Analysis by DNA barcoding of the heterogeneous response to anticancer drugs by different subpopulations of lung cancer cells

Houssein Chhouri

Normandie Université

THÈSE

Pour obtenir le diplôme de doctorat

Spécialité ASPECTS MOLECULAIRES ET CELLULAIRES DE LA BIOLOGIE

Préparée au sein de l'Université de Rouen Normandie

Analysis by DNA barcoding of the heterogeneous response to anticancer drugs by different subpopulations of lung cancer cells

Présentée et soutenue par
HOUSSEIN CHHOURI

Thèse soutenue le 10/11/2022 devant le jury composé de		
M. ANTONIO MARAVER	CHARGE DE RECHERCHE, UNIVERSITE DE MONTPELLIER	Rapporteur du jury
MME JULIE PANNEQUIN	DIRECTEUR DE RECHERCHE, UNIVERSITE DE MONTPELLIER	Rapporteur du jury
M. YOUSSEF ANOUAR	DIRECTEUR DE RECHERCHE, Université de Rouen Normandie	Membre du jury
M. VERA PANCALDI	CHARGE DE RECHERCHE, UNIVERSITE TOULOUSE 3 PAUL SABATIER	Membre du jury
M. OLIVIER SORIANI	PROFESSEUR DES UNIVERSITES, UNIVERSITE NICE SOPHIA ANTIPOLIS	Membre du jury
M. LUCA GRUMOLATO	MAITRE DE CONFERENCES, Université de Rouen Normandie	Directeur de thèse

Thèse dirigée par LUCA GRUMOLATO (Neuroendocrine EndOcrine and GeRminal DIfferentiation and Communication)

UNIVERSITÉ DE ROUEN NORMANDIE

ED 497 NBISE
Normande de Biologie Intégrative, Santé, Environnement

NorDiC

Acknowledgements

Ces travaux de thèse ont été réalisés sous la direction du Docteur Luca Grumolato, au sein du laboratoire de Différenciation et Communication Neuroendocrine, Endocrine et Germinale (NorDiC), INSERM 1239, dirigé par le Docteur Hervé Lefebvre.

Ces études ont été financées par une bourse de la Région Sud du Liban et une bourse fin de thèse de la Fondation pour la Recherche Médicale (FRM) et ont également bénéficié de l'aide des organismes suivants:

- L'Institut National de la Santé et de la Recherche Médicale (INSERM)
- L'Université Rouen Normandie
- L'Institut de Recherche et d'Innovation Biomédicale (IRIB)
- L'Ecole Doctorale Normande de Biologie Intégrative, Santé et Environnement (EDNBISE)
- L'Institut National du Cancer (PLBIO2017)
- La Ligue contre le cancer, comité de Seine maritime
- La Fondation ARC
- L'Agence National de la Recherche (ANR)
- La Fondation de France

First of all, I would like to thank my thesis supervisor, Dr. Luca Grumolato, for his excellent mentorship. For my almost four years in the lab as a PhD student, he has provided me with the best-in-class opportunity for scientific training and guided me through a challenging thesis project. I also appreciate that he could spend time with me in our regular individual meetings, which gave me the chance to keep my research in progress. This thesis would not have been possible without his endless support and encouragement. Special thanks to Dr. David Alexandre for all his invaluable input and suggestions throughout my PhD. Your help and advice was greatly appreciated. I also want to thank Dr. Youssef Anouar, for welcoming me to his team.

I am also very grateful to Dr. Julie Pannequin (Directrice de Recherche, Institut de Génomique Fonctionnelle de Montpellier), Dr. Antonio Maraver (Chargé de Recherche HDR, Institut de Recherche en Cancérologie de Montpellier), Dr. Vera Pancaldi (Chargé de Recherche HDR, Centre de Recherches en Cancérologie de Toulouse) and Pr. Olivier Soriani (Professeur des universités, Université Côte d'Azur), for the honor they have shown me by accepting to be a members of my PhD thesis committee.

The work presented in this thesis would not have been possible without collaboration. I would like to thank Dr. Vera Pancaldi and Dr. Nina Verstraete (Centre de Recherches en Cancérologie de Toulouse) for their help in the bioinformatics analysis. I want to thank Dr. Laura Holzhauer and Dr. Anke Deckers (Karlsruhe Institute of Technology, Germany) for providing the new anti-cancer compounds. I also thank Mr. Gaëtan Riou (Institut de Recherche et d'Innovation Biomédicale, Rouen) for his help with the FACS experiments.

Next, I want to thank my colleagues in Grumolato's Lab. I want to show special thanks to our lab manager Jiyoung Lee for her help in the experiments. Thank you for the warm welcome and for always being friendly and supportive. I also want to thank Lisa, Maria and Mirna for their support throughout my PhD. I would like to thank the members of the Neuroendocrine Plasticity and Pathophysiology team in particular Lina, Dorthe, Amine, Loubna, David, Christophe, Inès, Cédric, Isabelle, Christine, Maïté and Lydie.

Finally, I owe my deepest gratitude to my family. Thank you for encouraging and supporting me with your love throughout these years.

To my parents

Publications and Presentations

Publications

Brunet L, #Blanquer-Rosselló MDM, #Guernet A, **#Chhouri H**, Alexandre D, Kherrouche Z, Lee J, Yao S, Arabo A, Godefroy D, Jamme P, Vinchent A, Bérard C, Tuslane D, Arena S, Bardelli A, Coulouarn C, Cho, B. C., Aaronson SA, Cortot A, Anouar Y and Grumolato L: Prolonging lung cancer response to EGFR inhibition by inhibiting the selective advantage of resistant cells. In preparation (# equal contribution).

Swayden M, **Chhouri H**, Anouar Y and Grumolato L. (2020) Tolerant/Persister Cancer Cells and the Path to Resistance to Targeted Therapy. Cells 10.3390/cells9122601 [IF: 7.666].

Shaito A, Saliba J, Husari A, Harakeh M, **Chhouri H**, Hashem Y, Shihadeh A and El Sabban M. (2017) Electronic Cigarette Smoke Impairs Normal Mesenchymal Stem Cell Differentiation. Sci Rep 10.1038/s41598-017-14634-z [IF: 4.996].

Presentations

Brunet L, Blanquer- Rosselló M.M, Guernet A, **Chhouri H**, Khrrouche Z, Alexandre D, Yao S, Lee J, Tulasne D, Aaronson SA, Cortot A, Anouar Y and Grumolato L. Preventing acquired resistance to therapy by targeting tumor evolution: a new strategy to prolong lung cancer response to EGFR inhibition. 3rd edition Toulouse Onco Week, Toulouse, France. February 1–5, 2020.

Chhouri H, Guernet A, Brunet L, Blanquer Rosselló M.M, Alexandre D, Cortot A, Aaronson SA, Anouar Y and Grumolato L. CRISPR barcoding to investigate lung cancer resistance to targeted therapy in a context of intratumor heterogeneity. 4th EACR conference, Cancer Genomics, Cambridge, UK, June 23–26, 2019.

Abstract

Cancer is a dynamic disease, characterized by the co-existence of multiple subpopulations of tumor cells that can evolve in response to environmental changes. This intratumor heterogeneity has dramatic consequences, not only for cancer progression and metastatic spread, but also for treatment. Specific tyrosine kinase inhibitors are effective against non-small cell lung cancers harboring activating mutations of the epidermal growth factor receptor (EGFR), but the response is not durable, and cure remains elusive because of the inevitable development of acquired resistance. During therapeutic intervention, small subpopulations of resistant or tolerant cells are selected by virtue of their higher fitness, ultimately resulting in tumor relapse.

In this study, we used cellular barcoding to label several thousand populations of PC9 NSCLC cells and monitor their clonal dynamics in response to anti-cancer therapies. Our results revealed that some clones display a specific and predetermined response to treatment, indicating that cells that are primed to behave as tolerant or highly sensitive pre-exist in the original mass population. We extended these findings and showed that each type of anti-cancer drugs exerts a characteristic effect on the clonal architecture of the cell population, resulting in a specific barcode pattern that can be used as a signature to compare different compounds and investigate their mechanism of action. We have generated barcode profiles from 87 drugs targeting various pathways and used it to predict the mechanism of action of a new compound that can specifically inhibit NSCLC cell growth.

In the last part of the thesis, we took advantage of CRISPR/Cas9 technology to devise Barcode-Tracker, a new strategy to identify and isolate clones of interest from a mass population based on the recognition of a specific genetic barcode. Contrary to other approaches, Barcode-Tracker doesn’t involve drug selection of the cells and it can be used to sort clones according to their intrinsic capacity to behave in a particular manner in the presence of treatment, thus mimicking the response of a therapy-naïve tumor. By providing a better understanding of how treatment can shape clonal evolution, our studies should help to improve therapeutic strategies for NSCLC patients.

Keywords: Non-small cell lung cancer; EGFR; drug resistance; drug tolerance; anti-cancer therapies; cellular barcoding; CRISPR/Cas9; Barcode-Tracker.

Résumé

Le cancer est une maladie évolutive, caractérisée par l'existence d'un mélange complexe de plusieurs sous-populations cellulaires qui peuvent évoluer en réponse aux conditions environnementales. Cette hétérogénéité intratumorale a des conséquences extrêmement importantes, non seulement sur la progression tumorale, les métastases, mais aussi sur l'efficacité des traitements. Le cancer bronchique non à petites cellules (CBNPC) avec mutations activatrices de l'Epidermal Growth Factor Receptor (EGFR) est traité avec des inhibiteurs spécifiques de ce récepteur. Malheureusement, après une réponse favorable de plusieurs mois, ces tumeurs presque invariablement développent une résistance à ces médicaments. Au cours des traitements, des sous-populations de cellules résistantes ou tolérantes sont sélectionnées et vont émerger pour conduire à une rechute de la tumeur.

