *EGFR*, which stands out as the most frequently mutated targetable driver oncogene, other oncogenic drivers including *BRAF*, *ALK*, *ROS1* etc harbor activating mutations as well, which render the cancer cells dependent on the constitutive activity of these oncogenic mutants. Indeed, clinical trials for the kinase inhibitors hitting these targets have shown promising outcomes in selected patient populations [8-11, 19, 20] and a number of targeted drugs have been approved as first-line therapy for LUAD patients harboring these activating mutations.

Nonetheless, as with all the other cancer treatments, acquired resistance to targeted therapy is inevitable in lung cancer patients [21]. Mechanisms of resistance generally fall into two categories: on-target and off-target. On-target resistance is caused by secondary mutations in the original drug targets. The resistance conferring mutations are able to reduce the efficiency of drug binding and lead to reactivation of the oncogenic drivers through changing the affinity of the kinase for ATP or creating stearic hindrance [22, 23]. Examples of on-target mechanisms include the *EGFR T790M* mutation found in patients receiving first-gen EGFR TKIs and the *EGFR C797S* found in patients treated with third-gen EGFR TKI osimertinib [24-27]. Off-target resistance may result from activation of alternative signaling pathways that bypass the primary drug targets, such as *MET* and *ERBB2* amplification identified in TKI-treated *EGFR* mutant lung cancer patients [21, 28], or transition to another cell lineage that no longer depends on the
original oncogenes for survival, such as SCLC transformation and epidermal-to-mesenchymal transition (EMT) driving acquired resistance to EGFR TKIs [29].

**EGFR TKI Therapies in Mutant **EGFR** Lung Cancer**

Since LUADs driven by **EGFR** mutants are addicted to the constitutive signaling activity of EGFR pathway, lung tumor cells harboring **EGFR** activating mutations become sensitive to EGFR TKIs, which lead to clinical trials and official approval of EGFR TKIs as first-line therapy for **EGFR**-driven lung cancer [19, 20]. The schema shown in Figure 1.3A indicates the EGFR TKI therapies that a patient would receive after being diagnosed with mutant **EGFR** lung cancer. Initially, these patients have been treated with 1st or 2nd gen TKIs for about a year until they acquired resistance to these drugs. The mechanisms of resistance to 1st and 2nd gen TKIs were well studied (Figure 1.3B): the most frequent mechanism is secondary gatekeeper mutation T790M in active center of EGFR kinase domain, which results in an increase in the binding affinity for natural substrate ATP, outcompeting the drugs from their binding pocket [22]. Other less common resistant mechanisms include activation of bypass signaling molecules in other RTK pathways such as **MET/ERBB2** gene amplification and PI3K activating mutations, as well as histological transformation and phenotypic transition such as SCLC transformation and EMT [30].

To overcome T790M-mediated acquired resistance, 3rd gen TKI osimertinib was designed to specifically target and bind irreversibly to EGFR mutants [31, 32]. T790M+ patients with acquired resistance to 1st or 2nd gen TKIs will benefit from this drug for about another year before they become resistant again. The resistant mechanisms for osimertinib in 2nd line therapy have been gradually elucidated in the past several years [33, 34]. The general types of resistant mechanisms are very similar to what has been observed in patients treated with 1st and 2nd gen TKIs, although the distribution of specific mechanisms is different between the two scenarios (Figure 1.3C). On-target resistance conferring mutations were also identified in patients
receiving osimertinib: C797 and L718 were the most frequently mutated amino acid residues in osimertinib resistant tumors. C797S/G replaces the residue that osimertinib covalently binds to, thus dramatically reducing the efficiency of kinase inhibition [26, 27] whereas L718V/Q leads to bulkier side chains within the catalytic site that create steric hindrance for osimertinib [23].

Figure 1.3 EGFR TKI therapies approved for mutant EGFR lung cancer and mechanisms of acquired resistance to EGFR TKIs. (A) Summary of clinical use of EGFR TKIs in EGFR mutant lung cancer. (B-C) Mechanisms of acquired resistance to (B) 1st/2nd gen TKIs [30] and (C) 2nd line osimertinib [33,34].

Osimertinib has also been approved to be used in first-line therapy [19, 35]. Patients receiving first-line osimertinib show longer progression free survival, compared to 1st or 2nd gen TKIs (Figure 1.3A). However, they still inevitably develop acquired resistance. The resistant mechanisms for osimertinib in first-line setting have just started to be investigated. An early report on mechanisms of first-line osimertinib resistance in phase III FLAURA study suggest that MET amplification and EGFR C797S mutation are most frequently seen in patients with progressive disease as revealed by next-gen sequencing of circulating plasma DNA, although what causes acquired resistance in the majority of patients (about 70%) remain unknown [36].
More extensive research in clinical specimens as well as pre-clinical models is required to elucidate the complex spectrum of resistant mechanisms and develop strategies to overcome drug resistance.

**Tolerance to EGFR TKIs in Mutant EGFR Lung Adenocarcinoma**

*Incomplete Response to EGFR TKIs*

Lung adenocarcinomas driven by EGFR activating mutations become dependent on the constitutive activity of EGFR and therefore patients harboring these mutations show greater responses to EGFR tyrosine kinase inhibitors (TKIs) [37-39]. Three generations of EGFR TKIs have been developed so far and five targeted small molecule reagents have been approved to treat mutant EGFR lung cancer as first-line therapy. Third-gen TKI osimertinib, in particular, was designed to overcome secondary drug resistant mutation T790M and specifically target mutant EGFR, which spares wildtype EGFR and thus reduce potential side effects [31].

Despite the higher response rates and improved progression-free survival compared to traditional chemotherapy, EGFR TKIs are still not able to cause complete tumor regression in most cases, leaving a significant amount of tumor tissues even at the time point of maximal response (Figure 1.4) [19]. These drug tolerant residual tumor cells that survive TKI treatment serve as reservoir for relapse tumors, which is inevitable in patients receiving TKI therapies and is the major reason for treatment failure. In this regard, understanding the biology of drug tolerance may lead to identification of potential strategies to combat incomplete response and delay acquired resistance.

![Figure 1.4 Incomplete response to EGFR TKIs. Scan courtesy of Anne Chiang, MD, PhD](image)
**Current Knowledge about Drug Tolerance in Mutant EGFR Lung Cancer**

There has been extensive research on the persistent cancer cells in pre-clinical models of multiple types of cancer and various mechanisms of drug tolerance have been identified so far [40, 41].

Mutant *EGFR* lung cancer, in particular, has gained great research interest since drug tolerance to EGFR targeted therapy is prevalent in patients and always leads to acquired drug resistance. Studies of drug-tolerant persisters (DTPs) in established human *EGFR* mutant lung cancer cell lines have uncovered several pathways that contribute to drug tolerance including the NF-kB pathway and activation of NOTCH3 signaling through the non-canonical Wnt pathway [42, 43]. Additionally, the epigenetic regulator KDM5A was found to lead to a drug tolerant state that can be reversed upon TKI withdrawal [44] and upregulation of the anti-apoptotic factor BCL2 was shown to promote survival leading to TKI-tolerant residual cells [45]. More recent findings suggest that activation of Aurora kinase A was able to promote cell survival through mitigation of apoptotic program induced by EGFR TKIs [46]. Studies of the chromatin structures in DTPs revealed a heterochromatin state mediated by repressive epigenetic modifications, which facilitated cell survival through suppression of stress-induced *LINE-1* expression [47].

Despite these findings mostly from *in vitro* studies, knowledge of which pathways are responsible for drug tolerance *in vivo* in patients are still limited. Indeed, studying residual disease has been particularly challenging in lung cancer, largely due to the difficulty of acquiring patient specimens at the time of maximal response. A recent study on single-cell analysis of TKI-treated patient specimen suggest that residual tumor cells exhibited an increased expression of healthy lung alveolar gene signature, shedding light on a potential transition between cell states upon development of drug tolerance [48]. In order to gain deeper insights into the mechanisms of drug tolerance in *EGFR* mutant lung cancer using patient specimens,
we leveraged patient-derived models to investigate the cellular and molecular properties of the
cancer cells that persist despite dramatic tumor regression induced by EGFR TKI treatment,
and to identify potential mechanisms of drug tolerance.

Pre-clinical Models of Mutant EGFR Lung Adenocarcinoma

Cell Line

Cell lines used in mutant EGFR lung cancer research are derived from patient tumors
harboring EGFR activating mutations (Table 1.1) [49] and play important roles in studying
biology and mechanisms of resistance to EGFR TKIs in vitro [50]. In my thesis research, I will
mainly focus on three human cell lines: PC9 (delE746-A750), H1975 (L858R+T790M) and
HCC827 (delE746-A750). All the three lines show sensitivity to third-gen EGFR TKI osimertinib.

<table>
<thead>
<tr>
<th>Cell Line</th>
<th>EGFR mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC9</td>
<td>delE746-A750</td>
</tr>
<tr>
<td>H1975</td>
<td>L858R + T790M</td>
</tr>
<tr>
<td>HCC827</td>
<td>delE746-A750</td>
</tr>
<tr>
<td>H3255</td>
<td>L858R</td>
</tr>
<tr>
<td>HCC4006</td>
<td>delL747-A750&gt;P</td>
</tr>
</tbody>
</table>

Table 1.1 List of selected human mutant EGFR lung cancer cell lines

Patient-Derived Xenograft

Similar to human cell lines, patient-derived xenografts (PDXs) also originate from patient
specimens but are transplanted into immune-deficient mouse to have the tumor cells grow in vivo. PDX models preserve many biological features of the primary patient samples, including histology, drug sensitivity and diversity and heterogeneity of patient tumors [51, 52]. Such models provide an opportunity to directly study patient tumor cells in vivo. Through co-analysis of patient specimens and PDXs, we are able to obtain more clinically relevant information that
eventually improves the understanding of the biology of mutant EGFR lung cancer and
treatment outcomes.

Despite the lack of a fully functional immune system in the murine host of PDX models, they
are largely used to investigate tumor cell intrinsic mechanisms of drug tolerance and resistance.

**Genetically-Engineered Mouse Model**

In genetically-engineered mouse model of mutant *EGFR* lung cancer, human mutant *EGFR*
is transduced into mouse genome and its expression is induced specifically in mouse lung
epithelial cells to transform them into adenocarcinoma cells [53]. The mutant *EGFR*-driven
mouse lung tumors retain the same TKI sensitivity as in human lung cancer patients. As the
tumors are developed in immune-competent mice, transgenic mouse models allow for studies
on cancer biology and drug tolerant mechanisms in the context of fully functional tumor
microenvironment. In addition, mutant *EGFR* can be crossed with other transgenic alleles to
investigate the roles of certain genes in tumor progression and drug resistance or add some
features to the tumor cells (for example, lineage tracing by fluorescent marker).

In my thesis research, I will use a modified GEMM of mutant *EGFR* lung cancer as well as
PDX models to perform complementary studies on the mechanisms of TKI tolerance.
CHAPTER 2: MODELING RESIDUAL DISEASE IN PDX MODELS OF MUTANT EGFR LUNG CANCER.

Bomiao Hu¹, Fernando J. de Miguel², Zongzhi Liu¹, Lok-Hei Chan¹, Mary Ann Melnick², Anna Arnal Estape¹,²,³, Zenta Walther¹, Anna Wurtz², Rong Fan⁴, Qin Yan¹,², Robert Homer¹,², Don Nguyen¹,², Katerina Politi¹,²,³

¹Department of Pathology, ²Yale Cancer Center and ³Department of Medicine (Section of Medical Oncology), Yale School of Medicine, New Haven, Connecticut; ⁴Department of Biomedical Engineering, Yale School of Engineering & Applied Science, New Haven, Connecticut

# To whom correspondence should be addressed:
Katerina Politi, PhD
Departments of Pathology and Internal Medicine (Section of Medical Oncology)
Yale Cancer Center
Yale University School of Medicine
333 Cedar Street, SHM-I 234D
New Haven, CT 06510 USA
Office Tel: 203-737-5921
Lab Tel: 203-737-6215
Fax: 203-785-7531
Email: katerina.politi@yale.edu
SUMMARY

*EGFR* activating mutations are present in about 10-15% of lung adenocarcinomas and confer sensitivity to EGFR tyrosine kinase inhibitors. As a result of their efficacy in this patient population, EGFR TKIs are the first-line targeted therapy for LUAD patients with *EGFR* mutation positive tumors. Despite remarkable response rates to TKIs, tumor regression is not complete in most patients treated with EGFR TKIs and tumor cells persist despite treatment. Most prior studies have focused on understanding *in vitro* mechanisms of drug tolerance in human cell lines and the nature of the incomplete responses in patient specimens is still not fully understood. In order to learn more about the biology of residual cells in clinically relevant conditions, we leveraged patient-derived models to study drug tolerant residual disease. This chapter focuses on characterizing the PDX models and determining their capability to mirror incomplete response in patients, as well as gaining insights into the cellular and molecular properties of the PDX-derived drug-tolerant persisters. In a subset of PDXs that we tested, we observed a significant number of human mutant *EGFR* lung adenocarcinoma cells that survived osimertinib treatment. Subsequent cellular profiling of residual tumor cells indicated that, in these specific PDXs, they entered a quiescent state with cell cycle arrest and a low level of apoptosis. Through next-generation sequencing, we did not identify alterations in the mutational landscape of residual tumors compared to untreated tumors that could account for their persistence. However, we observed a dramatic change in the transcriptional profile between the two conditions. Thus, in summary, we show that PDXs can be used to model tolerance to EGFR TKIs.

INTRODUCTION

Mutations in the Epidermal Growth Factor Receptor (*EGFR*) gene are found in ~15% of lung adenocarcinomas in the US and at higher frequencies in East Asian populations [4]. Tumors harboring these mutations exhibit sensitivity to EGFR tyrosine kinase inhibitors (TKIs) [37-39]
and five TKIs (erlotinib, gefitinib, afatinib, dacomitinib and osimertinib) are approved for the first-line treatment of this disease. However, even with all of these agents and despite response rates of ~70-80%, tumors never regress completely and residual tumor remains [19, 20, 32]. These drug tolerant residual tumor cells then serve as reservoir for the subsequent emergence of acquired resistance, which inevitably develops in patients receiving EGFR TKI therapy, even with the TKI that has shown the greatest progression-free survival, osimertinib [26, 27, 33, 34]. Thus, it is critical for us to understand the mechanisms that underlie drug tolerance to develop strategies to limit the amount of residual disease and delay tumor relapse.

Studies of drug-tolerant persisters (DTPs) have been mostly focused on established human EGFR mutant lung cancer cell lines so far. Discoveries in vitro have uncovered various mechanisms of drug tolerance, including signaling pathways promoting cell survival such as the NF-κB pathway and NOTCH3 pathway [42, 43], repressive chromatin state through epigenetic regulation [47], upregulation of the anti-apoptotic factor BCL2 [45], activation of cell cycle-promoting factor Aurora kinase A [46] and metabolic shifts to fatty acid oxidation and enrichment of antioxidant program in cycling persisters [54]. However, mechanistic studies on TKI tolerance in vivo, especially in a clinically relevant setting, have just started to emerge. A recent study on single-cell analysis of EGFR TKI-treated patient specimens emphasized the role of both tumor cell intrinsic and microenvironmental programs in drug tolerant residual disease [48]. Despite the insights into drug tolerant mechanisms in human tumors, such study still has its limitation in that it is challenging to acquire sufficient patient specimens at the time of maximal response and the tumor cell abundance can vary between different specimens, which makes it difficult to analyze and interpret the data. One solution to addressing these challenges is to grow patient specimens as patient-derived xenografts (PDXs) in experimental animals like immunodeficient mice. PDX models are likely to preserve the biological features of patient tumors, especially tumor cell-intrinsic properties. The histological and pharmacological similarity between PDX and corresponding patient specimen makes it a good model to identify residual disease and
investigate drug tolerance *in vivo*. In this chapter, we characterized several mutant EGFR lung cancer PDXs to determine whether they can be used to study drug tolerance and leveraged these models to profile the cellular and molecular properties of the persistent cells and to identify potential mechanisms of drug tolerance.

**RESULTS**

*Identification of drug-tolerant residual disease in PDXs of EGFR-driven lung cancer treated with the EGFR TKI osimertinib.*

To determine whether we could observe and isolate residual disease grown *in vivo*, we studied three EGFR mutant lung adenocarcinoma Patient-Derived Xenografts (PDXs). All of the PDXs were grown from biopsies of tumors that had acquired resistance to first- or second-generation TKIs and all harbored an in-frame EGFR exon 19 deletion mutation and a secondary T790M mutation in *EGFR* known to confer resistance to 1st and 2nd generation TKIs (*Figure 2.1A-B*). Analysis of the TKI sensitivity of these PDXs revealed that they were all resistant to the first-gen TKI erlotinib (consistent with the clinical course in the patient) and sensitive to the third-gen TKI osimertinib (*Figure 2.1C-E and S2.1A-C*). To determine whether residual cancer cells were present in these PDXs following osimertinib exposure, we treated tumor-bearing mice with osimertinib for 4-6 weeks and monitored tumor size on a weekly basis. We found that, although YLR074 and YLR102 exhibited significant responses to osimertinib with an 85% decrease in tumor volume by the end of week 6 on drug, the response was incomplete and residual tumor remained. In contrast, YLR135 exhibited a complete response to osimertinib since its tumor volume shrank by over 98.5% after 4 weeks of treatment. These results indicate that different PDXs can exhibit variability in the depth of their response to osimertinib similar to that observed in patients (*Figure 2.1C-E*).

To determine whether residual cancer cells were present, we collected tissue that remained following treatment once the tumor volume had reached a stable plateau and the tumor was no
longer shrinking (after 6 weeks of osimertinib treatment for YLR074 and YLR102 and 4 weeks for YLR135). Histological analysis revealed that for all three of the PDXs residual lung adenocarcinoma cells were present although the abundance of these cells was lowest in YLR135 that had shown the greatest response to osimertinib (Figure 2.1F). In all cases, we observed an increase in the stromal compartment in the residual disease sample. We confirmed the presence of residual tumor cells following osimertinib treatment, by performing immunohistochemical (IHC) staining for Nkx2-1, a lung epithelial cell marker (Figure 2.1G). In addition, we also used an antibody that specifically recognizes the EGFR exon 19 DEL mutant [55] present in YLR074 and YLR102 to further positively identify residual tumor cells (Figure 2.1H). These experiments revealed that PDXs can be used to model and identify drug-tolerant disease. Considering the abundance of residual cells in YLR074 and YLR102, we used these PDXs for subsequent studies.

Given that activity of the EGFR pathway is critical for tumor cell proliferation and survival in EGFR-driven lung cancer [28, 56], we determined the levels of proliferation and apoptosis in the two PDXs using IHC staining for ki-67 and cleaved caspase-3, respectively. We observed a dramatic reduction in ki-67+ nuclei in both YLR074 and YLR102 shortly following osimertinib treatment, which remained at a low level in residual tumors (Figure 2.2A-B), indicating that cell cycle progression is inhibited in residual tumor cells. Cleaved caspase-3 staining was induced at the beginning of TKI treatment and then returned to low baseline levels in residual disease (Figure 2.2C-D), suggesting that apoptosis triggered by EGFR inhibition subsided in drug tolerant tumor cells. Additionally, Western blotting of lysates obtained from untreated and residual tumors from YLR074 and YLR102 PDXs revealed effective suppression of EGFR phosphorylation indicating that tumor cell persistence was not due to incomplete activity of osimertinib in these tumor cells (Figure 2.2E-F).

To investigate whether the residual state was reversible, we withdrew osimertinib in 2 mice for each PDX line after 4 weeks (for YLR102) or 6 weeks (for YLR074) of osimertinib treatment.
at the time when residual disease was present (Figure S2.2A). Following osimertinib withdrawal, the tumors grew out in the absence of osimertinib with a 3 to 10-fold increase in tumor volume after 2-3 weeks off drug (Figure S2.2A-C) and the relapsed tumors retained histological features characteristic of each PDX line (Figure S2.2D). Relapsed tumors continued to produce the EGFR DEL19 mutant as determined by IHC (Figure S2.2D).

Together these results indicate that drug-tolerant residual tumor cells are characterized by reduced proliferation and low levels of apoptosis, consistent with previous studies in human lung cancer cell lines [44, 46, 57].

The clonal and mutational landscape of osimertinib-tolerant disease was similar to that of untreated tumors.

To investigate the mechanisms that account for the presence of residual disease in the PDXs, we first examined whether resistance-conferring mutations were present in the residual tumor cells compared to the pre-treatment specimens. Resistance-conferring mutations may pre-exist at a low frequency in treatment-naïve tumors and become clonally dominant during the process of drug treatment. Supporting this possible scenario, in the context of mutant EGFR lung cancers treated with first-gen EGFR TKIs, in some cases, the T790M mutation has been observed in pre-treatment patient tumor specimens and parental cell lines [50, 58-60]. To determine whether potential resistance-conferring mutations could account for drug tolerance, we performed whole-exome sequencing (WES) on untreated and residual tumors and compared sequencing results between the two conditions for each PDX line. We identified 242 non-synonymous mutations in YLR074 and 194 non-synonymous mutations in YLR102 through this analysis. All of the mutations were present in both untreated and residual tumor cells, without loss or gain of any mutation following 6 weeks of osimertinib treatment.

We also compared the variant allele frequencies (VAF) for the mutations in untreated and residual disease samples and found that these were largely unchanged after osimertinib
treatment (Figure 2.3A-B) suggesting the absence of selection of specific mutations by drug treatment. Consistent with this, pyrosequencing to confirm the allele frequency of the two disease-related EGFR mutations (the exon19 deletion and T790M mutations) revealed that the abundance of the EGFR mutations remained constant between untreated and residual disease and that the VAFs were comparable with WES (Figure 2.3A-C). To add statistical power to our observations, we performed a Fisher’s Exact Test to compare the VAF for every non-synonymous mutation we identified in the two PDXs through WES. The majority of the mutations were not statistically significantly different (Figure S2.3) and for the small number of mutations that passed the test (7 in YLR074 and 11 in YLR102), we went on to determine whether these mutations affected the clonal architecture using the clonal analysis program PyClone [61]. As predicted by this program, the mutations were classified into different clusters based on the relative abundance of the mutations in the whole cell population and the cellular prevalence of each cluster did not change after osimertinib treatment in either of the PDXs, which suggests that development of drug tolerance did not alter the clonal architecture of the mutations in these mutant EGFR lung cancer PDXs (Figure 2.3D-E). The results of PyClone, together with the VAF analysis for the mutations identified in pre- and post-treatment PDXs, indicate that there is no specific genetic alteration being selected for or against during osimertinib treatment and the mutational landscape remains very similar between untreated tumor and residual disease.

Persistent tumor cells exhibited dramatic changes in transcriptional profiles compared to untreated tumor cells.

Having excluded genetic alterations as major contributors to the drug tolerant phenotype in our samples, we then turned to determine whether transcriptomic changes were present between the untreated and residual tumor using bulk RNA-sequencing. Alignment results showed that the majority of the sequencing reads were uniquely mapped to the human genome,
for all of the samples submitted for RNA-sequencing, indicating that samples were mostly composed of human tumor cells (Figure S2.4). Principal component analysis (PCA) of the data showed a clear demarcation between the PDXs from each patient, irrespective of treatment, indicating that features of the individual tumors dominate the transcriptional profiles. These data also revealed that YLR102 residual disease was transcriptionally very different from YLR102 untreated tumor while YLR074 untreated and residual disease had more similar transcriptional profiles (Figure 2.3F).

To identify pathways that were altered in the drug-tolerant disease in the PDXs, we performed pathway analysis using MetaCore and found deregulation of genes in development-associated pathways including Wnt- and Hedgehog-signaling pathways in both PDXs (Figure S2.5A), which are able to promote survival of tumor cells during drug treatment [62-65]. The WNT pathway, specifically, has been shown to be enriched in residual disease specimens from patients with EGFR mutant lung cancer after TKI therapy and this is likely related to the activation of cell renewal and injury repair program [48]. A Hypoxia-Inducible Factor-1alpha (HIF1α) regulated transcriptional program was the top downregulated pathway in both PDXs following osimertinib treatment, which is probably a consequence of tumor shrinkage and reduced hypoxia in tumor microenvironment due to increased access to oxygen (Figure S2.5A). This is in line with the low-oxygen environment identified within PDX tumors, especially when they reach a relatively big size [66, 67]. Consistent with the decreased Ki-67 staining in residual disease, another prominent feature with residual PDX tumors is negative enrichment of cell cycle pathways, indicating suppression of cell cycle progression and supporting the hypothesis that tumor cells can survive treatment-induced pressure by exiting the cell cycle and entering a dormant state [68] (Figure S2.5A-B).

Recent gene expression profiling of TKI-treated mutant EGFR lung cancer patient specimens revealed upregulation of a gene expression signature representative of normal alveolar cells [48]. In light of these findings, we examined the levels of expression of markers for
different lung epithelial cell types in our RNA-seq datasets [69]. Although the alveolar gene signature did not show significant enrichment in residual tumor cells in either of our PDX models (Figure S2.5C), the expression of specific genes in the list exhibited a constant increase in persister cells from both PDXs, including lung epithelial lineage factor Nkx2-1 and Aquaporin-4 (Table 2.1). Furthermore, both PDXs are characterized with upregulation of club cell marker SCGB3A2 and downregulation of goblet and mucous cell marker MUC5B (Figure S2.5D). The results of gene expression profiling for untreated and residual PDX tumors suggest that drug tolerant residual disease represent a transcriptional state that differ from pre-treatment tumors and might undergo a phenotypic change following TKI treatment.
Figure 2.1 PDXs of mutant \textit{EGFR} lung cancer showed incomplete responses to osimertinib treatment. (A) Experimental design. (B) Basic information of the PDXs used in this study. (C-E) Tumor growth curves showing normalized tumor size of each mouse bearing (C) YLR074, (D) YLR102 or (E) YLR135 PDX following osimertinib treatment. Three biological replicates were included for each PDX line. (F-H) Histological features of the PDXs before and after treatment, as indicated by (F) H&E staining and IHC staining for (G) Nkx2-1 and (H) \textit{EGFR} del19 mutant. Scale bar: 100μm.
Figure 2.2 Drug tolerant residual tumors exhibited low level of proliferation and apoptosis. (A-B) IHC staining for Ki-67 in untreated, short-term (3 days) treated and residual (6 weeks) tumors of YLR074 and YLR102. Quantification of Ki-67 staining is shown in (B). Scale bar: 100μm. (C-D) IHC staining for cleaved caspase-3 in untreated, short-term treated and residual tumors of YLR074 and YLR102. Quantification of cleaved caspase-3 staining is shown in (D). Scale bar: 50μm. Significance for IHC staining was determined by unpaired t-test between untreated and residual groups. Each group contains six different fields of view. n.s.: p>0.05; ****: p<0.0001. (E-F) Immunoblotting for phospho-EGFR and total EGFR in untreated and residual tumors in (E) YLR074 and (F) YLR102. Each lane represents protein lysates from an individual tumor.
Figure 2.3 Drug tolerant residual disease showed minimal changes in mutational landscape but dramatic changes in gene expression profile compared to untreated tumors. (A-B) Dot plots comparing variant allele frequency of non-synonymous mutations identified by WES between untreated and residual tumors in (A) YLR074 and (B) YLR102. Each dot represents one non-synonymous mutation and disease-related EGFR mutations are highlighted in black. The blue dashed line running diagonally indicates where VAF of a certain mutation in residual disease is equal to that in untreated tumor. (C) Pyrosequencing for EGFR del19 and T790M mutations in untreated and residual tumors of YLR074 and YLR102. Significance was determined by multiple t-test. n=2 biological replicates for untreated group and n=3 biological replicates for residual group. n.s.: p>0.05. (D-E) Changes in clonality of mutations identified in (D) YLR074 and (E) YLR102 after osimertinib treatment, as predicted by PyClone program. The mutations were separated into multiple clusters and the number of mutations in each cluster is indicated within the brackets. (F) PCA plot visualizing the difference in gene expression profiles between untreated and residual tumors for YLR074 and YLR102.
DISCUSSION

Incomplete response to drug treatment is almost inevitable in human cancer. For every therapy given to patients, whether it is radiotherapy, traditional chemotherapy or targeted therapy, tumor cells cannot be fully eradicated in most of the cases. During the clinical course, the tumor burden usually stays stable in the presence of treatment, followed by progression in either primary or metastatic sites upon acquired resistance. The biology of drug tolerant residual disease in mutant EGFR-driven lung cancer has mainly been learned from in vitro research on human cell lines and more recently from primary patient biopsies [40, 41, 48]. Appropriate models are required to gain deeper insights into the mechanisms of drug tolerance in human tumors. Patient specimens, depending on the sampling approach, are usually limited and contain only a small number of tumor cells. Expanding tumor cells in a robust system is thus critical to obtain sufficient materials for multitudinal profiling. We chose to grow patient tumor tissues as PDXs in immunodeficient mice given that this model can retain both biological and clinical traits of the original human samples [51, 52]. Indeed, PDX mouse models have been used to model drug tolerance and acquired resistance in melanoma and successfully mimic the treatment outcomes of targeted therapy in patients [70]. In the PDXs that we interrogated in this thesis research, they exhibited a histology of aggressive adenocarcinoma, harbored the driver mutations and maintained the TKI sensitivity, which are all characteristic of the patient tumors that they were derived from. Two out of the three models also showed incomplete responses to third-gen TKI osimertinib, providing opportunity to study drug tolerant residual disease in a more clinically-relevant setting.