Dans cette étude, nous avons utilisé une approche d'étiquetage cellulaire pour marquer plusieurs milliers de populations de cellules PC9 de CBNPC avec des codes-barres génétiques et suivre leur évolution clonale en réponse aux thérapies anticancéreuses. Nos résultats ont révélé que certains clones présentent une réponse spécifique et prédéterminée en fonction du traitement, ce qui indique que le phénotype tolérant ou encore hautement sensible représente une propriété intrinsèque de certaines cellules qui préexistent au sein de la population cellulaire initiale. Nous avons également montré que chaque type de médicament anticancéreux exerce un effet caractéristique sur l'architecture clonale de la population cellulaire, ce qui entraîne un profil de code-barres spécifique qui peut être utilisé comme signature pour comparer différents composés et étudier leurs mécanismes d'action. Nous avons généré une collection de profils de codes-barres à partir de 87 composés connus, agissant sur des processus cellulaires différents, et les avons utilisés pour identifier le mécanisme d'action d'une nouvelle molécule capable d'inhiber spécifiquement la croissance des cellules de CBNPC.

Dans la dernière partie de la thèse, grâce à la technologie CRISPR/Cas9, nous avons développé Barcode-Tracker, une nouvelle stratégie pour identifier et isoler des clones d'intérêt à partir d'une population de cellules tumorales, basée sur la reconnaissance d'un code-barres génétique spécifique. Contrairement à d'autres approches, Barcode-Tracker n'utilise pas

d’inhibiteur pour sélectioner des cellules, et peut être utilisé pour purifier les clones selon leur capacité intrinsèque à se comporter d'une manière particulière en présence d'un traitement, imitant ainsi la réponse d'une tumeur naïve. Avec une meilleure compréhension de la manière dont le traitement anticancéreux peut affecter l’évolution tumorale, ces études devraient permettre une amélioration des stratégies thérapeutiques pour les patients atteints de CBNPC.

Mots clés: Cancer bronchique non à petites cellules; EGFR; résistance à la thérapie; tolérance au traitement, thérapies anticancéreuses; étiquetage cellulaire; CRISPR/Cas9; Barcode-Tracker.

Table of Contents

ACKNOWLEDGEMENTS 5
PUBLICATIONS AND PRESENTATIONS 9
ABSTRACT 11
RÉSUMÉ 13
TABLE OF CONTENTS 15
LIST OF FIGURES 19
LIST OF TABLES 21
LIST OF ABBREVIATIONS 23
INTRODUCTION 29
1. Cancer as an evolutionary process 31
1.1. Clonal evolutionary trajectories: linear and branched evolution 31
2. Intratumor heterogeneity 33
2.1. Genetic intratumor heterogeneity 34
2.2. Non-genetic intratumor heterogeneity 34
2.3. Clinical implication of intratumor heterogeneity 35
3. Non-small cell lung cancer: a model for intratumor heterogeneity 36
3.1. Histological classification 36
3.2. Staging 37
3.3. Treatments for NSCLC 38
3.3.1. Surgery 38
3.3.2. Radiotherapy 38
3.3.3. Chemotherapy 39
3.3.4. Immunotherapy 40

3.3.5. Targeted therapy 41
3.4. Oncogenic driver mutations in NSCLC 42
3.4.1. KRAS mutations 42
3.4.2. ALK rearrangement 44
3.4.3. ROS1 rearrangements 45
3.4.4. MET aberrations 45
3.4.5. BRAF mutations 47
3.4.6. EGFR mutations 47
3.5. EGFR targeted therapies 49
3.5.1. First generation EGFR-TKIs 49
3.5.1.1. Primary or intrinsic resistance to EGFR-TKIs 49
3.5.1.2. Acquired resistance 51
3.5.2. Second generation EGFR-TKIs 53
3.5.3. Third generation EGFR-TKIs 53
3.5.3.1. Resistance to third generation EGFR inhibitors 55
3.5.4. Fourth generation EGFR-TKIs 55
3.5.5. EGFR degraders 56
4. Drug tolerant persister states 57
4.1. Persisters in NSCLC 57
4.2. Molecular characteristics of drug tolerance in NSCLC 58
4.2.1. Epigenetic modifications 58
4.2.2. Reactivation of EGFR signaling and up-regulation of other pathways 59
4.2.3. Metabolic reprogramming 61
4.3. Origin of DTP cells: Darwinian selection or Lamarckian induction? 62
5. Cellular barcoding 64
5.1. Brief history 64
5.2. Strategies for cellular barcoding 65
5.2.1. Cre-mediated recombination 66
5.2.2. Transposon-based approaches 67
5.2.3. CRISPR-mediated approaches 68

5.2.4. Lentiviral barcodes to investigate intratumor heterogeneity 73
5.2.5. CRISPR-barcoding 77
OBJECTIVE OF THE THESIS 79
RESULTS 81
1. Distinct predetermined fates emerge from a heterogeneous cell population in response to anti-cancer therapy 83
1.1. Diverse sensitivity to anti-cancer drugs by randomly isolated NSCLC clones 83
1.2. Functional characterization of the PC9-derived clones 86
1.3. Highly complex CRISPR-barcodes reveal predetermined phenotypes in response to different treatments 88
1.4. Potential involvement of epigenetic modifications on drug response 90
2. High throughput screen to investigate the heterogeneous response of different subpopulations of cancer cells to anti-cancer therapies 93
2.1. Generation of a highly complex barcode library to label PC9 cells 93
2.2. The barcode profile of different compounds is consistent across experiments 94
2.3. Distinct pre-existing subpopulations emerge in response to different treatments 96
2.4. Investigate the clonal dynamics of PC9 cells in response to different concentration of EGFR-TKIs 99
2.5. Comparison of our dataset with other drug similarity scores 100
2.6. Comparison of the barcode profiles to identify the mechanism of action of new anti-cancer drugs 102
3. Isolation and characterization of individual clonal lineages from a heterogeneous cell population 105
3.1. Barcode-Tracker workflow 105
3.2. Components of the Barcode-Tracker approach 106
3.2.1. dCas9-activator 106
3.2.2. Barcode-Tracker lentiviral vector 108
3.3. Optimization of the Barcode-Tracker approach 109
3.4. A toggle switch strategy for the Barcode-Tracker system 111
3.5. Components of the Barcode-Tracker toggle switch 112
3.5.1. Tet repressor 112

3.5.2. Cumate repressor 112
3.5.3. dCas9-repressor 112
3.6. Overview of the Barcode-Tracker toggle switch constructs 113
3.7. Workflow of the Barcode-Tracker toggle switch 114
3.8. Optimization of the Barcode-Tracker toggle switch approach 116
DISCUSSION 119
1. Elucidate the global effects of anticancer treatment on the clonal architecture of a mass population of NSCLC cells 121
2. Identify the mode of action of new anti-cancer agent 125
3. Isolate and characterize specific cell subpopulations displaying an intrinsic low or high sensitivity to treatment 127
CONCLUSIONS AND PERSPECTIVES 129
METHODS 133
1. Cell culture 135
2. Generation of PC9 single-cell subclones and drug-tolerant persisters 135
3. Cell viability assays 135
4. Western Blot 136
5. DNA damage assay 136
6. RNA extraction and qPCR 137
7. CRISPR-barcoding 137
8. Lentiviral barcode library construction 138
9. Lentivirus production and barcoding of PC9 cells 139
10. Drug screen 139
11. Bioinformatics analysis of the drug screen data 140
12. Barcode-Tracker activator constructs and cell line generation 140
13. Construction of Barcode-Tracker toggle switch plasmids 141
14. Generation of stable cell line expressing the toggle switch circuits 142
REFERENCES 153

List of Figures

Figure 1. Branched versus linear evolution 32
Figure 2. Tumor heterogeneity 33
Figure 3. A plan for stereotactic body radiation therapy to treat a patient with early stage NSCLC 39
Figure 4. Oncogenic mutations in NSCLC 42
Figure 5. RAS mediated intracellular signal transduction pathways 43
Figure 6. EML4-ALK rearrangements 44
Figure 7. MET alterations in NSCLC 46
Figure 8. EGFR signaling pathways in NSCLC 48
Figure 9. Schematic representation for the acquisition of drug-tolerant phenotype in NSCLC as described by Sharma and colleagues 58
Figure 10. Models for the emergence of drug-tolerant cells 64
Figure 11. Cre-loxP recombination for in vivo barcoding 67
Figure 12. Cellular barcoding through a Sleeping Beauty transposon system 68
Figure 13. CRISPR-based in vivo barcoding 71
Figure 14. Lentiviral-based in vitro barcoding 75
Figure 15. CRISPR-barcoding as a strategy for in vitro barcode delivery 78
Figure 16. Different sensitivities of PC9-derived clones to anti-cancer drugs 84
Figure 17. Cell viability assays showing the response of PC9-C1/C3/C5 to anti-cancer therapy 85
Figure 18. Characterization of PC9-derived clones 87
Figure 19. Clonal dynamics of PC9 cells in response to anti-cancer treatment using highly complex CRISPR-barcodes 89
Figure 20. Effects of TSA pre-treatment on the clonal evolution of PC9 cells 91
Figure 21. Lentiviral barcode strategy to investigate the variability of drug response within individual cancer cells 94
Figure 22. The drug response of the cell population is stable over time 95