Despite the similarity between PDXs and human tumors, as an in vivo model, PDXs are not perfect for detailed mechanistic studies such as genetic manipulation, CRISPR screening and subsequent functional experiments. To address the questions related to molecular mechanisms, ongoing work is focused on growing cancer cells as 2D or 3D cultures that can be genetically-modified and assayed more easily.
Cellular profiling of persistent tumor cells revealed that most of them entered a quiescent G0 stage as manifested by Ki-67 negativity and low level of apoptosis, which is consistent with cell cycle-arrested DTPs identified in human cell lines [44]. The non-proliferative cell state is able to protect malignant cells from cancer treatments, which preferentially target the fast-dividing cell population. This quiescent state shares a great extent of similarity in biological features with dormant tumor cells and in both scenarios cells are insensitive and refractory to treatment [68]. However, detailed molecular profiling suggests that DTP is likely a heterogeneous cell population. Recent single-cell analysis of persister cells from human EGFR mutant lung cancer cell lines revealed a DTP subpopulation that was not cell cycle arrested, which may contribute to the eventual tumor cell regrowth [54]. In line with the in vitro study, most up-to-date study on a new PDX sample in our cohort (data not shown) suggests that residual disease consisted of a large number of Ki-67+ cells, indicating the possibility that depending on the tumor intrinsic features drug tolerant cells can be cell cycle-active rather than quiescent.

Secondary drug resistant mutations in EGFR are major contributors to the mechanisms of acquired resistance to all the EGFR TKIs that have been approved so far. These resistant cells can arise either de novo or through selection of pre-existing mutation-positive subpopulation [50]. In order to understand the mutational landscape of residual tumor cells, which are prelude to relapse tumors, we performed WES to identify any resistance-conferring mutations that were enriched in osimertinib-tolerant cells compared to untreated tumors. We did not observe on-target EGFR mutations such as C797S and L718V/Q that are known to cause osimertinib resistance in either untreated or residual tumor cells [23, 26, 27], nor any activating mutations in KRAS that can bypass EGFR pathway [5, 23]. Furthermore, the allelic frequency and clonal architecture of all the non-synonymous mutations identified through WES remained quite similar between untreated cells and DTPs in the PDX models. The analyses of mutational status in the PDXs before and after treatment suggest that pre-existing drug resistant mutations may not contribute to the TKI tolerance in the PDX cohort that we examined. This is consistent with other
studies indicating that the transcriptional landscape is the major difference between untreated and residual disease.

Inter-tumoral heterogeneity is common in cancer patients. Tumors that are categorized in the same histological subtypes and harbor the same oncogenic driver mutations may still exhibit great extent of transcriptional difference due to genetic and epigenetic distinctions. The two PDXs profiled in this chapter showed clear difference in gene expression when not treated with TKI and osimertinib induced transcriptomic changes to different extent in different PDXs, as revealed by RNA-seq. How this transcriptional discrepancy between patients determines responses and tolerance to EGFR-targeted therapy is still not well understood. As discussed in the next chapter, this transcriptional heterogeneity is not limited to PDXs and can be extended to human lung cancer cell lines derived from different patients and the cellular contexts can affect the role of certain genes in driving TKI tolerance. To investigate drug tolerant mechanisms more comprehensively and understand the heterogeneity between patients, ongoing work is focused on characterizing additional patient-derived models and synthesizing information obtained from different sources.

**EXPERIMENTAL PROCEDURES**

*Establishing PDXs in immune-deficient mice.*

Immune-deficient NSG mice were purchased from The Jackson Laboratory. Mouse handling and *in vivo* experiments were performed following the protocol approved by Institutional Animal Care and Use Committee at Yale University.

Fresh or fresh-frozen tumor tissues were processed with tumor dissociation kit (Miltenyi Biotech, 130-095-929) to generate single cell suspension. Cells were spun down and resuspended in Matrigel (Corning, 356237), followed by subcutaneous injection into the flanks of NSG mice. Materials equal to 100mm³ tumor tissues were injected into each mouse. Tumors were let to grow freely until they reached desirable volume for further studies.
**Drug treatment in vivo.**

Tumor-bearing mice were treated with osimertinib at 25mg/kg for five consecutive days each week when their tumor size reached around 300mm³. The treatment lasted for 6 weeks before drug tolerant residual tumor was excised, unless stated otherwise. Tumor burden was measured by caliper once or twice per week.

Osimertinib (kindly provided by AstraZeneca) was dissolved in 0.5% methylcellulose (Sigma-Aldrich, M0512) at 5mg/ml, with overnight stirring.

**Immunohistochemistry staining.**

Tumor tissues were fixed in 4% paraformaldehyde (Electron Microscopy Services, 15714-S) and embedded in paraffin. FFPE tissue sections were deparaffinized in Clear-Rite 3 (Thermo Scientific, 6901) and then rehydrated in serially diluted ethanol before heat-induced epitope retrieval in antigen unmasking solution (Vector Laboratories, H-3300). The tissue sections were then treated with 3% hydrogen peroxide (J.T.Baker, 2186-01) to remove endogenous reactive oxygen species and blocked with goat serum (provided by VECTASTAIN Elite ABC-HRP kit, PK-6101), followed by incubation with primary antibody at 4 degree overnight. The sections were treated following the protocol in VECTASTAIN kit, stained with NovaRED substrate kit (Vector Laboratories, SK-4800) and counterstained with hematoxylin (Vector Laboratories, H-3401). Finally, sections were dehydrated and mounted in Permount (Fisher Scientific, SP15-100).

Primary antibodies used for IHC: Nkx2-1 (abcam, ab76013, 1:250), EGFRdel19 (CST, 2085, 1:200).

**Western blotting.**
Tissue powder or cells were lysed in RIPA buffer to extract proteins [39]. The protein concentration was quantified using DC protein assay (Bio-Rad, 5000116), followed by protein denaturation in 4X sample buffer (Bio-Rad, 1610747) supplied with 40mM DTT (Sigma-Aldrich, D9779). Proteins of equal amount for each sample were loaded onto a 4-20% SDS-PAGE gel (Bio-Rad, 4568093) and electrophoresis was performed at constant voltage until protein ladder (Bio-Rad, 1610375) was well-separated. Proteins were then transferred to nitrocellulose membrane (Bio-Rad, 1620112) and blocked in 5% non-fat dry milk (AmericanBio, AB10109-01000). Blots were incubated with primary antibody at 4 degree overnight with constant shaking. The membranes were then washed in TBS (Bio-Rad, 1706435) with 0.1% Tween-20 (Sigma-Aldrich, P7949), incubated with secondary antibody (CST, 7074, 1:2000) at room temperature for 1 hour and washed again in TBST. The blots were developed using SuperSignal West Pico (Thermo Scientific, 34580) or Femto (Thermo Scientific, 34096) chemiluminescent substrate and chemiluminescence was detected in ChemiDoc MP gel imaging system (Bio-Rad).

Primary antibodies used for western blotting: pEGFR (CST, 3777), EGFR (CST, 4267), EGFRdel19 (CST, 2085), Actin (Santa Cruz, sc47778 HRP, 1:2000). All the primary antibodies were diluted in 5% non-fat dry milk by 1:1000, unless stated otherwise.

**Whole exome sequencing.**

PDX tumors were harvested and flash-frozen in liquid nitrogen. Frozen tissues were ground into powder with mortar and pestle on top of dry ice. Genomic DNA was extracted from tissue powder using AllPrep DNA/RNA mini kit (Qiagen, 80204). Genomic DNA quality is determined by estimating the A260/A280 and A260/A230 ratios by nanodrop, both of which should be > 1.8, and by 1% agarose gel electrophoresis. High quality DNA will migrate as a single high molecular weight band. DNA samples were fragmented, ligated to adaptors and barcoded, followed by exome target capture with xGen Exome Panel (IDT). The libraries were sequenced.
on illumina HiSeq4000 sequencer at Yale Center for Genome Analysis (YCGA) using 100bp pair-end sequencing to generate data with 200X coverage per sample.

**Bulk RNA-sequencing.**

RNA was extracted from tissue powder or cell lines using AllPrep DNA/RNA mini kit. RNA integrity was determined by running an Agilent Bioanalyzer gel, which measures the ratio of the ribosomal peaks. Samples with RIN values of 7 or greater proceeded to library prep. RNA-seq libraries were constructed using KAPA mRNA HyperPrep Kit (Kapa Biosystems, KR1352) and sequenced on illumina HiSeq2500 sequencer using 75bp single-end sequencing to generate around 30M reads per sample.

**WES data analysis.**

WES raw data were processed and analyzed with similar methods as described before [17]. In brief, Xenome [71] was used to classify the origin of sequencing reads and filter out the reads originated from mouse. The remaining sequencing reads were aligned to the reference genome (hg19) using BWA-MEM [72] followed by GATK Best Practices workflow [73]. Somatic mutations were identified using MuTect2 [74] and Pisces [75] and in-house scripts were used to filter out potential false positives. Likely damaging variants were annotated including premature termination, canonical splice site, frameshift insertion/deletion, and missense mutation using Annovar [76]. All somatic mutations were manually confirmed by visualizing read plots.

**Estimation of tumor clonal population structure.**

PyClone [61] was used to infer the subclonality in tumor samples. The software was implemented to cluster somatic mutations incorporating the variant allele frequencies obtained by MuTect2, using the hierarchical Bayes statistical model and the cellular prevalence of
somatic mutations is estimated from a Dirichlet Process. Clusters containing at least 2 mutations and a mean of VAF greater than 5% were used.

**RNA-seq data analysis.**

RNA-seq raw data generated from PDXs were processed with BMX-seq pipeline as described before [77]. In brief, sequencing reads were trimmed of adaptor sequences using Trim Galore and mapped to a combined human (hg38) and mouse (mm10) genome using STAR [78] to separate the components of the two species and to a combined transcriptome (GENCODE v24 for human, vM10 for mouse) to add gene annotation. Then the uniquely mapped reads were counted per annotation using featureCounts [79]. Finally, the counting matrix went through DESeq2 package [80] to identify the genes that are differentially expressed between different conditions.

**Pathway enrichment analysis.**

DEG lists with log2FoldChange and p-value for individua genes were used as input for MetaCore software (https://portal.genego.com/). Enrichment analyses were performed only for genes with log2FC>1 or <-1 and p-value<0.05, unless stated otherwise.

**Gene Set Enrichment Analysis.**

Differentially expressed genes identified in bulk RNA-seq were sorted and ranked by their Wald statistic value as calculated by DESeq2 pipeline. Pre-ranked gene list was analyzed against proper gene signatures through the GSEA software 3.0 downloaded from GSEA website (https://www.gsea-msigdb.org/gsea/index.jsp) [81, 82]. Alveolar gene signature was obtained from previous studies [48]. Cell cycle-related gene signatures were downloaded from GSEA website.
Figure S2.1 Mutant EGFR lung cancer PDXs were resistant to erlotinib. (A-C) Tumor growth curves showing normalized tumor size of each mouse bearing (A) YLR074, (B) YLR102 or (C) YLR135 PDX following erlotinib treatment.
Figure S2.2 Tumor cells were still viable after development of drug tolerance. (A) Experimental schema of tumor relapse assay for YLR074 and YLR102. (B-C) Relapse tumor burden normalized to residual disease for (B) YLR074 and (C) YLR102 after 2 and 3 weeks off drug, respectively. (D) H&E staining and IHC staining for EGFR mutant on tissue sections from YLR074 and YLR102 relapse tumors. Scale bar: 100μm.
Figure S2.3 The changes in VAF after osimertinib treatment remained statistically insignificant for the majority of the non-synonymous mutations identified by WES. (A-B) Dot plots depicting VAF of the mutations in the two PDX lines, similar to what is shown in Figure 2.3A-B. The dots highlighted in red represent the mutations that pass the Fisher’s Exact Test (p-value<0.05).
Figure S2.4 Human tumor cells were major components of the PDX tumors. Bar plots showing the proportion of bulk RNA-seq reads mapped to human (hg) or mouse (mm) genome with or without assigned features for each single tumor sample of YLR074 or YLR102.
Figure S2.5 Transcriptional features of drug tolerant residual cells from PDX models. (A) Results of MetaCore pathway analysis showing the pathways predicted to be up- or down-regulated in residual disease in both PDXs. DEGs were filtered with threshold=1; p<0.05 (B) Enrichment plots for two cell cycle-related gene signatures analyzed in residual vs untreated tumors. (C) Enrichment plots for lung alveolar gene signatures analyzed in residual vs untreated tumors. (D) Expression level of selected lung epithelial marker genes in untreated and residual tumor cells from the PDXs.
| Gene ID | YLR074 | | YLR102 | |
|---------|--------|--------|--------|
|         | Log2FC | p-adj  | Log2FC | p-adj  |
| NKX2-1  | 0.43   | 0.039  | 0.44   | 5.55E-06 |
| AQP4    | 3.28   | 1.17E-33 | 6.08   | 6.47E-186 |

Table S2.1 Changes in the expression of selected lung alveolar signature genes in osimertinib-tolerant cells vs untreated cells from both PDXs.
CHAPTER 3: ASCL1 DRIVES TOLERANCE TO OSIMERTINIB IN EGFR MUTANT LUNG CANCER IN PERMISSIVE CELLULAR CONTEXTS.

Bomiao Hu¹, Fernando J. de Miguel², Zongzhi Liu¹, Lok-Hei Chan¹, Mary Ann Melnick², Anna Arnal Estape¹,²,³, Zenta Walther¹, Anna Wurtz², Rong Fan⁴, Qin Yan¹,², Robert Homer¹,², Don Nguyen¹,², Katerina Politi¹,²,³

¹Department of Pathology, ²Yale Cancer Center and ³Department of Medicine (Section of Medical Oncology), Yale School of Medicine, New Haven, Connecticut; ⁴Department of Biomedical Engineering, Yale School of Engineering & Applied Science, New Haven, Connecticut

# To whom correspondence should be addressed:
Katerina Politi, PhD
Departments of Pathology and Internal Medicine (Section of Medical Oncology)
Yale Cancer Center
Yale University School of Medicine
333 Cedar Street, SHM-I 234D
New Haven, CT 06510 USA
Office Tel: 203-737-5921
Lab Tel: 203-737-6215
Fax: 203-785-7531
Email: katerina.politi@yale.edu
SUMMARY

Histological transformation and lineage transition are well-described mechanisms of acquired resistance to EGFR TKIs whereas detailed cellular and molecular profiling of drug tolerant persister cells that precede these changes in relapsed tumors remain poorly delineated. To gain deeper insights into the potential lineage plasticity and evolution of DTRCs, we further investigated the gene expression profiles of untreated and residual tumor cells in PDXs as discussed in last chapter. Through these analyses we found upregulation of the neuroendocrine lineage transcription factor ASCL1 in one of the PDX models following osimertinib treatment. ASCL1 expression was accompanied by an increase in the expression of its target genes in residual cells. Single-cell RNA-seq revealed that this ASCL1-mediated transcriptional program was enriched in a subpopulation of untreated tumor cells that shared similar transcriptomic features with DTRCs, suggesting that persister cells may arise from the selection of pre-existing TKI tolerant clones. Functional studies of ASCL1 in established human cell lines indicate that in a specific cellular context ASCL1 was able to induce an EMT signature which likely mediated drug tolerance.

INTRODUCTION

LUAD tumors harboring EGFR activating mutations are more sensitive and thus benefit more from EGFR TKI therapy, however, acquired resistance inevitably occurs to all EGFR TKIs. Despite the relatively thorough investigation of on-target EGFR resistance conferring mutations and bypass pathway activation [21, 30], the biology and evolution process underlying another resistance mechanism, such as phenotypic transformations of the tumors, are not well understood. Mutant EGFR-driven lung adenocarcinoma cells that undergo either SCLC transformation or EMT become less dependent on EGFR activity and lose sensitivity to EGFR TKIs. Despite the active searching for therapeutic approaches to overcome or prevent these mechanisms of acquired resistance [49, 83], chemotherapy is still the most common therapeutic
option for patients. This lineage plasticity that can lead to phenotypic transformation is not limited in EGFR mutant lung cancer and has been identified in LUAD treated with other targeted therapies (e.g. ALK inhibitor) or immune checkpoint inhibitors and other types of cancer as well [84-87]. For example, in prostate cancer patients receiving anti-androgen therapy, adenocarcinoma cells can undergo neuroendocrine differentiation giving rise to an aggressive and drug-resistant cancer [87]. Besides trans-differentiation into neuroendocrine lineage, squamous cell carcinoma transformation has also been identified in lung cancer patients receiving first-line osimertinib [88]. This commonality across various clinical settings clearly indicates the existence of mechanisms to overcome the lineage barrier between adenocarcinoma and neuroendocrine cancer or squamous cell carcinoma under the selective pressure of targeted and other therapies. However, little is known about the determinants of lineage transition in that patient subset and how it develops in a stepwise manner. Drug tolerant residual disease, as an intermediate stage, is a critical step during the development of tumor relapse and the studies of its biology pave the way for understanding the lineage plasticity and heterogeneity of cancer cells and identification of potential therapeutic vulnerabilities.

This chapter is focused on the context-dependent role of NE lineage factor ASCL1 in driving TKI tolerance and promoting an EMT signature in mutant EGFR lung cancer. ASCL1 encodes a basic helix-loop-helix transcription factor, which plays a central role in both physiological and pathological processes, including the development of lung neuroendocrine cells and initiating small cell lung cancer in both humans and genetically-engineered mouse models [89-92]. Approximately 8% of lung adenocarcinomas are also positive for ASCL1 and define a subset of these tumors with neuroendocrine features [93]. Using single-cell RNA-sequencing, we confirmed the presence of a population of persister-like cells that express ASCL1 signature genes at a high level even before TKI treatment. Moreover, we demonstrate functionally that ASCL1 led to drug tolerance in cells that were permissive to activation of an epithelial to mesenchymal gene expression program. All the results that we obtained from different model
systems suggest that \textit{ASCL1} may be a potential driver of tolerance to EGFR tyrosine kinase inhibitors in mutant \textit{EGFR} lung cancer.

\textbf{RESULTS}

\textit{An ASCL1-mediated transcriptional program was present in drug tolerant residual disease.}

The dramatic difference in the transcriptional profiles of untreated tumor vs. residual disease in YLR102 (Figure 2.3F) raised the possibility that drug tolerance, in this case, could be due to a population of cells with a new TKI resistant identity altogether. Further examination of individual differentially expressed genes between residual disease and untreated tumor revealed a dramatic upregulation of \textit{ASCL1} in YLR102 (but not in YLR074) (Figure 3.1A). This was especially intriguing because \textit{ASCL1} is a neuroendocrine (NE) lineage transcription factor that is highly expressed in one of the molecular subtypes of SCLC [94]. Since SCLC transformation is one of the mechanisms of resistance to TKI therapies in mutant \textit{EGFR} lung cancer [29], \textit{ASCL1} upregulation in response to osimertinib treatment could be associated with the neuroendocrine differentiation of these tumors. Moreover, \textit{ASCL1} has been shown to be a pioneer factor given its ability to reprogram fibroblasts to neuronal cells [95].

To further confirm these observations, we examined expression of several canonical \textit{ASCL1} downstream targets including \textit{BCL2, DLL3, PROM1} and \textit{RET} [92, 93]. All of these targets were among the top upregulated genes in YLR102 residual disease (Figure 3.1B-E), and the changes were retained at the protein level (Figure 3.1F). Indeed, while phosphorylation of EGFR was effectively suppressed in the residual tumor, \textit{ASCL1} and its downstream targets were all abundantly produced while they were undetectable by immunoblotting in the bulk untreated tumor. Gene set enrichment analysis (GSEA) revealed that two previously published \textit{ASCL1} gene signatures [92, 93] were positively enriched in YLR102 residual tumor cells.
Together these data indicate the presence of an ASCL1-mediated transcriptional program in the residual disease from the YLR102 PDX.

Next, given the role that ASCL1 plays in the establishment of a neuroendocrine program we investigated whether the YLR102 residual disease exhibited features of neuroendocrine differentiation. In this regard, YLR102 residual disease retained histological features of lung adenocarcinoma (Figure 2.1F) and did not show a dramatic increase in classical neuroendocrine markers such as Synaptophysin, Chromogranin A and INSM1 (Figure S3.1).

Collectively, these data suggest that an ASCL1-mediated transcriptional program is enriched in TKI-tolerant residual disease of EGFR mutant lung cancer despite no evidence in neuroendocrine differentiation.

**Rare ASCL1-high tumor cells pre-existing in untreated tumors were selected for upon osimertinib treatment.**

Drug tolerant residual cells can emerge through the selection of a rare drug-tolerant population of cells present in the bulk tumor. Alternatively, treatment-induced changes in a subset of cancer cells could lead to drug tolerance. To determine the origin of the drug-tolerant ASCL1+ positive cells, we stained untreated and residual tumors for ASCL1. ASCL1 showed clear nuclear staining in YLR102 residual tumor samples, which was consistent with its role as a transcription factor (Figure 3.2A-B). ASCL1+ cells were present, but sparse, in untreated tumor samples, as evidenced by immunohistochemistry (Figure 3.2A-B). This led us to hypothesize that a small number of ASCL1^hi^ tumor cells pre-existed in untreated tumors and got selected for by osimertinib treatment. To validate this hypothesis, we used single-cell RNA sequencing to determine whether we could identify a rare drug tolerant cell population in the untreated tumor.

Live EpCAM+ cells were sorted from a dissociated tumor cell suspension to specifically enrich for human tumor cells (Figure S3.2A) since isolated EpCAM+ cells expressed mutant EGFR (Figure S3.2B) and the EPCAM gene is expressed in both untreated and residual
tumors, as shown by RNA-sequencing (Figure S3.2C). After processing of raw data, we identified 5663 untreated cells and 1168 residual cells that would be incorporated into the following analyses. Consistent with findings from our bulk RNA-seq data (Figure 2.3F), the clustering results show a clear separation between untreated and residual cells (Figure 3.2C-D). We also identified a small group of untreated cells that clustered together with residual tumor cells in cluster 2 (Figure 3.2C-D and H), underscoring the transcriptional similarity between these two cell populations in this specific cluster, whereas the majority of untreated cells clustered separately from residual cells in the UMAP plot.

Similar to bulk RNA-seq findings, ASCL1, BCL2 and RET were more highly expressed in drug tolerant residual cells (DTRCs) compared with untreated tumor cells (Figure 3.1A-E and 3.2E-G). The subpopulation of untreated tumor cells that clustered with DTRCs also showed enrichment of ASCL1 and its target genes (Figure 3.2I) providing further evidence that ASCL1 positive DTRC-like tumor cells pre-exist in the untreated tumors and are selected for during treatment with osimertinib. To examine whether selection of a pre-existing DTRC-like subpopulation was unique to the YLR102 PDX we also performed single cell analysis of untreated and residual cells from the YLR074 PDX and found that clusters 6 and 12 were composed of cells from both conditions, indicating that cells transcriptionally resembling DTRCs were also present in untreated YLR074 tumors, even though ASCL1 was not the driver of TKI tolerance in this PDX line (Figure S3.3). These data indicate the presence of a population of untreated tumor cells having properties of DTRCs which upon osimertinib treatment survive and become the main cell population present in residual tumors.

**ASCL1 triggered epithelial-to-mesenchymal transition (EMT) which was the likely mediator of drug tolerance.**

To investigate if and how ASCL1 causes drug tolerance, we turned to human mutant EGFR lung adenocarcinoma cell lines for mechanistic studies. We used doxycycline-inducible vectors
to express ASCL1 in HCC827, PC9 and H1975 cells and evaluated the impact of ASCL1 on colony formation and viability in the presence of osimertinib (Figure 3.3A and S3.4A-B). In HCC827 cells transfected with the empty vector (HCC827-EV), osimertinib eradicated colonies. In contrast, in ASCL1-transfected cells (HCC827-ASCL1), the colony size was reduced by osimertinib but a significantly higher number of colonies survived treatment (Figure 3.3B-C). Further, HCC827-ASCL1 cells also showed a decrease in osimertinib sensitivity compared to EV cells, as evidenced by a shift of the dose-response curve in a growth inhibition assay (Figure 3.3D). The differential behavior of HCC827-EV and -ASCL1 cells when treated with osimertinib indicate that ASCL1 promotes survival and drives drug tolerance in this cell line. In PC9 and H1975 cells, ectopic ASCL1 expression did not affect the clonogenic activity of the cells nor did it affect osimertinib sensitivity (Figure S3.4C-F). These data suggest that the ability of ASCL1 to cause tolerance to EGFR TKIs depends on the cellular context.

To gain mechanistic insights into ASCL1-mediated TKI tolerance and determine what distinguishes cells that are permissive to the phenotypic changes driven by ASCL1 and those that are not, we performed RNA-seq on HCC827 and PC9 cell lines transfected with the EV or ASCL1 vectors in the absence of osimertinib (vehicle) or treated with osimertinib for 3 and 10 days to acquire short-term treated cells and drug tolerant persisters, respectively. Gene expression profiling of the cell lines in different conditions indicated a more dramatic transcriptional change in the untreated HCC827 cell line compared to PC9 cells upon ASCL1 overexpression as revealed through principal component analysis (Figure S3.5A). These findings were consistent with the observation that ASCL1-transfected HCC827 cells exhibited morphological changes including increased number of protrusion-like structures and growth in individual cells rather than in clusters, which were not seen in PC9 cells (Figure 3.4A). Pathway analyses performed on the RNA-seq data from the two cell lines using MetaCore indicated that EMT was the top upregulated pathway in HCC827 cells following ASCL1 expression but it was not enriched in PC9-ASCL1 cells (Table 3.1). Given the known association of EMT with TKI
resistance and activation of the EMT program upon ASCL1 expression in HCC827 cells, we decided to further investigate the relationship between ASCL1-mediated EMT and TKI tolerance. We examined the expression of individual epithelial and mesenchymal markers in our RNA-seq datasets. Although these markers did not change significantly in PC9-ASCL1 cells versus EV cells, their expression was altered dramatically in HCC827 cells following ectopic ASCL1 expression, with a significant downregulation of epithelial markers (E-cadherin and Epcam) and upregulation of mesenchymal markers (N-cadherin and Vimentin), indicative of EMT (Figure 3.4B and S3.5B). This notion was further supported by GSEA, which showed positive enrichment of an EMT gene signature in HCC827-ASCL1 cells as compared to EV cells, but not in PC9 cells (Figure 3.4C).

The EMT phenotype identified in ASCL1-transfected HCC827 cells was strengthened by TKI treatment. As drug tolerance developed, HCC827-ASCL1 cells exhibited greater enrichment of EMT features compared to untreated cells (Figure 3.4B-C and Table 3.1). In contrast to the EMT-mediated drug tolerance in ASCL1-overexpressed cells, EV-transfected HCC827 cells did not persist upon osimertinib treatment and lacked robust enrichment of EMT features (Figure 3.3B and 3.4B). Inspection for the upstream EMT-related transcription factors revealed that ZEB family members shared a similar transcriptional pattern with mesenchymal lineage markers after ASCL1 overexpression and osimertinib treatment, and such a pattern was largely missing for SNAI and TWIST family members (Figure 3.4B and S3.5C). This suggests that the ASCL1-induced EMT phenotype in HCC827 cell line may be mediated by ZEB family members.

We also checked the bulk RNA-seq datasets of the PDX samples for several EMT markers. Consistent with the discoveries from cell lines, in YLR102 residual disease where we identified ASCL1 upregulation, E-cadherin was downregulated and Vimentin was upregulated as drug tolerance developed and such changes were not observed in the ASCL1-unrelated case YLR074 (Figure S3.5D).
Taken together, the results obtained both in vitro and in vivo suggest that ASCL1 is able to drive tolerance to EGFR TKI in mutant EGFR lung cancer by triggering EMT program.