Figure 23. Drug-drug correlation clustered heatmap 97
Figure 24. Effect of different drugs targeting the EGFR and its downstream signaling pathways on the clonal evolution of PC9 NSCLC 99
Figure 25. Clonal dynamics of PC9 NSCLC in response to different concentrations of osimertinib 100
Figure 26. Comparison between the drugs used in our screen with other drug similarity scores 101
Figure 27. Annotated drug clustering based on barcodes signatures 103
Figure 28. Screen of novel small molecules on the growth of PC9 and SKNAS cells 104
Figure 29. Diagram describing the Barcode-Tracker approach 106
Figure 30. RNA-guided transcriptional activation 107
Figure 31. Diagram illustrating the Barcode-Tracker reporter based on the dCas9-SAM system 108
Figure 32. Optimization of the Barcode-Tracker system 110
Figure 33. Diagram illustrating the concept of a genetic toggle switch 111
Figure 34. The basic components of our synthetic gene network 113
Figure 35. Schematic diagrams of the toggle switch constructs 114
Figure 36. Schematic representation of the Barcode-Tracker-toggle switch strategy 115
Figure 37. Generation PC9 cells expressing the toggle switch components 117
Figure 38. Distinct response to osimertinib and pemetrexed 122
Figure 39. PC9 barcoded cells transplanted into immunodeficient mice 123
Figure 40. Predefined response to dual EGFR/MEK inhibition 124

List of Tables

Table 1. TNM stage grouping for NSCLC 38
Table 2. Classification of chemotherapeutic agents based on their mechanism of action..... 40
Table 3. Comparison of EGFR-TKIs..... 50
Table 4. Overview of the currently available techniques for cellular barcoding..... 72
Table 5. Classification of the drugs used in the screen based on their mechanism of action..... 96
Table 6. List of oligonucleotides used for CRISPR-Barcoding..... 143
Table 7. List of oligonucleotides used to generate the lentiviral barcode library..... 145
Table 8. Compounds used for this study..... 145
Table 9. Primers used for deep sequencing for lentiviral barcode library..... 148
Table 10. List of oligonucleotides used to generate the Barcode-Tracker constructs 150
Table 11. List of sgRNA used to knock in repressor 2 into the AAVS1 locus. 151

List of Abbreviations

AAVS1: Adeno-associated virus integration site 1 locus

ATP: Adenosine triphosphate

AKT: AKT serine/threonine Kinase 1

ALDH: Aldehyde dehydrogenase

ALK: Anaplastic lymphoma kinase

ARE: Anti-repressor elements

AURKA: Aurora kinase A

AXL: Receptor tyrosine kinase AXL

BIM: Bcl-2-like protein 11

BRAF: V-Raf Murine sarcoma viral oncogene homolog B

CSCs: Cancer stem cells

CTLA4: cytotoxic T-lymphocyte-associated antigen 4

CIN: Chromosomal instability

CRISPR: Clustered regularly interspaced short palindromic repeats

CARLIN: CRISPR array repair lineage tracing

CaTCH: CRISPRa tracing of clones in heterogeneous cell populations

CuO: Cumate responsive elements

CymR: Cumate repressor

Cum: Cumic acid

CMV: Cytomegalovirus

DSB: Double-stranded break

Dox: Doxycycline

DTP: Drug-tolerant persisters

DTEP: Drug-tolerant expanded persisters

DUSP6: Dual specificity phosphatase 6

dCas9: dead Cas9

ER: Endoplasmic reticulum

EZH2: Enhancer of zeste homolog 2

EGF: Epidermal growth factor

EGFR: Epidermal growth factor receptor

EMT: Epithelial-to-mesenchymal transition

EMA: European Medicines Agency

ERK: Extracellular signal-regulated kinase

FAO: Fatty acid oxidation

FISH: Fluorescence *in situ* hybridization

FACS: Fluorescence-activated cell sorting

FAK: Focal adhesion kinase

FDA: Food and Drug Administration of the United States

GESTALT: Genome editing of synthetic target arrays for lineage tracing

GPX4: Glutathione peroxidase 4

GAP: GTPase-activating protein

GDP: Guanosine biphosphate

GTP: Guanosine triphosphate

GFP: Green fluorescent protein

GEM: Gel bead in emulsion

gDNA: Genomic DNA

HRAS: Harvey rat sarcoma viral oncogene homolog

HSCs: Hematopoietic stem cells

HGF: Hepatocyte growth factor

HDAC: Histone deacetylase

HDR: Homology-directed repair

HER2: Human epidermal growth factor receptor 2

HER3: Human epidermal growth factor receptor 3

HSB: Hyperactive Sleeping Beauty transposase

hgRNA: Homing gRNA

ICBs: Immune checkpoint blockades

IGF-1R: Insulin-like growth factor 1 receptor

IASLC: International association for the study of lung cancer

JAK: Janus Kinase 2

KRAS: Kirsten rat sarcoma viral oncogene homolog

KDM5A: Lysine K-specific demethylase 5A

KRAB: Krüppel-associated box

Lgr5: Leucine rich repeat containing G-protein coupled receptor 5

LARRY: Lineage and RNA recovery

LINNAEUS: Lineage tracing by nuclease-activated editing of ubiquitous sequences

LINE-1: Long interspersed repeat element 1

LTR: Long terminal repeat

MET: Hepatocyte growth factor receptor

MEK: Mitogen-activated protein kinase kinase

MAPK: Mitogen-activated protein kinases

MARC1: Mouse for actively recording cells 1

MOI: Multiplicity of infection

MeCP2: Methyl CpG binding protein 2

mSCRIBE: Mammalian synthetic cellular recorders integrating biological events

mTOR: Mammalian target of rapamycin

NAC: N-acetylcysteine

NRAS: Neuroblastoma RAS viral oncogene homolog

NGS: Next-generation sequencing

NHEJ: Non-homologous end joining

NSCLC: Non-small cell lung cancer

NCI-60: National cancer institute panel of 60 human cancer cell lines

ORR: Overall response rates

OXPHOS: Oxidative phosphorylation

PD1: Programmed cell death-1

PDL1: Programmed cell death-ligand 1

PKC: protein kinase C

PTEN: Phosphatase and tensin homolog

PIP-3: Phosphatidylinositol 3,4,5-triphosphate

PIP2: Phosphatidylinositol 4,5-bisphosphate

PI3K: Phosphatidylinositol 4,5-bisphosphate 3-kinase

PIK3CA: Phosphatidylinositol 4,5-bisphosphate 3-kinase, catalytic subunit alpha

PLCγ: Phospholipase Cγ

PolyA: Polyadenylation signal

PCR: Polymerase chain reaction

PFS: Progression-free survival

PROTACs: Proteolysis targeting chimeras

PAM: Protospacer-adjacent motif

PRISM: Profiling relative inhibition simultaneously in mixtures

ROS: Reactive oxygen species

RTK: Receptor tyrosine kinase

RFP: Red fluorescent protein

ROS1: ROS1 proto-oncogene 1 receptor tyrosine kinase

STAT: Signal transducer and activator of transcription

SCLC: Small-cell lung cancer

SOS: Son of sevenless

SPRY4: Sprouty RTK signaling antagonist 4

SBRT: Stereotactic body radiation therapy

SAM: Synergistic activation mediator

stgRNA: Self targeting guide RNA

shRNA: Short hairpin RNA

scRNA-Seq: Single cell RNA sequencing

sgRNA: Single-guide RNA

ssODN: Single-stranded DNA oligonucleotide

TMB: Tumor mutation burden

T2A: Self-cleaving peptides

TNM: TNM tumor, noduli, metastasis

TEAD: Transcriptional enhanced associate domains

TGF-α: Transforming growth factor alpha

Tn: Transposon element

TSA: Trichostatin A

TME: Tumor microenvironment

TKIs: Tyrosine kinase inhibitors

TetO: Tet operator

tTR: Tet repressor

UTR: Untranslated region

Wnt: Wingless integration site family

WPRE: Woodchuck hepatitis virus posttranscriptional regulatory element

YAP: Yes1-associated protein

ZEB: Zinc-finger E-box-binding

Introduction

1. Cancer as an evolutionary process

In his seminal work "*On the Origin of the Species"*, Charles Darwin described an evolutionary framework to understand the genesis and diversification of the different types of plants and animals that live on earth (Darwin, 1859). This theory of evolution by natural selection is based on three basic principles:

a. Variation: high diversity of phenotypes amongst individuals within a population
b. Heredity: phenotypic variation is heritable and can be passed to future generations
c. Selection: phenotypes are associated with different fitness, which results in different survival and/or reproduction rates

The original concept that cancer can follow a similar evolutionary model has been attributed to Peter Nowell. In 1976, he proposed that tumors arise from a single neoplastic cell, and progression depends on a process of clonal evolution driven by the sequential accumulation of a series of new mutations capable of improving cell fitness, thus resulting in the selection of increasingly aggressive subpopulations (Nowell, 1976). In the last four decades, the idea of cancer as an evolutionary process has been broadly supported by an overwhelming body of evidence, and the concept of clonal selection and the resulting intratumor heterogeneity have been formally demonstrated by next-generation sequencing (NGS) analysis of human tumors (Welch et al., 2012; Yates et al., 2015).

1.1. Clonal evolutionary trajectories: linear and branched evolution

The tremendous advances in sequencing technologies have been crucial in uncovering the genetic landscape of different types of tumors. The type of aberrations ranges from single nucleotide substitutions to chromosomal translocations and their number is highly variable and depends on the tumor type (Martincorena and Campbell, 2015). According to their contribution for cancer development, these mutations can be classified as either driver or passenger: the former provide a selective advantage to cancer cells and they can play a significant role in promoting tumor growth and progression, whereas the latter are phenotypically silent and confer a negligible benefit on cellular fitness (Stratton et al., 2009).

Since Nowell's landmark paper, different models of tumor evolution have been proposed (Davis et al., 2017). Linear evolution implies that, when a new driver mutation arises in a tumor cell, this clone replaces the other cell populations through selective sweeps (Yates and Campbell, 2012). According to this model, tumors are expected to be mostly homogenous, composed of clonally identical cells (Fig. 1A). Branched evolution represents an alternative model, where multiple subclones can diverge in parallel from a common ancestor through the acquisition of different driver mutations (Fig. 1B) (Yates and Campbell, 2012). This mode of evolution would result in the coexistence of several distinct clonal subpopulations within the tumors. While some cancers, such as acute lymphoid leukemia (Anderson et al., 2011), can evolve through linear steps, the large majority grows following a branched pattern, as demonstrated by the high levels of genetic intratumor heterogeneity, found for example in breast (Yates et al., 2015), ovarian (McPherson et al., 2016), colorectal (Sottoriva et al., 2015) and prostate cancers (Gundem et al., 2015).