**ASCL1 was expressed and drove neuroendocrine phenotype in a subset of TKI-resistant EGFR mutant lung cancer with SCLC transformation.**

ASCL1 has been implicated as a dominant molecular identifier and master regulator of NE differentiation in a subgroup of primary SCLC [94] but little is known about the molecular subtyping of mutant EGFR lung adenocarcinoma that transforms to SCLC upon TKI exposure. In order to gain insights into the drug resistance mechanisms in a clinically-relevant setting, we developed repeat biopsy program at Yale Cancer Center that allowed us to collect treatment-naïve and TKI resistant tumor specimens from the same patients and use these samples for a variety of applications including histological analysis, NGS, growing them as PDXs and cell lines etc. We chose to analyze the specimens that showed SCLC transformation as mechanism of acquired resistance to EGFR TKIs in our patient collection and characterized their genetics and expression of neuroendocrine marker genes. Since RB1 and TP53 loss are the most common genetic alterations identified in SCLC [96], we first examined the mutational status and copy number changes of these two genes in the patient specimens where genetic profiling data are available. Consistent with SCLC, both RB1 and TP53 were frequently mutated in SCLC transformed EGFR mutant lung cancer as well (3 out of 4) (Figure 3.5A). Copy number variants for the two genes were not commonly identified in most of the patient samples except for YUL0125, which showed a copy number gain of RB1 (Figure 3.5B). YUL0125 harbors a frameshift deletion in RB1 so the amplified gene copies are likely to still carry this deleterious mutation given the role of RB1 as a tumor suppressor gene. Further detection of RB1 proteins in tumor cells by IHC staining is required to validate the functional consequences of the genetic alterations.
As revealed by recent studies, SCLC can be divided into multiple molecular subgroups based on the expression of signature transcriptional regulators including ASCL1, NEUROD1, POU2F3 and YAP1 [94]. We performed histological analyses on patient tumor samples that underwent SCLC transformation (Figure 3.5C) and examined the expression pattern of three molecular identifiers of SCLC (ASCL1, NEUROD1 and YAP1). Three out of five tumors (YLR051, YUL0025 and YUL0125) exhibited positive staining for ASCL1 and one tumor specimen (YLR035) showed weak positivity for NEUROD1 (Figure 3.5D-E). YAP1 only showed focal staining in YLR051 and YLR035 and is thus unlikely to be the transcriptional regulator that defines these cases (Figure 3.5F). YUL0003 is negative for all the three markers and the regulator of NE program in this particular case requires further investigation. The heterogeneous expression of these transcriptional regulators in SCLC transformed EGFR mutant lung cancer is consistent with what is identified in SCLC in general, highlighting that the transcriptional heterogeneity of SCLC is recapitulated in EGFR mutant tumors that transform to SCLC upon TKI treatment.

ASCL1 is able to drive neuroendocrine differentiation and required for maintaining NE status in NE-high SCLC [91, 92, 97, 98]. To characterize the NE features in ASCL1-positive SCLC transformation cases, we examined the expression level of multiple NE markers in pre-treatment and resistant YUL0025 tumor samples through bulk RNA-seq. We observed a dramatic upregulation of ASCL1 and NE marker genes as the tumor cells acquired resistance to EGFR TKI and transformed into SCLC (Table 3.2), indicating a neuroendocrine phenotype in resistant tumor cells. The cellular and molecular profiling of the SCLC transformation cases suggest that expression of SCLC molecular identifiers is heterogeneous, with a subgroup being positive for ASCL1, which is correlated with a strong NE signature. The expression of neuroendocrine lineage factor ASCL1 in TKI-tolerant residual disease might pave the way for NE differentiation and lead to SCLC transformation and eventually acquired resistance.
Figure 3.1 An ASCL1 gene expression program was specifically activated in YLR102 residual tumor cells. (A-E) Expression level of ASCL1 and some of its target genes in untreated and residual tumors of YLR102, as evidenced by bulk RNA-seq. Significance for gene expression changes was determined by unpaired t-test between untreated and residual groups. n=2 biological replicates for untreated group and n=3 biological replicates for residual group. n.s.: p>0.05; **: p<0.01; ***: p<0.001; ****: p<0.0001. (F) Immunoblotting to detect protein level of ASCL1 and its downstream targets. Each lane represents protein lysates from an individual tumor. (G) GSEA results showing enrichment of two ASCL1 gene signatures in residual tumor cells versus untreated ones.
Figure 3.2 ASCL1\textsuperscript{hi} persister-like cells pre-existed in the YLR102 untreated tumor. (A) IHC staining for ASCL1 on untreated and residual tumor sections of YLR102. Red arrow heads point to the nuclei stained as positive. Scale bar: 100μm. (B) Quantification of ASCL1 staining in YLR102. Significance for IHC staining quantification was determined by unpaired t-test between untreated and residual groups. Each group contains six different fields of view. ****: p<0.0001. (C-D) UMAP plots showing clustering of YLR102 untreated and residual tumor cells by single-cell RNA-sequencing. (E-G) Feature plots depicting expression of ASCL1-related genes in individual cells. (H) Number of untreated and residual tumor cells in each individual cluster. Cluster 2 is highlighted in red frame. (I) Heatmap showing relative expression level of ASCL1 signature genes per cluster per condition.
Figure 3.3 ASCL1 overexpression caused TKI tolerance in HCC827 cell line. (A) Immunoblotting for ASCL1 in HCC827 cells transfected with empty vector or ASCL1 vector, upon doxycycline treatment with dose escalation. (B) Clonogenic assay for HCC827-EV or -ASCL1 cells treated with different concentrations of osimertinib for 10 days. (C) Quantification of colony number in (B). Significance was determined by multiple t-test. n=3 technical replicates for each condition. **: p<0.01; ****: p<0.0001. (D) Drug sensitivity assay measuring readout of CellTiter Blue for HCC827-EV and -ASCL1 cells when treated with osimertinib at different doses. n=4 technical replicates per condition.
Figure 3.4 **ASCL1 expression led to EMT phenotype in permissive cellular contexts.** (A) Morphology of PC9 and HCC827 cells transfected with EV or **ASCL1**. Scale bar: 100μm. (B) Expression level of EMT marker genes in PC9 and HCC827 with or without **ASCL1** overexpression, treated with or without osimertinib. Significance was determined by unpaired t-test. n=3 technical replicates for each condition. n.s.: p>0.05; *: p<0.05; **: p<0.01; ***: p<0.001; ****: p<0.0001. V=vehicle-treated; O=osimertinib-treated at 100nM for 3 days; P=persister after 10 days of osimertinib treatment at 100nM. (C) GSEA results showing the enrichment of EMT gene signature in untreated HCC827-ASCL1 cells compared to HCC827-EV cells; HCC827-ASCL1 persister cells compared to untreated HCC827-ASCL1 cells; and untreated PC9-ASCL1 cells compared to PC9-EV cells.
Figure 3.5 TKI resistant EGFR-driven LUAD with SCLC transformation exhibited heterogeneous expression of subtype-defining transcriptional regulators of SCLC. (A-B) Mutational and copy number status RB1 and TP53 in SCLC transformation cases. (C-F) H&E staining (C) and IHC staining for subtype-defining factors of SCLC including (D) ASCL1, (E) NEUROD1 and (F) YAP1. Scale bar: 100μm.
DISCUSSION

Histological transformation such as SCLC transformation and EMT are well-established mechanisms of TKI resistance and have been reliably observed in lung cancer patients receiving EGFR TKIs. The factors that determine which subset of patients undergo this phenotypic change and the detailed evolutionary process through which lineage transition is developed remain poorly understood. Investigating the cellular and molecular features of drug tolerant residual disease provides an opportunity to understand resistant mechanisms more comprehensively. Through transcriptomic profiling of residual tumors in our PDX models, we identified upregulation of *ASCL1* in persistent tumor cells compared to untreated cells in one of the PDXs. *ASCL1* is an important regulator of NE differentiation in SCLC however NE markers did not show enrichment in drug tolerant persister cells (Figure S3.1) and the residual tumor retained histology of adenocarcinoma without NE features (Figure 2.1F), suggesting that lineage transition did not occur in residual tumor cells despite increased expression of *ASCL1*. Previous genetic analyses of *EGFR* mutant lung cancer with SCLC transformation revealed a common *RB1* loss, which is consistent with our patient cohort (Figure 3.5A). The expression of *RB1* appeared to be intact in YLR102 PDX in which *ASCL1* upregulation was identified given that it did not harbor any *RB1* mutation and this gene was normally transcribed as indicated by RNA-seq. Therefore, additional genetic or epigenetic alterations may be required to trigger histological transformation. Ongoing work is focused on determining whether *RB1* loss is required for NE phenotype to be manifested through knockout of *RB1* in PDX and cell lines of mutant *EGFR*-driven lung cancer in the presence of ASCL1.

Through *in vitro* assays for TKI sensitivity, we determined the role of *ASCL1* in driving osimertinib tolerance in established human lung cancer cell lines. *ASCL1* caused TKI tolerance by inducing an EMT signature specifically in HCC827 but not in the other lines, suggesting that cellular context can affect the lineage transforming function of *ASCL1*. Indeed, ASCL1 functions as a pioneer transcription factor to reprogram various cell types into neuronal cells only in the
presence of permissive epigenetic landscape [95]. Given the tendency of the HCC827 cell line to undergo EMT upon resistance to osimertinib, the current working hypothesis is that the genes encoding EMT-related regulators and markers are more accessible in this line compared to the others. To determine the chromatin accessibility for EMT-related genes in the human LUAD cell lines with or without ASCL1 overexpression, we are performing Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) for HCC827 and PC9 cell lines in different conditions to address the complexity of permissiveness underlying ASCL1-induced EMT.

Intratumoral heterogeneity plays a critical role in the development of acquired resistance to targeted therapies. At the genetic level, subpopulations harboring resistance-conferring mutations quickly adapt to drug treatment and becomes the dominant clone in relapse tumor [50, 59, 60]. At transcriptional level, cells in a pre-existing drug tolerant state can survive treatment and be enriched in residual disease, which further evolves into acquired resistance [70]. In the mutant EGFR lung cancer PDXs that we examined in this study, transcriptional changes appeared to be the major contributor to TKI tolerance and single cell RNA-seq revealed pre-existing DTRC-like subclones, supporting the hypothesis of subclonal selection. One question that remains to be answered is whether adaptation to osimertinib may have any effect on the drug tolerant phenotypes. Despite the fact that the results of scRNA-seq clearly support the hypothesis of selection, we cannot exclude that adaptive changes induced by osimertinib may also be occurring and conferring drug tolerance. Future studies combining intricate cell barcoding technology and NGS may facilitate the efforts to track down the cell fate of individual clones and distinguish between the roles of selection and adaptation.

**EXPERIMENTAL PROCEDURES**

**Cell lines.**

All of the human lung adenocarcinoma cell lines (PC9, H1975 and HCC827) and 293T/17 cell line used in my thesis research were purchased from ATCC. The LUAD cell lines were
cultured in RPMI 1640 medium (gibco, 11875-093) supplemented with 10% fetal bovine serum (gibco, 16140-071), 1mM sodium pyruvate (gibco, 11360-070), 10mM HEPES (gibco, 15630-080), 100U/ml penicillin and 100ug/ml streptomycin (gibco, 15140-122). 293T cells were cultured in DMEM medium (gibco, 11965-092) supplemented 10% FBS, 1mM sodium pyruvate, 100U/ml penicillin and 100ug/ml streptomycin. Cells involved in doxycycline induction experiments were cultured with tet system approved FBS (Takara Bio, 631101) instead. Cells were frozen down in FBS plus 10% DMSO (AmericanBio, AB00435-00500) for preservation.

**Immunohistochemistry staining.**

Experimental details are the same as described in chapter 2.

Primary antibodies used for IHC: ASCL1 (abcam, ab211327, 1:500), NEUROD1 (abcam, ab213725, 1:1000), YAP1 (CST, 14074, 1:400).

**Western blotting.**

Experimental details are the same as described in chapter 2.

Primary antibodies used for western blotting: ASCL1 (abcam, ab211327), BCL2 (CST, 4223), DLL3 (CST, 71804), RET (CST, 14556), Nkx2-1 (abcam, ab76013), EpCAM (abcam, ab32392), Vimentin (CST, 5741). All the primary antibodies were diluted in 5% non-fat dry milk by 1:1000, unless stated otherwise.

**Bulk RNA-sequencing.**

Experimental details are the same as described in chapter 2.

**RNA-seq data analysis.**

RNA-seq raw data generated from human cell lines were trimmed of adaptor sequences using Trim Galore (v0.5.0, https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) and
mapped to the human reference genome (hg38) using HISAT2 [99]. The sequencing reads were annotated according to GENCODE v27 model and gene expression levels were quantified using StringTie [100]. Differentially expressed genes were identified using DESeq2 [80].

*Pathway enrichment analysis.*  
Experimental details are the same as described in chapter 2.

*Gene Set Enrichment Analysis.*  
Experimental details are the same as described in chapter 2. ASCL1 gene signatures were obtained from previous studies [92, 93]. EMT gene signature was downloaded from GSEA website.

*Single-cell RNA-sequencing.*  
Subcutaneous PDX tumors were resected from mice and processed with tumor dissociation kit. The single-cell suspension was stained with human EpCAM antibody (BioLegend, 324209, 1:100) and LIVE/DEAD cell stain (invitrogen, L34971, 1:250), and subjected to FACS to isolate live tumor cells, based on the gating strategy shown in Figure S3.2A. The sorted cells were re-suspended in FACS buffer and loaded on the Single Cell A Chip together with RT Master Mix and barcoded gel beads to form nanoliter-scale Gel Beads in Emulsions (GEMs) through Chromium Controller. The single-cell cDNA libraries were constructed using 3' library prep kit v3 (10X genomics) and then sequenced on illumina HiSeq4000 sequencer using 100bp pair-end sequencing to generate approximately 150M reads per sample.

*Single-cell RNA-seq data analysis.*  
The single-cell sequencing raw data were processed and mapped to a human-mouse hybrid genome to distinguish between human and mouse cells. Then gene expression matrix for
human cells is generated by Cell Ranger pipeline and analyzed using Seurat package (v3.0) in R language [101, 102]. Genes whose expression was identified in at least three cells would be considered as features and cells with number of feature genes greater than 200 but smaller than 7000 and percentage of mitochondria reads less than 50% were included in the following analyses. The single-cell expression data were normalized by SCTransform embedded in Seurat v3 and then underwent linear dimensional reduction through PCA. The first 100 principal components were used to cluster cells, with a resolution of 0.4. The clustering results were visualized with UMAP plot. Expression level of individual genes was visualized with feature plots and heatmaps.

**ASCL1 overexpression in human LUAD cell lines.**

Human *ASCL1* cDNA was amplified from phASCL1-N106 (addgene, #31781) with Q5 HF master mix (NEB, M0492S), ASCL1_F and ASCL1_R, and purified with Monarch PCR & DNA Cleanup Kit (NEB, T1030S). Cleavage sites for NheI (NEB, R3131S) and AgeI (NEB, R3552S) were added to both ends during PCR. After digestion of the PCR products and pCW57-MCS empty vector (addgene, #71782) with both restriction enzymes and purification of the digestion products, *ASCL1* cDNA insert was ligated to plasmid backbone using T4 DNA ligase (NEB, M0202S). The ligation products were transfected into stbl3 competent E. coli (invitrogen, C7373-03) and bacteria were spread on ampicillin LB agar plate. Single clones were cultured in LB broth (gibco, 10855-021) with 100ug/ml ampicillin (AmericanBio, AB00115-00010) overnight and plasmids were extracted by QIAprep spin miniprep kit (Qiagen, 27106). Successful ligation was confirmed by sanger sequencing using the sequencing primer.

The pCW57-hASCL1 vector and empty vector were then packaged into lentivirus with TransIT-Lenti reagent (Mirus Bio, MIR6604). Target vectors and two packaging plasmids – pMD2.G (addgene, #12259) and psPAX2 (addgene, #12260) were transfected into 293T cells following instructions of TransIT-Lenti. Lentivirus-containing cell medium was collected 2 days
following transfection and ran through 0.45um filter (Millipore, SLHV033RS) to remove cell debris. Lentivirus was concentrated through ultracentrifugation at 18000rpm, 4 degree for 2h and resuspended with PBS. Lentiviral particles were added to cell medium one day after cell seeding, supplemented with 8ug/ml TransduceIT reagent (Mirus Bio, MIR6620) to facilitate infection. Lentivirus-infected cells were treated with puromycin (gibco, A11138-03) at appropriate concentrations (0.7ug/ml for PC9 and 0.5ug/ml for H1975 and HCC827) for one week to generate stable cell lines.

Cells transfected with EV or ASCL1 vector were maintained in RPMI 1640 cell medium supplemented with 10% tet system-certified FBS and 1ug/ml doxycycline (Sigma-Aldrich, D9891), unless stated otherwise.

Primers used in this section are listed here. ASCL1_F: 5’-AGCTAGCGCGGCCGCCACC-3’. ASCL1_R: 5’-GACCGGTTCGATCTAGA-3’. Sequencing primer for pCW57-hASCL1: 5’-CGTATGTCGAGGTAGGCGTG-3’.

All the plasmids were obtained from Addgene. phASCL1-N106 was a gift from Jerry Crabtree (Addgene plasmid # 31781). pCW57-MCS1-2A-MCS2 was a gift from Adam Karpf (Addgene plasmid # 71782). pMD2.G and psPAX2 were gifts from Didier Trono (Addgene plasmid # 12259 and 12260).

**Clonogenic assay.**

Cells were seeded with doxycycline in 6-well plates and treated with drug vehicle or Osimertinib at various concentrations for 10 days. Seeding density is 250 cells/well for H1975 EV/ASCL1 cells and 500 cells/well for the other cell lines. Cell medium containing drugs was replenished every three days. Cell clones formed in the absence or presence of osimertinib were fixed with 4% PFA and stained with crystal violet (Sigma-Aldrich, C3886). Tissue culture plates were then washed with water to remove excess stain and dried for pictures and
quantification. Osimertinib was dissolved in DMSO and final concentration of drug vehicle was no more than 1:1000 in cell culture medium.

**Drug sensitivity assay.**

Cells were seeded with doxycycline in 96-well plates and treated with drug vehicle or Osimertinib at escalating concentrations for 3 days. Seeding density is 3500 cells/well for PC9 EV/ASCL1 cells, 3000 cells/well for H1975 EV/ASCL1 cells, 3500 cells/well for HCC827 EV cells and 2500 cells/well for HCC827 ASCL1 cells. Cell viability was determined with CellTiter Blue reagent (Promega, G8081) following manufacturer’s protocol. Fluorescence was measured with SpectraMax M3 plate reader.
Figure S3.1 YLR102 residual disease did not show histological transformation into SCLC. The table showed the RPKM value of individual NE marker genes and SCLC subtype identifiers in each of the YLR102 samples as well as the log2-scaled fold change in gene expression in residual disease group as compared with untreated group.
Figure S3.2 Human tumor cells were enriched for single-cell RNA-sequencing by FACS. (A) Experimental scheme. (B) Expression of lung epithelial markers in EpCAM+ and EpCAM- cells sorted from YLR102 PDX. (C) Expression level of EPCAM in untreated and residual tumors of YLR102.
Figure S3.3 DTRC-like tumor cells pre-existed in untreated YLR074 PDX. (A) UMAP plots showing clustering of YLR074 untreated and residual tumor cells by single-cell RNA-sequencing. (B) Number of untreated and residual tumor cells in each individual cluster. Cluster 6 and 12 are highlighted in red frame.
Figure S3.4 ASCL1 overexpression did not affect osimertinib sensitivity in PC9 or H1975 cell lines. (A-B) Immunoblotting for ASCL1 in (A) PC9 and (B) H1975 cells transfected with empty vector or ASCL1 vector, upon doxycycline treatment with dose escalation. (C-D) Clonogenic assay for EV or ASCL1 transfected (C) PC9 and (D) H1975 cells treated with different concentrations of osimertinib for 10 days. (E-F) Drug sensitivity assay for (E) PC9-EV/ASCL1 and (F) H1975-EV/ASCL1 cells when treated with osimertinib at different doses.
(A) Gene expression clustering analysis.

(B) Western blot analysis showing protein expression levels. Please note the specific proteins and their bands.

(C) Quantitative analysis of gene expression levels showing the following conditions:
- **SNAI1**
- **SNAI2**
- **TWIST1**
- **TWIST2**

(D) RPKM (Reads Per Kilobase of transcript per Million mapped reads) analysis for **CDH1_074&102** and **VIM_074&102** genes in different conditions:
- Untreated
- Residual

**Conditions**
- HCC827ASC1_O
- HCC827ASC1_P
- HCC827ASC1_V
- HCC827EV_O
- HCC827EV_V
- PC9ASC1_0
- PC9ASC1_V
- PC9EV_O
- PC9EV_V

**Protein Bands**
- PC9 ASC1
- PC9 EV
- HCC827 ASC1
- HCC827 EV

**Gene Expression**
- **CDH1_074&102**
- **VIM_074&102**

**Statistical Significance**
- n.s.: not significant
- *: significant
- ***: highly significant
**Figure S3.5** *ASCL1* expression led to EMT phenotype in permissive cellular contexts. (A) PCA plot visualizing the gene expression profiles of different experimental groups. (B) Immunoblotting of epithelial and mesenchymal markers in PC9 and HCC827 cell lines following *ASCL1* overexpression. (C) Expression level of SNAI and TWIST transcription factors in PC9 and HCC827 cells under different experimental conditions. (D) Expression of epithelial and mesenchymal markers in untreated and residual tumors in YLR074 and YLR102 revealed by RNA-seq. Significance was determined by unpaired t-test. n=2 biological replicates for untreated group and n=3 biological replicates for residual group. n.s.: p>0.05; *: p<0.05; ***: p<0.001.
### Top 5 pathways upregulated in HCC827ASCL1-untreated vs HCC827EV-untreated

<table>
<thead>
<tr>
<th>#</th>
<th>Pathways</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Development_Regulation of epithelial-to-mesenchymal transition (EMT)</td>
<td>3.08E-11</td>
</tr>
<tr>
<td>2</td>
<td>Stromal-epithelial interaction in Prostate Cancer</td>
<td>2.29E-09</td>
</tr>
<tr>
<td>3</td>
<td>Development_Negative regulation of WNT/Beta-catenin signaling at the receptor level</td>
<td>6.37E-09</td>
</tr>
<tr>
<td>4</td>
<td>Signal transduction_WNT/Beta-catenin signaling in tissue homeostasis</td>
<td>2.39E-08</td>
</tr>
<tr>
<td>5</td>
<td>Development_Negative feedback regulation of WNT/Beta-catenin signaling</td>
<td>4.62E-08</td>
</tr>
</tbody>
</table>

### Top 5 pathways upregulated in HCC827ASCL1-persister vs HCC827ASCL1-untreated

<table>
<thead>
<tr>
<th>#</th>
<th>Pathways</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Development_Regulation of epithelial-to-mesenchymal transition (EMT)</td>
<td>1.55E-12</td>
</tr>
<tr>
<td>2</td>
<td>Cell adhesion_ECM remodeling</td>
<td>4.74E-11</td>
</tr>
<tr>
<td>3</td>
<td>IL-1 beta- and Endothelin-1-induced fibroblast/ myofibroblast migration and extracellular matrix production in asthmatic airways</td>
<td>1.26E-10</td>
</tr>
<tr>
<td>4</td>
<td>Development_Regulation of cytoskeleton proteins in oligodendrocyte differentiation and myelination</td>
<td>1.91E-10</td>
</tr>
<tr>
<td>5</td>
<td>TGF-beta-induced fibroblast/ myofibroblast migration and extracellular matrix production in asthmatic airways</td>
<td>2.65E-10</td>
</tr>
</tbody>
</table>

### Top 5 pathways upregulated in PC9ASCL1-untreated vs PC9EV-untreated

<table>
<thead>
<tr>
<th>#</th>
<th>Pathways</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PXR-mediated direct regulation of xenobiotic metabolizing enzymes / Rodent version</td>
<td>1.42E-04</td>
</tr>
<tr>
<td>2</td>
<td>PXR-mediated direct regulation of xenobiotic metabolizing enzymes / Human version</td>
<td>1.65E-04</td>
</tr>
<tr>
<td>3</td>
<td>1-Naphthylamine and 1-Nitronaphtalene metabolism</td>
<td>1.90E-04</td>
</tr>
<tr>
<td>4</td>
<td>2-Naphthylamine and 2-Nitronaphtalene metabolism</td>
<td>3.38E-04</td>
</tr>
<tr>
<td>5</td>
<td>Heme metabolism</td>
<td>9.95E-04</td>
</tr>
</tbody>
</table>

Table S3.1 Top 5 upregulated pathways in HCC827 and PC9 cells after ASCL1 overexpression and osimertinib treatment as predicted by MetaCore software.
Table S3.2 Expression level of NE markers and subtype-defining transcriptional regulators of SCLC in pre- and post-TKI *EGFR*-driven lung cancer. The table showed the FPKM value of individual NE marker genes and SCLC subtype identifiers in patient specimens obtained from pre-TKI (YUL0025P) and TKI-resistant SCLC-transformed tumors (YUL0025R) from the same patient. Expression levels of dramatically upregulated genes were highlighted in red for post-TKI tumor sample.

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>YUL0025P</th>
<th>YUL0025R</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYT</td>
<td>0.25</td>
<td>19.22</td>
</tr>
<tr>
<td>CHGA</td>
<td>0.43</td>
<td>0.48</td>
</tr>
<tr>
<td>INS1M1</td>
<td>0</td>
<td>5.59</td>
</tr>
<tr>
<td>SCG3</td>
<td>0.16</td>
<td>9.82</td>
</tr>
<tr>
<td>GRP</td>
<td>0</td>
<td>72.44</td>
</tr>
<tr>
<td>CALCA</td>
<td>0</td>
<td>5.05</td>
</tr>
<tr>
<td>SYN1</td>
<td>0.13</td>
<td>5.37</td>
</tr>
<tr>
<td>ASCL1</td>
<td>0.28</td>
<td>21.6</td>
</tr>
<tr>
<td>NEUROD1</td>
<td>0</td>
<td>0.03</td>
</tr>
<tr>
<td>POU2F3</td>
<td>2.75</td>
<td>1.98</td>
</tr>
<tr>
<td>YAP1</td>
<td>50.12</td>
<td>16.91</td>
</tr>
</tbody>
</table>

- **NE Markers**
- **Subtype-Defining Factor of SCLC**
CHAPTER 4: MODELING RESIDUAL DISEASE IN GENETICALLY-ENGINEERED MOUSE MODEL OF MUTANT EGFR LUNG CANCER.

Bomiao Hu\textsuperscript{1}, Fernando J. de Miguel\textsuperscript{2}, Lok-Hei Chan\textsuperscript{1}, Mary Ann Melnick\textsuperscript{2}, Rong Fan\textsuperscript{4}, Qin Yan\textsuperscript{1,2}, Robert Homer\textsuperscript{1,2}, Don Nguyen\textsuperscript{1,2}, Katerina Politi\textsuperscript{1,2,3}

\textsuperscript{1}Department of Pathology, \textsuperscript{2}Yale Cancer Center and \textsuperscript{3}Department of Medicine (Section of Medical Oncology), Yale School of Medicine, New Haven, Connecticut; \textsuperscript{4}Department of Biomedical Engineering, Yale School of Engineering & Applied Science, New Haven, Connecticut

# To whom correspondence should be addressed:
Katerina Politi, PhD
Departments of Pathology and Internal Medicine (Section of Medical Oncology)
Yale Cancer Center
Yale University School of Medicine
333 Cedar Street, SHM-I 234D
New Haven, CT 06510 USA
Office Tel: 203-737-5921
Lab Tel: 203-737-6215
Fax: 203-785-7531
Email: katerina.politi@yale.edu
SUMMARY

PDX models serve as critical tools for understanding drug tolerant residual disease from a clinical perspective given that these preserve a number of biological traits of patient specimens. However, the lack of a functional immune system in an NSG mouse host results in a defective interaction between tumor cells and the microenvironment. To gain more comprehensive knowledge about TKI tolerance in an immunocompetent system, we performed complementary research using a refined genetically-engineered mouse model (GEMM), in which lung tumor cells are labeled with the fluorescent marker mKate to enable cell sorting. We characterized multiple aspects of this model, including tumor development, histology, response to EGFR TKIs and isolation of untreated and residual tumor cells. Gene expression profiling of mKate+ cells sorted from tumor bearing mice with or without osimertinib treatment revealed common pathways that are deregulated in both GEMM and PDX models.

INTRODUCTION

Cancer can be modeled both in vitro as 2D or 3D cultures and in vivo with tumors growing in experimental animals. In chapter 2 and 3, we discussed our studies on drug tolerant residual disease in PDX models, which provided insights into the mechanisms of drug tolerance in a clinically-relevant setting. In order to reduce graft rejection and ensure successful engraftment of human tumor tissues in a murine system, the mouse hosts are engineered to have a severe immune deficiency [103, 104]. The lack of a functional immune system in PDXs means that interactions between tumor cells and immune cells in the microenvironment are not recapitulated in these models. Furthermore, in most PDX models, tumor cells are grown subcutaneously rather than in the organs where patient biopsies were obtained. Therefore these models do not fully represent tumor development and drug response in the most physiologically-relevant conditions. In order to gain comprehensive knowledge about mechanism of TKI
tolerance in \textit{EGFR} mutant lung cancer, we performed complementary studies on the biology of tumor cells in an immune-competent genetically-engineered mouse model.