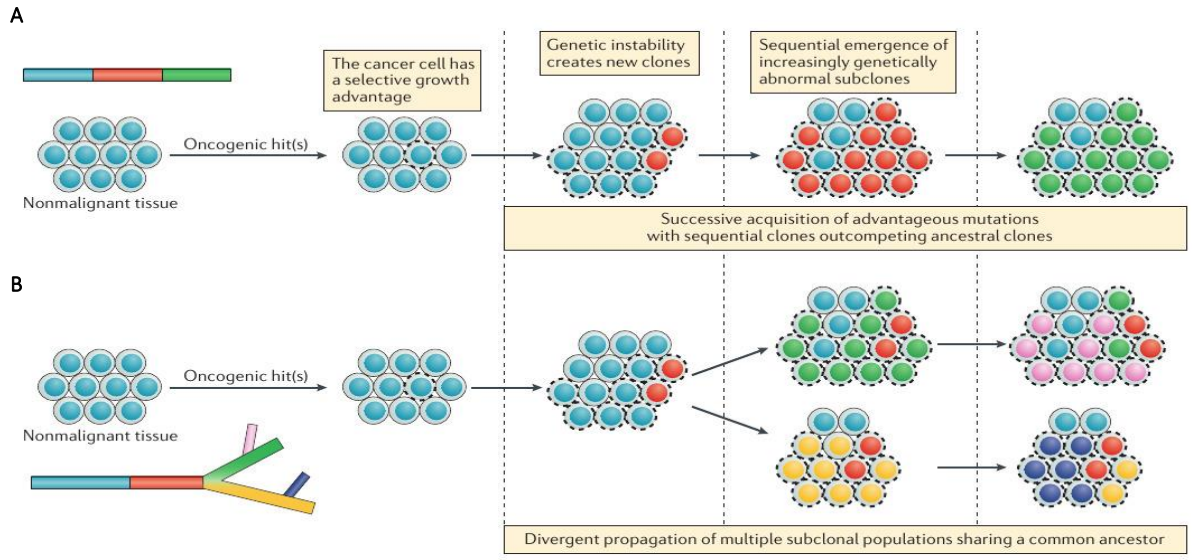


Figure 1. Branched versus linear evolution. Schematic illustration of the tumor evolution models. (A) The linear model is characterized by the sequential accumulation of mutations with each new clone outcompeting the previous one. (B) In the branched model, mutations randomly accumulate in different cells, leading to the formation of multiple subclonal populations. From Dagogo-Jack and Shaw, Nat. Rev. Clin. Oncol., 2017.

2. Intratumor heterogeneity

Tumor heterogeneity may be divided into three general categories: interpatient, intermetastatic and intratumor heterogeneity. Interpatient heterogeneity refers to differences between patients harboring tumors of the same histological type (Fig. 2A), while intermetastatic heterogeneity describe the variation between different metastatic lesions (Fig. 2B) (Vogelstein et al., 2013). The coexistence of molecularly and phenotypically distinct populations of cancer cells within a tumor is known as intratumor heterogeneity (Fig. 2C). This heterogeneity was first described nearly 100 years ago by Theodor Boveri, who reported morphological differences between tumor cells when observed under the microscope, well before higher resolution cytogenetic approaches documented the diversity of structural aberrations in tumors (Hansford and Huntsman, 2014). The fact that tumors are constituted of a mosaic mixture of subclonal cell populations with different phenotypes has important implications for cancer progression, metastasis and resistance to therapy.

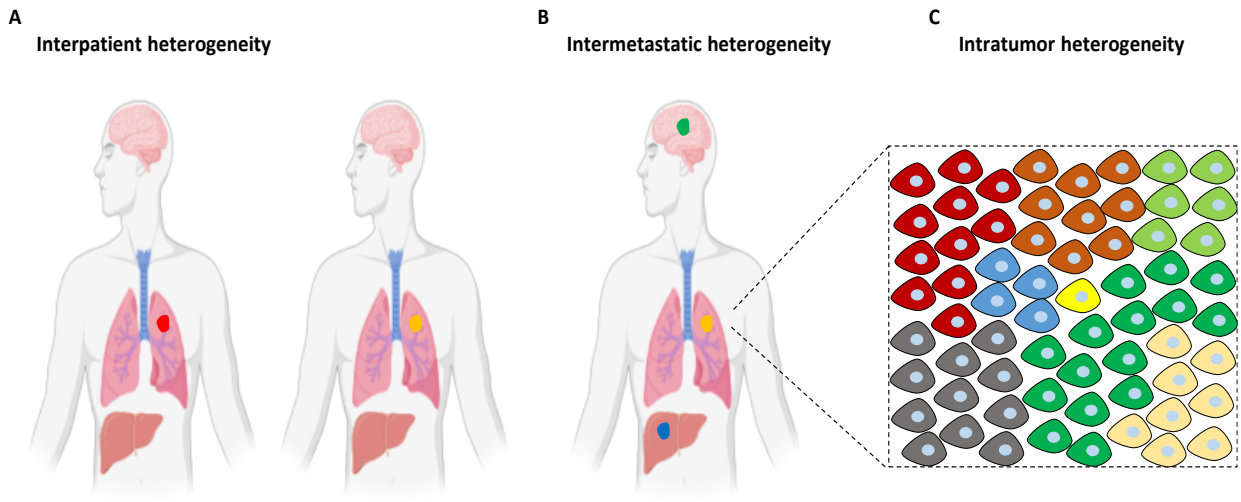


Figure 2. Tumor heterogeneity. Schematic representation of different types of tumor heterogeneity. (A) The variation among tumors from different patients is termed interpatient heterogeneity. (B) The existence of various tumor clones in different metastatic lesions of the same patient is defined as intermetastatic heterogeneity. (C) Intratumor heterogeneity refers to the presence of heterogeneous subclones within individual tumors.

2.1. Genetic intratumor heterogeneity

Malignant transformation of normal cells is a multistep process characterized by the sequential acquisition of driver mutations of genes that are generally involved in the regulation of cell proliferation and survival. These genes fall into two categories: oncogenes and tumor suppressor genes (Vogelstein and Kinzler, 2015). Oncogenes are genes whose activation promotes growth of cancer cells (Hall, 1984), while tumor suppressors are genes that function, in normal conditions, to slow down cell division, repair DNA mismatches or induce cell death, but are inactivated by the mutation (Yeo, 1999). As a tumor progresses, cancer cells typically acquire more and more mutations. This is often facilitated by inactivating mutations of genes involved in DNA repair, which results in genomic instability. This term refers to an enhanced tendency of tumor cells to acquire new mutations during cell division. The accumulation of these molecular alterations contributes to genetic and phenotypic diversity among cancer cells. Genomic instability can involve chromosomal alterations, as well as point mutations (Almendro et al., 2013). Chromosomal instability (CIN) describes an increased rate of change in chromosome number or structure, including a gain or loss of whole chromosomes, translocations, inversions and deletions. CIN is a major driver source of intratumor genetic heterogeneity, allowing tumor cells to adapt and survive environmental pressures (Gisselsson et al., 2000).

2.2. Non-genetic intratumor heterogeneity

The mechanisms that enable a single fertilized egg to give rise to hundreds of different cell types, tissues, and organs to make a complex organism is one of the greatest mysteries of developmental biology. This diversity arises from genetically identical cell populations and depends on a spatiotemporal regulation of gene expression during development. This cellular and phenotypic diversity within the developing embryo represents a typical example of non-genetic heterogeneity. Similarly, cancer cells with the same genetic profile can be phenotypically heterogeneous because of variations in the expression of genes that can have a profound impact of cell physiology and differentiation.

Cancer stem cells (CSCs) are defined as a subset of cancer cells that possess indefinite self-renewal ability to initiate and maintain tumor growth, based on epigenetic mechanisms (DNA methylation, histone modification and chromatin architecture) (Kreso and Dick, 2014). These cells can divide to produce either a daughter CSCs or progeny that will proliferate and differentiate into different cell lineages that form the bulk of the tumor (Prasetyanti and Medema, 2017). The first evidence for the existence of CSCs came from xenograft studies with acute myeloid leukemia cells, showing that only a small subpopulation of $CD34^{+}/CD38^{-}$ cells could recapitulate the original tumor heterogeneity (Bonnet and Dick, 1997). CSCs has also been described in multiple types of solid tumors, including breast (Al-Hajj et al., 2003), colon (O'Brien et al., 2007) and pancreas (Li et al., 2007). Even if there is still a debate on the specificity of the markers that can be used to identify and isolate CSCs and it has been shown that, in some cases, differentiated cancer cells can revert to CSCs (Vlashi and Pajonk, 2015), these cells undoubtedly represent an example of cancer non-genetic heterogeneity.

Another source of heterogeneity can derive from the tumor microenvironment (TME), a highly heterogeneous milieu consisting of various types of cells, including immune cells, endothelial cells, fibroblasts and adipocytes. Interactions between cancer cells and the associated stroma, as well as changes in the level of hypoxia or acidity, affect tumor development by establishing conditions that can favor cancer progression (Marusyk et al., 2012). The TME is a critical player for the selection process and fate determination of cancer cells. For example, it has been shown that hypoxia regulates the activity of histone modifying enzymes, which can promote an invasive phenotype in distinct subpopulations of cancer cells (Chen et al., 2011; Ramón Y Cajal et al., 2020; Wilson and Hay, 2011). Thus, the TME can contribute to the generation of cellular diversity, and is considered as a major determinant of non-genetic intratumor heterogeneity.