Transgenic mice have been widely used to study the biology and drug resistance of \textit{EGFR} mutant lung adenocarcinoma [23, 53, 105]. Transgenes encoding oncogenic drivers are integrated into the mouse genome and their expression is induced specifically in lung epithelial cells using a doxycycline-inducible genetic approach, resulting in transformation of healthy lung epithelial cells into lung adenocarcinoma cells. The mouse tumors therefore develop orthotopically in the lungs in the presence of the lung microenvironment and tumor-bearing mice can be used to study the biology of the tumors, therapeutic responses, drug resistance and, importantly for this project, drug tolerant persister cells.

GEMMs have been successfully used to model acquired resistance to EGFR TKIs and identify resistance mechanisms. In general, the tumor-bearing mice respond quite well in the beginning of TKI treatment and the mouse lungs appeared to be normal through imaging and histological analysis, however, TKI resistant tumor nodules develop after a prolonged and continuous treatment [23, 106], indicating the presence of drug tolerant residual cells despite the almost complete drug response.

In this chapter, we refined and characterized a GEMM of \textit{EGFR}-driven lung cancer such that cancer cells were labeled with the fluorescent protein mKate that enabled them to be sorted. mKate+ cells isolated from untreated and treated mice can be genetically and transcriptionally profiled to identify mechanisms of drug tolerance. Although tumor-bearing mice showed an almost complete response to osimertinib, mKate+ cells were still present following TKI treatment. Through targeted deep sequencing, we did not find any \textit{EGFR} mutations in the residual cells that are known to confer resistance. Comparison of the gene expression profile between untreated and residual cells revealed some gene signatures that were shared between residual diseases identified in PDX models and GEMM as well as features that were unique to the transgenic mouse model.
RESULTS

Identification and isolation of drug tolerant residual cells in a GEMM of EGFR mutant lung cancer.

To induce the expression of mutant EGFR and label tumor cells with a fluorescent marker, we crossed mice carrying a doxycycline-inducible mutant EGFR transgene with a conditional allele that encodes the rtTA transcriptional activator and red fluorescent protein mKate that were expressed upon deletion of a lox-stop-lox cassette by Cre recombinase [107]. A CreER transgene was under control of Sftpc promoter, which enabled the expression of recombinase specifically in type II alveolar lung epithelial cells in a tamoxifen inducible manner [108]. Additionally, a Trp53 floxed allele is incorporated into the model to accelerate tumor development and increase the aggressiveness of tumor cells [109, 110]. Mice of the appropriate genotype (Figure 4.1A-B) were administered tamoxifen to activate CreER recombinase, which initiated rtTA and mKate expression and deleted Trp53 gene in lung epithelial cells. Mice were then fed with a doxycycline diet to induce expression of the EGFR L858R mutant, which drove transformation of healthy lung epithelial cells into lung adenocarcinoma cells. Theoretically, all the cells that express mutated EGFR would be labeled with mKate and can be isolated through FACS.

To verify tumor development, we monitored the tumor volume in mouse lungs by magnetic resonance imaging (MRI). Typically, after 6-10 weeks on a DOX diet and initial tamoxifen treatment, mice developed diffused tumor lesions in the lungs that were responsive to both the first generation TKI erlotinib and third generation TKI osimertinib (Figure 4.1C-D and S4.1A-B). The lung tumors regressed almost completely after treatment with osimertinib for 4 weeks and the histology of the lungs of TKI-treated mice reverted from harboring lung adenocarcinomas to looking histologically normal and similar to those in healthy mice (Figure 4.1E and S4.1C). In contrast, tumor bearing mice only showed partial responses to erlotinib in most cases (Figure
S4.1C), indicating a more potent anti-tumor effect of osimertinib compared to erlotinib. Whether this reflects a difference in the ability of the TKIs to inhibit EGFR in this model remains to be determined.

One of the advantages of this transgenic mouse model is expression of fluorescent marker in lung epithelial cells, including tumor cells. To validate expression of mKate and confirm the presence of drug tolerant residual cells in mice with complete tumor regression, we probed for mKate+ cells in mouse lungs before and after osimertinib treatment by IHC staining and FACS. Pre-treatment tumor cells showed strong positivity for EGFR mutant and mKate (Figure 4.2A). The mKate+ cells could also be detected by flow cytometry both before and after osimertinib treatment (Figure 4.2B-C) despite the absence of sizable post-treatment tumor lesions as revealed by MRI and histological analysis (Figure 4.1C-E). To determine the appropriate time point to collect DTRCs, we treated tumor-bearing mice with osimertinib for 4 or 6 weeks until the tumors were no longer visible and compared the number of mKate+ cells isolated from mouse lungs between the two groups. There was no significant difference between the 4- and 6-week groups, indicating that maximal tumor regression had been achieved by week 4 on drug (Figure 4.2B). Therefore, we collected mKate+ cells after 4 weeks of osimertinib treatment. Besides mKate+ cells isolated from TKI-naïve and treated mice, we were also able to collect them from healthy mice that do not harbor mutant EGFR transgene through sorting (Figure 4.2B and 4.3A). The cellular profiling of the lung tumors suggests that tumor cells still persist even after they become undetectable by imaging and the DTRCs, although few in number, can be collected through FACS.

**Transcriptional profiling of the tumor cells isolated from the pre- and post-treatment GEMM revealed pathways deregulated in DTRCs.**

In order to identify the mechanisms that drive TKI tolerance in the GEMM, we sought to investigate and compare the molecular features of untreated tumor cells and drug tolerant
persisters. As GEMMs of mutant \( \text{EGFR} \) lung cancer are often used to study on-target mechanisms of acquired resistance to EGFR TKIs [23, 106], our first step for this molecular profiling was to search for mutations in \( \text{EGFR} \) that conferred resistance to first-, second- and third-generation TKIs in pre- and post-treatment tumor cells isolated from the GEMM (Figure 4.3A). We performed targeted deep sequencing for the \( \text{EGFR} \) coding regions that covered the codons that were known to be mutated and cause acquired resistance, including L718V/Q, T790M and C797S [23-27]. We also examined the presence of the \( \text{EGFR} \) L858R mutation encoded in the transgene as positive control to ensure proper sequencing and data analysis. The common drug resistant \( \text{EGFR} \) mutations were not detected in treatment-naïve, osimertinib-treated or erlotinib-treated tumor cells whereas the L858R mutation was uniformly present (Figure S4.2), suggesting that resistance conferring mutations in \( \text{EGFR} \) were not the major contributors to the drug tolerant phenotypes in the TKI-treated GEMM.

Since on-target genetic mechanisms were not identified in post-osimertinib mKate+ cells, we then went on to profile the transcriptomic changes in mKate+ cells after 4 weeks of osimertinib treatment versus mKate+ cells prior to TKI treatment. We sorted mKate+ cells from TKI naïve and long-term osimertinib treated (4 weeks) tumor bearing mice, as well as from healthy mice in which lung epithelial cells were labeled with mKate but lacked mutant \( \text{EGFR} \) expression (Figure 4.3A). Bulk RNA-seq of mKate+ cells collected from healthy, untreated and osimertinib-treated mice revealed dramatic differences in gene expression between the three groups (Figure 4.3B). This result further indicates that mKate+ cells isolated from healthy mice and tumor bearing lungs following complete response to osimertinib were not transcriptionally identical despite the similar histology shared by healthy tissue and osimertinib treated lung tissue (Figure 4.1E). The expression level of lung epithelial markers was largely retained across different conditions, indicating the isolation of correct cell types through mKate sorting (Table 4.1). In order to systematically analyze the gene expression profiles, we performed the MetaCore pathway analysis for the bulk RNA-seq data of mKate+ cells to identify the pathways that were potentially
altered in post- vs pre-treatment mKate+ cells. Such analyses revealed transcriptional programs that were positively or negatively enriched in osimertinib-treated mKate+ cells compared to untreated counterparts. Top pathways that were upregulated in post-treatment cells included a lipid biosynthesis pathway mediated by SREBP and development-related NOTCH pathway. The functional and mechanistic implications of these pathways remain to be studied in order to address how they are related to TKI tolerance. The downregulated pathways in mKate+ cells following osimertinib treatment were mainly centered in cell cycle and target genes of HIF-1α (Figure 4.3C). It is worth noting that similar programs were also implicated in pathway analysis for residual tumor cells in our PDX models, including upregulation of development-related pathways such as Wnt signaling and Hedgehog signaling and downregulation of cell cycle-related pathways and HIF-1α activity (Figure S2.5A-B). The commonality between GEMM and PDX models suggests that signaling networks critical for normal growth and development may be hijacked by residual tumor cells to promote survival and that treatment-induced cell cycle arrest and reduced hypoxia due to tumor regression are hallmarks of residual disease.
Figure 4.1 Tumor-bearing transgenic mice of mutant EGFR LUAD showed complete responses to osimertinib. (A) Transgenic composition of the GEMM used for study of drug tolerance. (B) Experimental schema for tumor development and TKI treatment. (C) Tumor burden in osimertinib-treated mice as measured by MRI. (D) MRI images of pre- and post-treatment mouse lungs. (E) H&E staining showing the histology of the lungs from treatment-naïve, osimertinib-treated and healthy mice. Scale bar: 100μm.
Figure 4.2 Isolation of mKate+ cells from the GEMM. (A) IHC staining for EGFR L858R mutant and mKate. Scale bar: 50 μm. (B) Number of mKate+ cells isolated from tumor-bearing mice after osimertinib treatment for 4 or 6 weeks and from healthy mice. (C) Flow charts showing the presence of mKate+ cells in mouse lungs both before and after treatment. The genotypes and drug treatment schedules of the mice exemplified in each group are listed below each flow chart.
Figure 4.3 Transcriptomic profiling of mKate+ cells isolated from the GEMM. (A) Experimental schema for molecular profiling of mKate+ cells isolated in various conditions. (B) PCA plot showing the differences in gene expression profiles of mKate+ cells between healthy, untreated and osimertinib-treated groups. (C) Results of MetaCore pathway analysis showing the pathways up- or downregulated in osimertinib-treated mKate+ cells as compared to untreated cells.
DISCUSSION

The transgenic mouse model of mutant EGFR lung cancer provides further insights into mechanisms of TKI tolerance, which are complimentary to the PDX models as discussed in Chapter 2 and 3. We performed different types of NGS for the tumor cells collected from the two models and profiled the molecular properties of drug tolerant residual disease in vivo. One prominent feature identified through DNA sequencing that is shared across different mouse models is lack of resistance conferring mutations in both pre- and post-TKI treatment tumor cells. In PDXs, we identified neither osimertinib resistant mutations in EGFR that emerged following treatment nor any dramatic changes in allelic frequency of already existing mutations in residual disease compared to untreated tumors. In GEMMs, there was no selection or enrichment of common mutations in EGFR transgene that cause resistance to first-, second- or third-gen TKIs in DTRCs. These results suggest that at least in the models that we have examined, TKI resistant mutations, if any, tend to develop de novo during treatment rather than pre-exist in treatment-naïve tumors. To support this idea, studies of acquired resistance to osimertinib in GEMMs suggest that it generally takes 5-6 months to obtain resistant tumor nodules [23] and 4 weeks of TKI treatment is probably not sufficient for observing a dominant clone with secondary mutations. The PDX models (YLR074 and YLR102) require even longer period of treatment to develop acquired resistance. Pilot studies suggest that tumors were still sensitive to osimertinib following treatment for about a year with drug-on and -off schedule. Therefore, TKI resistant mutations are unlikely to be present at high frequency in persistent tumor cells after only six weeks of treatment. Furthermore, in vitro analyses of human mutant EGFR lung cancer cell lines indicated the presence of a subpopulation free of pre-existing drug resistant T790M mutation, which developed acquired resistance and gained T790M mutation after long-term treatment with first-gen EGFR TKI, indicating the possibility of de novo mutagenesis given enough time of drug treatment [50].
In the DTRCs isolated from GEMM, we identified an upregulated transcriptional program that was related to cholesterol and fatty acid biosynthesis and controlled by the transcription factor SREBP. This finding led to a hypothesis that changes in metabolic pathways could be underlying the drug tolerant phenotypes in tumor cells. In fact, metabolic reprogramming has been characterized in residual tumor cells following oncogene deprivation in transgenic mouse models of breast cancer [111]. The tumor cells that persisted after cessation of oncogene expression shifted to lipid metabolism, which resulted in production of higher levels of reactive oxygen species (ROS) and altered redox balance within the cells. The metabolic shifts caused increased accumulation of DNA damage mediated by ROS and eventual tumor recurrence. In order to understand how SREBP-related transcriptional changes shape the metabolic status in drug tolerant persisters, future experiments will be focused on profiling the regulators and enzymes in different metabolic pathways to determine the relationship between metabolic reprogramming and survival of DTRCs.

The Cre recombinase-mediated conditional rtTA expression system in this refined GEMM can be adapted in various ways. When combined with lentiviral-Cre and other genetic tools, this model enables direct genetic modifications and cell barcoding in mouse lung tumor cells. Recent CRISPR screening studies in the GEMM of mutant EGFR lung cancer revealed the possibility to edit multiple tumor suppressor genes simultaneously and compare tumor development and TKI responses in different genetic contexts [110]. A refined model like this is useful in terms of investigating the biology of drug tolerant persisters in the presence of various co-occurring mutations identified in lung cancer patients.

In the GEMM that we discussed in this chapter, we identified mKate+ cells in tumor-bearing mice with complete responses to osimertinib. One potential problem with this model, especially when drug response is profound and tumor lesion is invisible, is that mKate-expressing cells may not be transformed and retain the phenotypes of healthy epithelial cells. In the context of drug tolerant residual disease, these mKate+ normal cells can be sorted out together with
persister tumor cells, confounding the molecular profiling studies of drug tolerant mechanisms. A potential solution to this problem is to perform single-cell sequencing on the isolated mKate+ cells to distinguish between healthy and tumor cells and to untangle the complexity resulting from a mixed cell population.

EXPERIMENTAL PROCEDURES

Initiate tumor growth in the GEMM.