2.3. Clinical implication of intratumor heterogeneity

Genetic and non-genetic intratumor heterogeneity is a hallmark of nearly all types of cancer, contributing to the extraordinary capacity of tumor cells to adapt to environmental conditions. Neoplasm can be viewed as a complex dynamic ecosystem where heterogeneity provides high levels of plasticity, promoting tumorigenesis and disease progression

(McGranahan and Swanton, 2015). Within a tumor, cancer cells interact, cooperate and compete with each other, as well as with different cells in their microenvironment, for the same resources, space and nutrients in order to proliferate and survive (Merlo et al., 2006; Tabassum and Polyak, 2015). Different subpopulations of cancer cells can cooperate with each other to overcome many biological constraints during their development. It has been shown that this interaction can play an important role in metastatic spread. By mixing tumor cells expressing either green fluorescent protein or mCherry, Aceto et al. generated multicolored primary breast tumors in mice. In the bloodstream of these mice, the authors found oligoclonal clusters of cancer cells, bound to each other by cell-to-cell junctions. Compared to single circulating tumor cells, these clusters showed much higher metastatic potential (Aceto et al., 2014), indicating that cooperation between distinct subpopulations can play an important role in the spread of cancer cells.

Intratumor heterogeneity has also major implications in cancer response to treatment. The selective pressure represented by therapy drives the emergence of the cell subpopulations displaying low or no sensitivity to the treatment, ultimately leading to the acquisition of drug resistance and tumor relapse. This point will be discussed in more details in the next section.

3. Non-small cell lung cancer: a model for intratumor heterogeneity

Lung cancer is one of the most commonly diagnosed tumors and the leading cause of cancer-related deaths worldwide, claiming an estimated 1.8 million lives annually. In France, there were around 40,000 new cases and 37,000 deaths in 2020 according to the World Health Organization. Tobacco is the major cause of lung cancer, estimated to account for approximately 80% of all cases. In addition to active inhalation, secondhand smoke can also play a role in lung carcinogenesis. Besides tobacco, other known risk factors include environmental or occupational exposure to radon gas, air pollution and asbestos (Malhotra et al., 2016).

3.1. Histological classification

Histologically, lung cancer can be divided into two groups; small-cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). SCLCs constitute approximately 15% of the cases and

are believed to arise from neuroendocrine cells in the airways. They are characterized by the expression of common neuroendocrine markers, such as synaptophysin and chromogranin A. SCLC is a very aggressive malignancy with early and frequent metastasis and is strongly related to cigarette smoking. Bi-allelic inactivation of tumor suppressors TP53 and RB1 are the most frequent genetic alterations in this type of malignancy (Gazdar et al., 2017).

NSCLC represents the remaining 85% of lung cancers and includes three major histological subtypes: adenocarcinoma, squamous cell carcinoma and large cell carcinoma (Herbst et al., 2018). Adenocarcinoma is the most common type of NSCLC, accounting for 40-50% of the cases. It usually develops in the periphery of the lung and originate from type II alveolar cells, which secrete mucus and other substances. These tumors show features of glandular differentiation and are characterized by the expression of the thyroid transcription factor 1 and cytokeratin 7 (Chen et al., 2014). Squamous cell carcinoma is responsible for about 30% of NSCLCs. This type of cancer arises most frequently in the proximal bronchi and it generally has the strongest association with smoking (Kenfield et al., 2008). Large cell carcinoma is the least frequent subtype of NSCLC and accounts for 10-15% of all cases. This type of carcinoma may begin anywhere in the lungs and tends to grow quickly. It appears large and relatively undifferentiated, and it is diagnosed by exclusion of other histological NSCLC subtypes (Davidson et al., 2013).

3.2. Staging

Staging is used to describe the extent of cancer in the body. Like all solid tumors, NSCLC is clinically staged based on the TNM (tumor, noduli, metastasis) system proposed by the International Association for the Study of Lung Cancer (IASLC) (Table 1). T refers to the size of the primary tumor and ranges from 0 to 4. T0 means that the tumor can't be found or is *in situ*; while T1-T4 are used to identify the extension of the tumor (higher T numbers mean a larger tumor). N followed by a number from 0 to 3 indicates whether the cancer has spread to lymph nodes: N0 is used for localized tumors and N3 means an extensive regional dissemination. Finally, M describes the presence of distant metastases and is either 0 (no metastasis), or 1 (metastasis has occurred) (Goldstraw et al., 2016). The proper determination of stage for NSCLC patients is crucial, since it will guide treatment decisions.

Table 1. **TNM stage grouping for NSCLC.**

N / T	0	1	2	3
1a ≤ 1 cm	IA1	IIB	IIIA	IIIB
1b > 1–2 cm	IA2	IIB	IIIA	IIIB
1c > 2–3 cm	IA3	IIB	IIIA	IIIB
2a > 3–4 cm	IB	IIB	IIIA	IIIB
2b > 4–5 cm	IIA	IIB	IIIA	IIIB
3 > 5–7 cm	IIB	IIIA	IIIB	IIIC
4 > 7 cm	IIIA	IIIA	IIIB	IIIC
M1a (any T and N)	IVA	IVA	IVA	IVA
M1b (any T and N)	IVA	IVA	IVA	IVA
M1c (any T and N)	IVB	IVB	IVB	IVB

M1a: separate tumor nodule(s) in a contralateral lobe; tumor with pleural nodules or malignant pleural/ pericardial effusion. M1b: single extrathoracic metastasis in a single distant organ. M1c: multiple extrathoracic metastases in one or more distant organs.

3.3. Treatments for NSCLC

3.3.1. Surgery

Surgical resection of the primary tumor remains the best curative option for NSCLC patients diagnosed with an early stage disease (stage I, II and IIIA). This is possible if the patient is in good physical conditions and if the lesions are localized and can be completely removed. After surgery, patients usually receive adjuvant therapy, such as chemotherapy or radiotherapy, to eliminate any tumor cells left behind and to reduce the risk of relapse (Howington et al., 2013).

3.3.2. Radiotherapy

Radiation therapy is the use of high energy x-rays or other particles to destroy tumor cells. Patients with early-stage NSCLC who have a single small nodule in the lung without any metastases to nearby lymph nodes and who are not candidates for surgical resection due to comorbidities are often treated with radiotherapy. The most common type of radiation treatment is called external-beam radiation therapy, such as stereotactic body radiation therapy (SBRT). This technique uses CT scans or PET scans to precisely locate the tumor, which enables the

delivery of few high-dose fractions of radiation to a small target area (Fig. 3), minimizing the risk of affecting healthy regions of the body (Cuaron et al., 2013; Navarria et al., 2013; Timmerman et al., 2010).

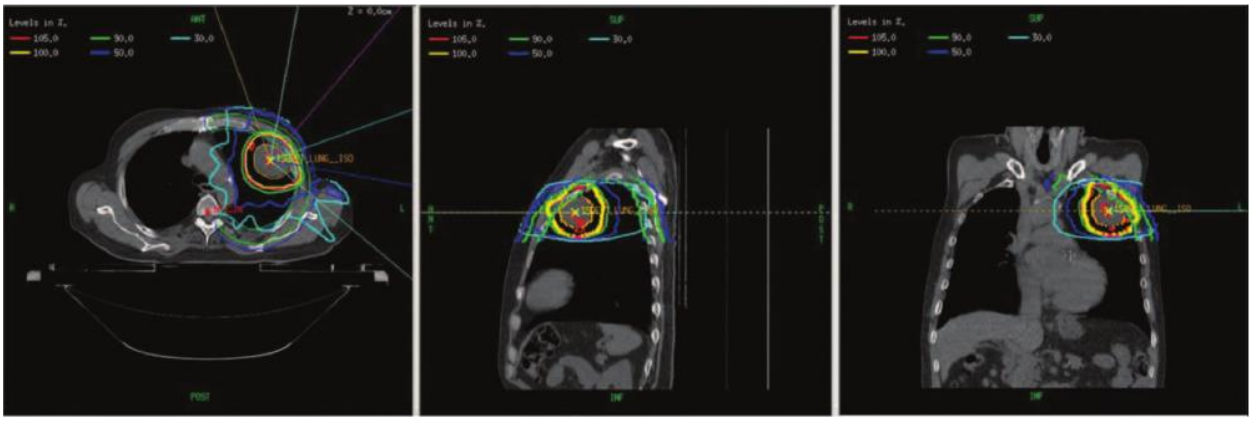


Figure 3. A plan for stereotactic body radiation therapy to treat a patient with early stage NSCLC. Delivery of very high doses of radiation to the tumor sites with high precision. From Cuaron et al., J. Thorac. Oncol. 2013.

3.3.3. Chemotherapy

Chemotherapy is the most common option for patients with advanced NSCLC who are no longer amenable to local therapies (Zappa and Mousa, 2016). Chemotherapeutic agents usually work by killing rapidly proliferating cells. These drugs can be divided into several classes according to their mechanisms of action, including alkylating agents, antimetabolites, topoisomerase inhibitors and mitotic inhibitors (Table 2) (Anand et al., 2022; Bukowski et al., 2020). For lung cancer patients with good performance status, the first-line treatment consists of a platinum-based doublet chemotherapy, *e.g.* cisplatin or carboplatin in combination with a third generation cytotoxic drug, such as pemetrexed, paclitaxel, docetaxel, vinorelbine or gemcitabine. In contrast, patients with poor performance status are not eligible to receive drug combination regimens but they may benefit from single agent therapy (Novello et al., 2016). Due to the lack of specificity against malignant cells, these chemotherapy agents can cause undesirable side effects on multiple tissues and organs, including gastrointestinal tract and bone marrow. In addition, the use of chemotherapy may induce multi-drug resistance, thus limiting the duration of the response to the treatment. Chemotherapy is still the major option for patients for whom no driver mutations are identified or targeted therapies are not available.

Table 2. Classification of chemotherapeutic agents based on their mechanism of action. Adapted from Anand et al., Genes Dis. 2022.

Class	Subclass	Representative drug	Approval	Mechanism of action
Alkylating	Platinum drugs	Cisplatin, carboplatin	FDA/EMA approved*	Formation of DNA crosslinks and bulky adducts inhibits DNA synthesis
	Mustard gaz derivatives	Cyclophosphamide	FDA/EMA approved	
Mitotic	Taxane	Paclitaxel, docetaxel	FDA/EMA approved*	Inhibit spindle mitotic formation
	Vinca alkaloids	Vinorelbine		
Antimetabolites	Purine antagonist	6-Mercaptopurine	FDA/EMA approved	Interfere with the production of purine and pyrimidine nucleotide
	Pyrimidine antagonist	Gemcitabine	FDA/EMA approved*	
	Antifolate	Pemetrexed		Inhibit the folate pathway
Topoisomerase	Topoisomerase I	Irinotecan	FDA/EMA approved**	Interfere with the action of topoisomerase enzymes
	Topoisomerase II	Etoposide		

*Approved for NSCLC, **Approved for SCLC, EMA: European Medicines Agency, FDA: Food and Drug Administration of the United States.

3.3.4. Immunotherapy

In recent years, immune checkpoint blockades (ICBs), including cytotoxic T-lymphocyte-associated antigen 4 (CTLA4), programmed cell death-1 (PD1) and PD-ligand 1 (PDL1) inhibitors have become routinely part of the clinical approach for management of NSCLC. Under normal physiologic conditions, immune checkpoints function as negative feedback to regulate the immune responses following T-cell activation. It has been shown that tumor cells can induce the expression of checkpoints proteins to evade destruction by the immune system. For example, cancer cells expressing PDL1 interact with PD1 on T cells to inhibit their activation. ICBs act by blocking the inhibitory interactions between these checkpoints molecules allowing activation of the immune system, which can identify and eliminate tumor cells (Sharma and Allison, 2015). So far, the best predictive marker for response of ICBs is PDL1 expression in the tumor (Hanna et al., 2020). Another promising marker is tumor mutation burden (TMB). It has been shown that patients with higher TMB might benefit more from immunotherapy, as non-synonymous mutations produce more neoantigens that facilitate recognition of tumor cells by the immune system (Rizvi et al., 2015).

Three classes of ICBs targeting CTLA4, PD1 or PDL1 have been approved for NSCLC. Ipilimumab (Yervoy, Bristol Myers Squibb), a monoclonal antibody directed against CTLA4, was the first FDA approved ICBs for the treatment of metastatic melanoma. Recently, ipilimumab in combination with anti-PD1 and platinum-doublet chemotherapy was approved to treat patients with advanced NSCLC (Vellanki et al., 2021). The PD1 inhibitors nivolumab (Opdivo, Bristol-Myers Squibb) and pembrolizumab (Keytruda, Merck) have both shown good results in NSCLC, with prolonged overall survival and reduced toxicity in comparison to chemotherapy (Brahmer et al., 2015; Herbst et al., 2016). Currently, pembrolizumab is recommended as first-line treatment for NSCLC patients with high PDL1 expressing tumors (Grant et al., 2021). It is also worth noting that antibodies directed against PDL1, such as atezolizumab (Tecentriq, Genetech), durvalumab (Imfinzi, Astrazeneca) and avelumab (Bavencio, Merck/Pfizer), have also been used to treat NSCLCs (Wu et al., 2019). However, although these ICBs are usually well tolerated, some severe immune-related adverse effects can occur in the course of treatment (Martins et al., 2019).

3.3.5. Targeted therapy

In the last twenty years, the identification of key genetic events driving tumor growth and survival has dramatically redefined the treatment of NSCLC based on their molecular characteristics (Wang et al., 2021). These genetic aberrations occur in certain oncogenes and can serve as drug targets. This is mainly due to the dependency of certain tumors on a single dominant oncogenic protein or pathway to sustain their proliferation (Weinstein, 2002). Inhibition of this specific oncogene can be sufficient to induce substantial growth arrest, resulting in tumor shrinkage. Many driver mutations have been identified in NSCLC (Fig. 4). The most common are represented by activating mutations of the Kirsten rat sarcoma viral oncogene homolog (KRAS) and the epidermal growth factor receptor (EGFR), observed in 30% and 15% of the patients, respectively. Several other genetic aberrations potential have been identified at lower frequencies in NSCLCs, including translocations involving the anaplastic lymphoma kinase (ALK) or the ROS1 proto-oncogene receptor tyrosine kinase (ROS1), amplification or mutations of the MET tyrosine kinase receptor and point mutations of the BRAF serine/threonine kinase (Wang et al., 2021). In recent years, drugs targeting these pathways have

been developed and some have been clinically approved for NSCLC patients. Because of their high response rates and increased specificity compared to standard chemotherapy, targeted agents are used in the first line for the treatment of certain types of patients.

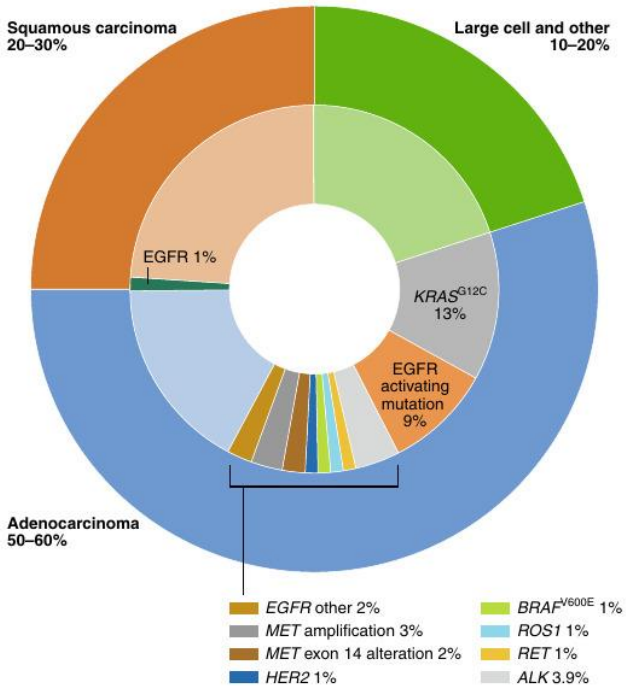


Figure 4. Oncogenic mutations in NSCLC. From Wang et al., Nat. Med. 2021.

3.4. Oncogenic driver mutations in NSCLC

3.4.1. KRAS mutations

KRAS is the most commonly mutated oncogene in NSCLC, representing approximately 20-30% of the cases (Skoulidis and Heymach, 2019). This intracellular guanine nucleotide-binding protein (G protein) belongs to a small GTPases family that also includes NRAS and HRAS. KRAS acts as a molecular switch by cycling between a GDP-bound off state and a GTP-bound on state in response to extracellular stimuli, such as those mediated by receptor tyrosine kinases (RTKs; Fig. 5A). This transitions between both states is tightly regulated by guanine nucleotide-exchange factors, such as son of sevenless (SOS), which loads GTP and activates KRAS, and by GTP hydrolysis, which is catalyzed by GTPase-activating proteins (GAP), such as neurofibromin, to inactivate KRAS (Simanshu et al., 2017). The KRAS-GTP complex triggers

the activation of several downstream signaling effectors, including the canonical mitogen-activated protein kinases (MAPK) and the PI3K/AKT/mTOR pathways. In NSCLC, KRAS activating mutations, which primarily (>95%) occur on exon 2 at codons 12 and 13, alter the GTPase activity of this protein and confer resistance to GTPase activators, leading to constitutively active GTP-bound state (Fig. 5B) (Karachaliou et al., 2013). The most common alterations in KRAS-mutant NSCLC are G12C (46%), G12V (21%), G12D (17%) and G12A (11%) (Moore et al., 2020). KRAS-G12C/V/A have been associated with tobacco, whereas KRAS-G12D is the most frequent mutation among patients without any smoking history (Dogan et al., 2012). Until recently, KRAS has been described as an undruggable target. On May 2021, the FDA granted accelerated approval to the first KRAS inhibitor, sotorasib (Lumakras, Amgen), for the treatment of locally advanced or metastatic NSCLC with KRAS-G12C mutations (Hong et al., 2020; Reck et al., 2021). Sotorasib (also known as AMG510) is a covalent inhibitor that irreversibly binds to cysteine 12 of mutant KRAS, locking the protein in the inactive GDP-bound state (Canon et al., 2019).