6-8 weeks old mice with appropriate genotypes were given two consecutive doses of tamoxifen (Sigma-Aldrich, C5648) at 0.25mg/g to active Cre recombinase. Mice were then fed with doxycycline diet (Envigo) to induce expression of oncogenic EGFR. Tumor development was monitored by MRI until the mice were ready to be treated.

Tamoxifen was dissolved at 25mg/ml in corn oil (Sigma-Aldrich, C8267), with overnight shaking.

Drug treatment in vivo.

Tumor-bearing mice were treated with osimertinib at 25mg/kg for five consecutive days each week when their tumor size reached around 300-500mm³. The treatment lasted for 4 weeks before mouse lungs harboring drug tolerant residual tumor was excised. Tumor burden was measured by MRI at least once before treatment and once before sacrifice.

Magnetic resonance imaging.

Experimental details of magnetic resonance imaging on mice were described previously [112]. In brief, mice were anesthetized and maintained on isoflurane/O2 (2–2.5% v/v) throughout the imaging process. MR images were collected using a 4T small-animal Bruker horizontal-bore spectrometer (Bruker AVANCE). MR data collection was synchronized with animal respiration through a small-animal monitoring and gating system (SA Instruments, Inc),
to make sure that all the images were captured during postexpiratory periods. Tumor burden was determined by quantification of lung opacity area present in MR images using BioImage Suite 3.01 [113].

**Immunohistochemistry staining.**

Experimental details are the same as described in chapter 2.

Primary antibodies used for IHC: EGFR L858R mutant (CST, 3197, 1:100), mKate (Evrogen, AB233, 1:1000).

**Sorting of mKate+ cells.**

Mouse lungs excised from mice with desired genotypes and receiving appropriate treatments were processed with mouse lung dissociation kit (Miltenyi Biotech, 130-095-927). The single-cell suspension was subjected to FACS to identify and isolate mKate+ cells, based on the gating strategy shown in Figure 4.2C.

**Targeted deep sequencing for EGFR transgene.**

Genomic DNA was extracted from sorted mKate+ cells with AllPrep DNA/RNA Micro Kit (Qiagen, 80284) and used as template to amplify two coding regions in mutant EGFR transgene: cDNA 2065-2355 and 2337-2604, which cover the common TKI resistant mutations in EGFR kinase domain. The PCR was performed with HotStarTaq Master Mix Kit (Qiagen, 203443). Sequencing libraries were prepared from PCR products with KAPA HyperPrep Kit (Kapa Biosystems, KK8504), KAPA Dual-Indexed Adapter Kit (Kapa Biosystems, KK8722), and KAPA Pure Beads (Kapa Biosystems, KK8000). The libraries were pooled and sequenced on illumina NovaSeq 6000 sequencer at YCGA to generate 150bp paired end reads. The raw sequencing reads were first aligned to the EGFR amplicon sequences using the Burroughs-
Wheeler aligner [72], and an in-house script was used to calculate the allelic frequency at each position seen in the amplicon reference.

**Bulk RNA-sequencing.**

Total RNA was extracted from sorted mKate+ cells with AllPrep DNA/RNA Micro Kit (Qiagen, 80284). RNA integrity was determined by running an Agilent Bioanalyzer gel, which measures the ratio of the ribosomal peaks. Samples with RIN values of 7 or greater proceeded to library prep. RNA-seq libraries were constructed using SMART-Seq v4 Ultra Low Input RNA Kit (Takara Bio) and sequenced on illumina HiSeq2500 sequencer using 75bp single-end sequencing to generate around 30M reads per sample.

**Bulk RNA-sequencing data analysis.**

Low quality reads were trimmed and adaptor contamination were removed using Trim Galore (v0.5.0, https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Trimmed reads were mapped to the mouse reference genome (mm10) using HISAT2 (v2.1.0) [99]. Gene expression levels were quantified using StringTie (v1.3.3b) [100] with annotations based on gene models (vM15) from the GENCODE project. Differentially expressed genes were identified using DESeq2 (v 1.22.1) [80].

**Pathway enrichment analysis.**

Experimental details are the same as described in chapter 2.
Figure S4.1 Tumor-bearing transgenic mice of mutant EGFR LUAD showed partial responses to erlotinib. (A) Tumor burden in erlotinib-treated mice as measured by MRI. (B) MRI images of pre- and post-treatment mouse lungs. (C) Waterfall plot showing the degree of drug response in osimertinib- and erlotinib-treated mice.
Figure S4.2 On-target resistance conferring EGFR mutations were not enriched in mKate+ cells isolated from TKI-treated GEMM. Bar plots showing the frequency of wildtype alleles at the codon encoding EGFR mutations known to confer resistance to first-, second- and third-generation EGFR TKIs in mKate+ cells sorted from untreated and TKI-treated mice. EGFR L858R mutation was analyzed as a positive control and the bar plot for L858R showed the frequency of mutant allele.
<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Normalized Gene Counts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Healthy</td>
</tr>
<tr>
<td>Nkx2-1</td>
<td>8703.3</td>
</tr>
<tr>
<td>Stipc</td>
<td>4.42e6</td>
</tr>
<tr>
<td>Epcam</td>
<td>9764.4</td>
</tr>
<tr>
<td>Scgb1a1 (CC10)</td>
<td>32827.7</td>
</tr>
</tbody>
</table>

Table S4.1 Expression of lung epithelial markers in mKate+ cells isolated from healthy, untreated and osimertinib-treated mice.
CLOSING REMARKS

Incomplete response to therapies occurs in nearly all types of cancer. In many cases, tumor lesions in patients receiving anti-tumor agents initially undergo regression, followed by an intermediate drug tolerant stage where tumor burden remains constant as treatment continues with eventual tumor progression (acquired resistance). In mutant $EGFR$-driven lung cancer for which targeted therapy is available, residual disease is a major challenge since acquired resistance to EGFR TKIs is inevitable and the main reason for treatment failure. Studies on drug tolerant persisters in TKI-treated lung cancer cell lines have broadened our knowledge about the biology of persister cells in vitro. Further investigation of residual disease from a more clinically and physiologically relevant perspective will deepen our understanding about drug tolerant mechanisms in vivo and shed light on the evolution of acquired resistance in more detail. To pursue this and gain insights into the biology of residual disease, we studied the biological properties of drug tolerant residual disease in PDX mouse models, which more resembled the complexity and diversity of clinical specimens than established cell lines.

In the three PDXs that we characterized in my thesis project, two of them showed incomplete tumor regression in response to the third-generation TKI osimertinib, which mimicked the residual disease observed in mutant $EGFR$ lung cancer patients receiving EGFR TKIs. The cellular features of persistent tumor cells included cell cycle arrest and low levels of apoptosis, as evidenced by IHC staining for proliferation and apoptosis markers. To identify mechanisms underlying TKI tolerance in DTRCs, we examined the genomic and transcriptomic profiles of residual tumor cells and explored the changes in the mutational landscape and gene expression before and after osimertinib treatment. As indicated by WES, the mutational landscape remained very similar between untreated tumor and residual disease. However, gene expression profiling revealed a number of differentially expressed genes in residual tumor cells compared to the untreated ones in both PDXs, indicating that transcriptional rather than genetic mechanisms potentially accounted for TKI tolerance in persister cells.
Close examination of gene expression data for PDX tumor samples led to identification of ASCL1 upregulation in one of the PDXs post-osimertinib treatment. Gene-set enrichment analysis suggested that ASCL1 functioned as a transcription factor and a gene signature associated with it was induced in the residual cells. To understand how the tumor cells evolved into DTRCs during TKI therapy, we performed single-cell RNA-sequencing for untreated and residual tumor cells, which revealed the presence of a pre-existing DTRC-like subpopulation in untreated tumors characterized by high expression of ASCL1. The intratumoral heterogeneity identified in PDXs raised the possibility that residual cells resulted from the selection of a subpopulation with drug tolerant state by EGFR TKIs.

Subsequent functional and mechanistic studies demonstrate that ASCL1 overexpression led to osimertinib tolerance in one out of the three mutant EGFR lung adenocarcinoma cell lines that we tested. In the cell line that showed TKI tolerant phenotypes but not in the others, ectopic expression of ASCL1 induced morphological changes and an EMT gene signature, indicating a context-dependent role of ASCL1 in driving phenotypic changes in EGFR mutant lung cancer.

In addition to studying PDX models, to get a more comprehensive understanding of mechanisms of TKI tolerance, we also interrogated the biological features of drug tolerant persister cells in an immune-competent GEMM of EGFR mutant lung cancer in which lung epithelial cells were labeled with fluorescent marker to enable FACS. Transcriptomic profiling of the mKate+ cells isolated from mice before and after TKI treatment by RNA-seq indicated a variety of deregulated gene expression programs in residual cells, which were involved in regulatory and survival pathways, including upregulation of cholesterol/fatty acid biosynthesis and NOTCH signaling pathway, as well as downregulation of cell cycle-related pathways and HIF1α-regulated program. Further studies to understand the biological implications of deregulation of these pathways remain to be performed.

Despite all the in vivo and in vitro findings that lead to deeper insights into residual disease in EGFR-driven lung cancer, questions regarding the connections between drug tolerance and
acquired resistance and the therapeutic approaches to target residual tumor cells still remained to be addressed.

**Evolving from drug tolerant residual disease into acquired resistance.**

Five EGFR TKIs have been approved as the first-line treatment for *EGFR* mutant lung cancer, however acquired resistance inevitably occurs in patients receiving any of these drugs. On-target resistance conferring mutations in *EGFR* are the major mechanisms of resistance for all of the three generations of TKIs [30, 33, 34, 36]. Although previous studies have demonstrated that the T790M mutation, for example, can pre-exist in TKI-naïve tumors at low frequency and become enriched following treatment with first-generation TKIs, T790M- tumor cells can still survive TKI treatment as drug tolerant persisters and eventually develop such resistance mutation *de novo* [50]. Drug tolerant residual disease acts as an intermediate stage between pre-treatment tumor and acquired resistance and understanding its mutational landscape thus provides insights into how the resistance-conferring mutations come into existence. In either the PDX models or the GEMM that we studied in this thesis project, the on-target *EGFR* mutations were not identified in residual tumor cells and the overall mutational landscape remained largely unchanged compared to untreated tumors. Indeed, increasing evidence suggests that TKI tolerance in *EGFR* mutant lung cancer is caused by epigenetic mechanisms and transcriptional changes rather than resistance conferring mutations [41, 44, 47]. The absence of genetic drivers in TKI tolerant cells suggests that mutagenesis is a relatively inefficient process and it takes a long time for the right gene target to become mutated and for drug resistant clones to become dominant in recurrent tumors. In the mouse models described in previous chapters, the residual tumor cells were collected following only 1-2 months of TKI treatment, which was likely to be still at an early stage considering the time frame during which acquired resistance usually developed (~1.5 years for osimertinib). Further characterization of residual tumor cells collected longitudinally along the treatment course would
provide more evidence to elucidate when certain resistance-associated genetic events occur in the genome.

Phenotypic transition and histological transformation constitute another big category of mechanisms of TKI resistance, for which detailed evolutionary processes remain poorly understood. We identified upregulation of a neuroendocrine lineage factor ASCL1 in post-osimertinib residual tumors from one of the PDX models that we examined, albeit without significant neuroendocrine phenotypes. Given the role of ASCL1 as a neuroendocrine lineage transcription factor, the residual disease in that PDX might represent a transition state prior to the acquisition of neuroendocrine features. RB1 has been implicated as being commonly mutated in TKI-resistant tumors with SCLC transformation [114]. Therefore, one hypothesis is that histological transformation requires the presence of additional genetic alterations like RB1 loss and functional studies in PDX and cells are underway to prove this.

**Upstream regulators for ASCL1 in EGFR mutant lung cancer.**

Gene expression profiling through bulk RNA-sequencing and single-cell RNA-sequencing indicated enrichment of an ASCL1 gene signature in TKI tolerant residual cells and a pre-existing subpopulation of untreated tumor cells from one PDX in our cohort. Despite the fact that a subset of lung adenocarcinomas are characterized by ASCL1 positivity [93], it still remains unclear what upstream regulators and/or pathways mediate ASCL1 expression in these LUAD cases. In the context of SCLC, the histone demethylase KDM5A has been shown to sustain ASCL1 expression through suppression of NOTCH signaling activity, highlighting an epigenetic mechanism underlying the maintenance of the NE differentiation state [97]. As future directions, we would like to determine the role of KDM5A as well as other epigenetic regulators in mediating ASCL1 expression and whether or not they can affect lineage transition in PDXs and cell lines of EGFR mutant lung cancer. Ongoing experiments are designed to overexpress
ASCL1 in established human cell lines, followed by functional testing including target gene expression, drug sensitivity and clonogenic assay.

**Targeting drug tolerant residual disease to delay acquired resistance.**

One of the reasons for studying mechanisms of drug tolerance is to identify drug targets to repress or eliminate persister cells to delay or prevent the emergence of acquired resistance. Based on the results from this thesis project, we identified ASCL1 as a driver of TKI tolerance in mutant EGFR lung cancer. Studies in SCLC have provided clear evidence that ASCL1 is a critical regulator of NE differentiation and required for tumor initiation and maintenance [89, 92, 97, 98]. In addition, SCLC transformation is one of the common TKI resistance mechanisms. Therefore, targeting ASCL1 could be of potential use in treating patients with incomplete response to EGFR TKIs and upregulation of ASCL1 in residual cells. As ASCL1 functions as a transcription factor, developing its small molecule inhibitor is proved to be challenging. Alternative strategies include identifying druggable targets among upstream regulators and downstream gene targets of ASCL1. One canonical target gene regulated by ASCL1 encodes the anti-apoptotic protein BCL2, for which BH3-mimetic small molecule venetoclax has been approved to treat cancers dependent on the anti-apoptotic effects of BCL2 [115, 116]. The combination of EGFR TKIs and BCL2 inhibitors could be used as a follow-up treatment for leftover tumor cells after initial targeted therapy [45]. Another ASCL1-targeting strategy under development is to directly modulate its protein levels. Depending on the chemical nature of a gene product, it is possible to design protein degraders that target either the protein-of-interest (POI) itself or its binding partners that stabilize the POI. Recent genetic and chemical screening identified CDK2 as a stabilizer of ASCL1 protein and degradation of CDK2 by PROTAC reduced ASCL1 protein levels, providing a potential strategy to suppress ASCL1+ TKI tolerant residual cells [117].
BIBLIOGRAPHY


104. Shultz, L.D., et al., Human Lymphoid and Myeloid Cell Development in NOD/LtSz-scid IL2R