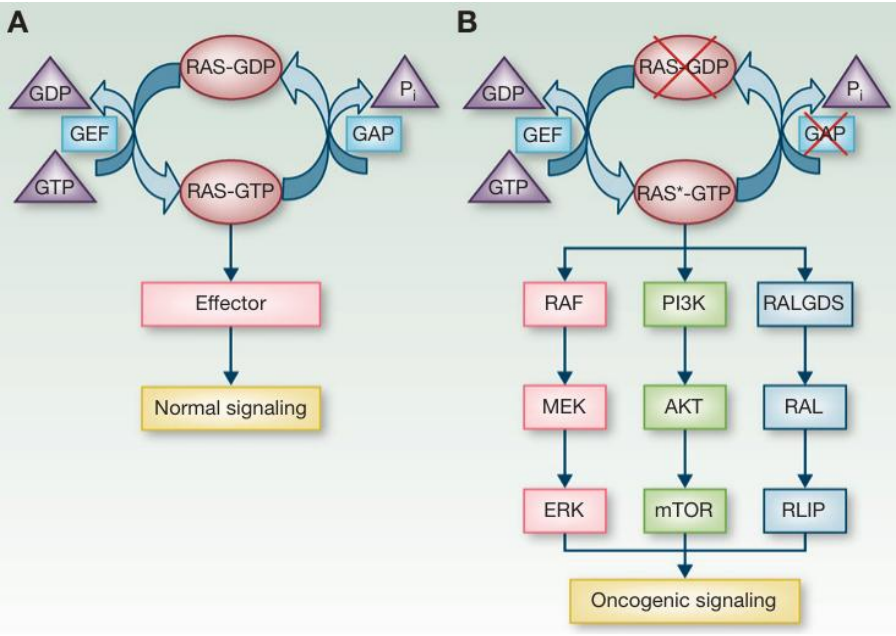


Figure 5. RAS mediated intracellular signal transduction pathways. (A) In normal conditions RAS GTP/GDP bound states is regulated by GEF and GAP proteins. (B) Mutation of RAS impairs the GTPase activity and prevents GAP from promoting the hydrolysis of GTP to GDP, resulting in sustained activation of PI3K, MAPK and RAL pathways, thus promoting cell survival and proliferation**.** From Vasan et al., Clin. Cancer Res. 2014.

3.4.2. ALK rearrangement

ALK is a RTK belonging to the insulin receptor superfamily. It is comprised of an extracellular domain, a transmembrane segment, and a cytoplasmic kinase domain (Shaw and Solomon, 2011). Like other RTKs, ALK is activated by ligand binding, leading to receptor dimerization and autophosphorylation. Heparin (Murray et al., 2015) and augmentor α and β (FAM150) (Guan et al., 2015; Reshetnyak et al., 2015) are known ligands for this receptor. ALK activation has been shown to activate a number of pathways including the MAPK, PI3K/AKT/mTOR, JAK/STAT and phospholipase Cγ (PLCγ) (Lin et al., 2017). In NSCLC, ALK rearrangements are seen in approximately 1-7% of the cases, and they appear to be more frequent in never smokers, younger patients with adenocarcinoma histology (Devarakonda et al., 2015). The echinoderm microtubule-associated protein-like 4 (EML4) is the most common ALK fusion partner in ALK-rearranged NSCLC. EML4-ALK rearrangement occurs through an inversion in chromosome 2 (p21p23), leading to the expression of a chimeric tyrosine kinase, in which the N-terminal of EML4 is fused to the C-terminal intracellular region of ALK containing the kinase domain (Soda et al., 2007). This fusion provokes aberrant expression and ligand independent activation of ALK (Fig. 6). Currently, five ALK inhibitors are FDA approved for the treatment of patients with ALK-rearranged NSCLC: crizotinib (XaLKori, Pfizer), ceritinib (Zykadia, Novartis), alectinib (Alecensa, Genentech), brigatinib (Alunbrig, Takeda) and lorlatinib (Lorviqua, Pfizer).

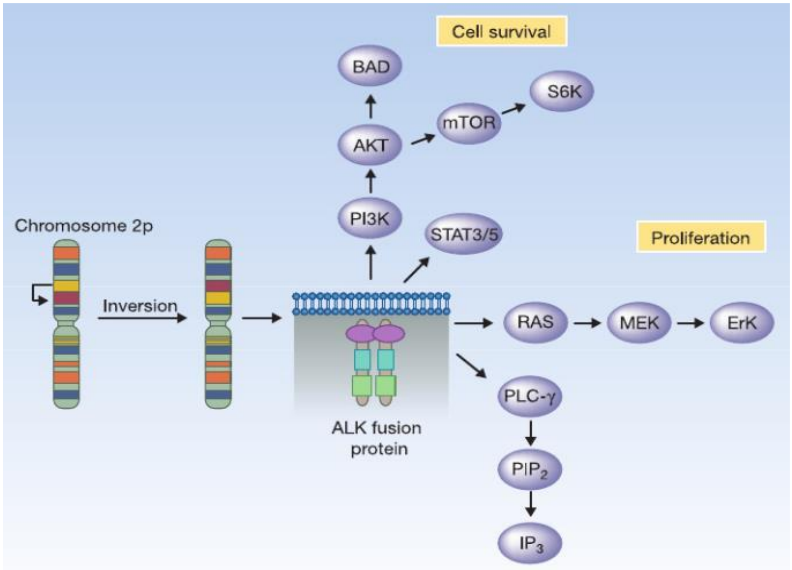


Figure 6. EML4-ALK rearrangements. Inversion of chromosome 2p results in the constitutively active EML4-ALK fusion protein, leading to activation of downstream signaling pathways including, MAPK, PI3K/AKT, and PLCγ, which can promote cell survival and proliferation. From Shaw et al., Clin. Cancer Res. 2011.

3.4.3. ROS1 rearrangements

ROS1 rearrangements are present in approximately 1-2% of NSCLCs and are associated with younger age and the absence of smoking history (Bergethon et al., 2012). The *ROS1* gene is located on chromosome 6q22 and encodes a RTK of the insulin receptor family closely related to ALK (Lin and Shaw, 2017). Chromosomal rearrangements involving the ROS1 gene were initially identified in glioblastoma (Birchmeier et al., 1987) and later reported by Rikova and colleagues in NSCLC (Rikova et al., 2007). In ROS1-rearranged NSCLC, the most frequent fusion partner is CD74. Other partners have also been identified, including EZR, FIG1, TPM3, CCD6, KDELR2, SDC4, SLC34A2 and LRIG3 (Sehgal et al., 2018). This fusion involves the C-terminal sequence of ROS1, including the catalytic domain and the N-terminal region of the partner gene, and it results in constitutive activation of the ROS1 kinase, which subsequently stimulates downstream signaling, such as JAK/STAT, PI3K/AKT/mTOR and MAPK, thus promoting cell survival and proliferation (Gainor and Shaw, 2013). Because of the significant homology between ROS1 and ALK tyrosine kinase domains, some ALK inhibitors were tested in NSCLC patients harboring ROS1 rearrangements (Drilon et al., 2017; Shaw et al., 2014, 2017), leading to the clinical approval of crizotinib and entrectinib (Rozlytrek, Genentech) by the FDA in 2016 and 2019, respectively (Drilon et al., 2020; Shaw et al., 2019).

3.4.4. MET aberrations

MET is a transmembrane RTK that is activated by the hepatocyte growth factor (HGF) and is involved in multiple processes, such as cell proliferation, survival, invasion, and angiogenesis (Comoglio et al., 2018). Splicing mutations of the *MET* gene resulting in exon 14 skipping are also found in 2% of patients with NSCLC. The prevalence of these mutations are higher in elderly patients (older than 70 years) and in adenosquamous (8.2%) or sarcomatoid carcinoma (7.7%) than in adenocarcinoma (2.9%) (Schrock et al., 2016). The exon 14 of MET encodes the intracellular juxtamembrane domain, containing the Y1003 residue, which serves as binding site for the casitas B-lineage lymphoma (CBL) E3 ubiquitin ligase, required for degradation and internalization of MET (Peschard et al., 2001). Therefore, loss of MET exon 14 impairs CBL-mediated receptor ubiquitination and degradation, leading to increased MET stability and prolonged activation of MET signaling (Fig. 7A) (Awad, 2016). Other residues could also

participate in the oncogenic potential of exon 14 skipping, such as serine 985, which is phosphorylated by protein kinase C (PKC) and participates in the regulation of MET activity, as well as a caspase cleavage site, which can result in the generation of the p40MET fragment involved in apoptosis (Cortot et al., 2017). The MET inhibitors, tepotinib (Tepmetko, Merck) and capmatinib (Tabrecta, Novartis), have been approved by the FDA for the treatment of NSCLC patients harboring MET exon alterations (Mathieu et al., 2022; Paik et al., 2020). Besides exon 14 skipping, MET can also be aberrantly activated as a result of gene amplification, which occurs in 3% of NSCLC patients (Rotow and Bivona, 2017). MET amplification can arise either by polysomy due to an increase in the number of copies of chromosome 7, or by focal amplification (Fig. 7B) (Recondo et al., 2020).

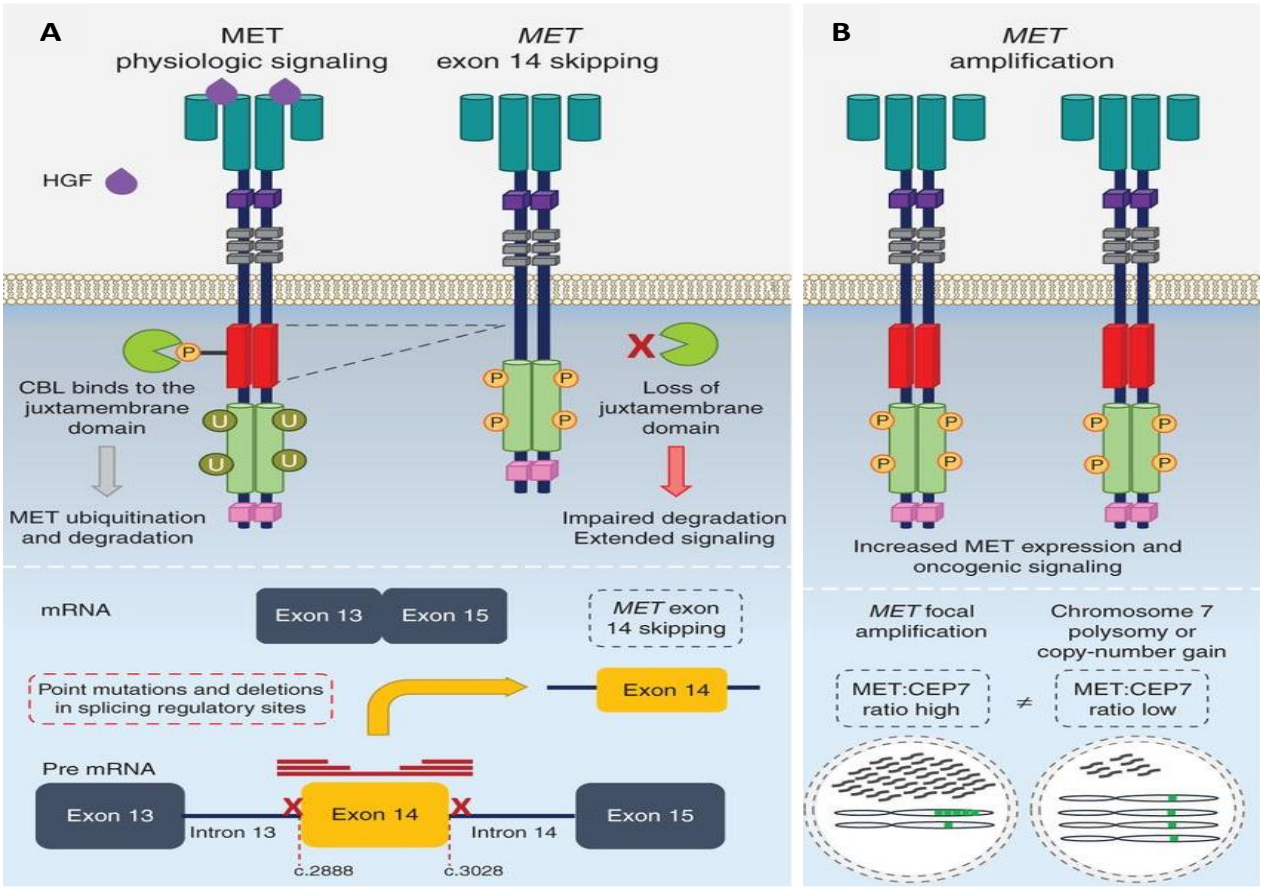


Figure 7. MET alterations in NSCLC. (A) The wild-type MET receptor is activated upon HGF binding. In this physiological condition, MET receptor can be targeted by E3-ubiquitin ligase CBL and directed for lysosomal degradation. In NSCLC, splicing mutations, resulting in exon 14 skipping and loss of the juxtamembrane domain, impair MET ubiquitination and degradation. (B) MET amplification can also lead to aberrant activation of MET signaling. From Recondo et al., Cancer Discov. 2020.

3.4.5. BRAF mutations

BRAF is a serine/threonine kinase that mediates proliferation and cell survival. It is a member of the RAF kinase family, which also includes ARAF and CRAF (also known as RAF1). Similar to its other isoforms, BRAF is activated by RAS proteins. This upstream signal leads to the formation of RAF homodimers (BRAF-BRAF) or heterodimers (BRAF-CRAF or BRAF-ARAF) (Caunt et al., 2015; Yaeger and Corcoran, 2019). Activated RAF can then bind and phosphorylate the mitogen-activated protein kinase kinase 1 and 2 (MEK1 and MEK2), which, in turn, activate the extracellular signal-regulated kinase1 and 2 (ERK1 and ERK2). Mutations in BRAF are generally found in former or current smokers. They are identified in 1-5% of NSCLC patients and they are classified according to their effect on BRAF function. Class I corresponds to the substitution of valine 600 with glutamate, which can be found in about half of BRAF mutant NSCLCs (Rotow and Bivona, 2017). This mutation induces constitutive BRAF activation in its monomeric form in the absence of activated RAS, leading to increased ERK signaling (Yao et al., 2015). Class II mutations include K601E, L597Q and G469A, while the substitutions G466V and D594G/N belong to class III. Class II BRAF mutants can signal as constitutive, RAS independent dimers (Yao et al., 2015), while class III mutants have impaired BRAF kinase activity, but they bind more tightly to RAS and function as heterodimers with wild-type RAF (Yao et al., 2017). Currently available drugs, such as vemurafenib (Zelboraf, Roche) and dabrafenib (Tafinlar, GlaxoSmithKline), are only inhibiting class I BRAF mutants. A combination of dabrafenib and the MEK inhibitor trametinib (Mekinist, Novartis) was FDA approved for the treatment of NSCLC patients harboring the BRAF-V600E mutation.

3.4.6. EGFR mutations

EGFR belongs to the ErbB family of RTKs and binds ligands such as EGF, transforming growth factor alpha (TGF-α) and amphiregulin. The ErbB family consists of four related members, EGFR/ErbB1/HER1, ErbB2/HER2, ErbB3/HER3 and ErbB4/HER4, which are implicated in a wide range of biological processes (Roskoski, 2014). Under physiological conditions, ligand binding to EGFR *via* its extracellular domain triggers receptor homo- and/or heterodimerization with other ErbB members, resulting in auto- and transphosphorylation of the intracellular domain on tyrosine residues (Fig. 8A). The activated receptors signal to downstream

pathways, including MAPK, PI3K/AKT/mTOR and JAK/STAT, to regulate cell proliferation and survival. Somatic activating mutations of EGFR have been reported in 15% of NSCLC patients, and are responsible for constitutive ligand-independent receptor signaling (Rotow and Bivona, 2017). The prevalence is significantly higher in patients of East Asian ethnicity, women and non-smokers (Zhang et al., 2016). These mutations are generally found in adenocarcinomas and they provoke a conformational change that shifts the tyrosine kinase domain towards an active state (Fig. 8B). Exon 19 deletions and L858R substitution represent 90% of EGFR activating mutations in NSCLC (Kobayashi et al., 2013; Shan et al., 2012). Other less common aberrations have been reported, including G719X, L861X, S768I substitutions and exon 20 insertions (Kate et al., 2019; Sharma et al., 2007).

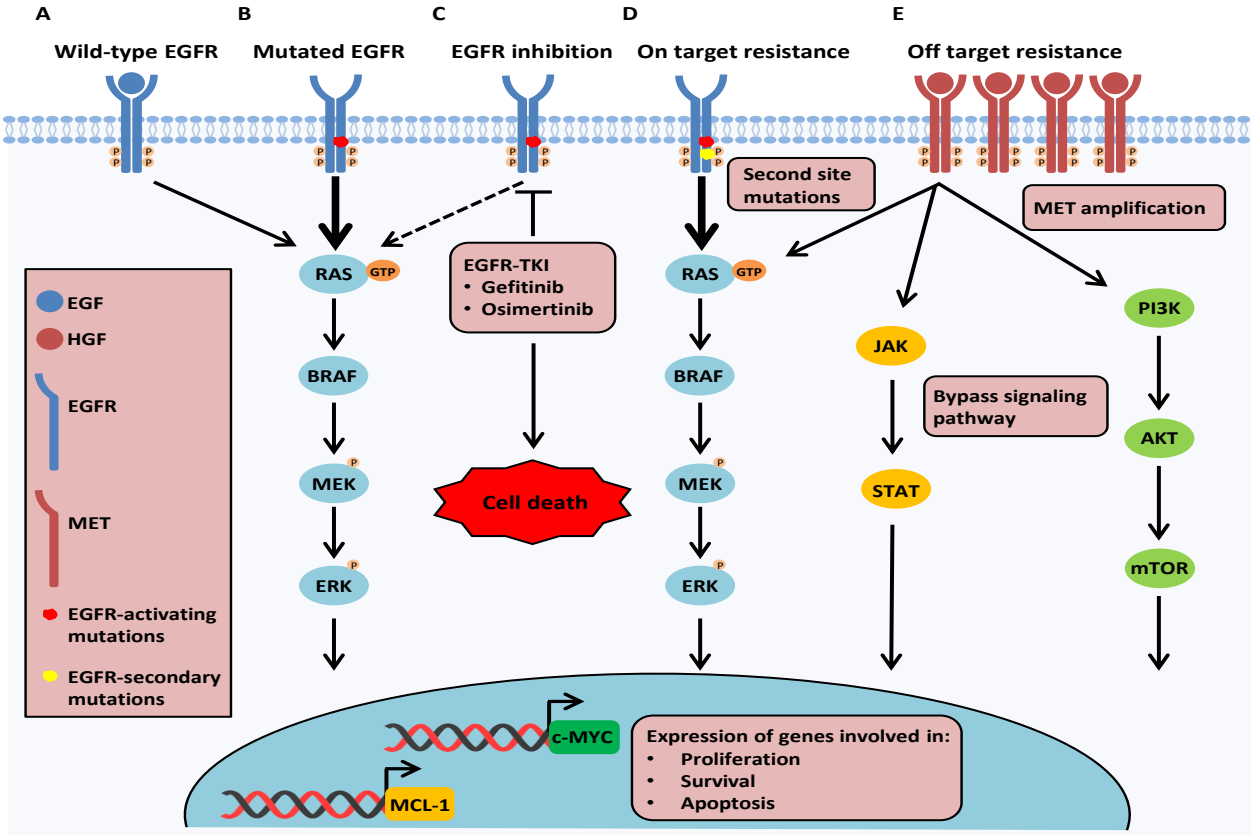


Figure 8. EGFR signaling pathways in NSCLC. (A) Wild-type EGFR is activated upon the binding of EGF, leading to the activation of MAPK signaling pathways that mediate cell survival and proliferation. (B) Oncogenic activating mutations in EGFR, such as exon 19 deletions and L858R point mutation, induce constitutive activation of EGFR that is independent of ligand binding. (C) Pharmacological inhibition of EGFR provokes down-regulation of MAPK pathway, leading to a massive cell death. (D-E) Resistance mechanisms to EGFR-TKIs include second-site EGFR kinase domain mutations or activation of bypass signaling pathway.

3.5. EGFR targeted therapies

The presence of EGFR-activating mutations has been associated with dramatic responses to treatment with EGFR tyrosine kinase inhibitors (EGFR-TKIs) in patients with advanced NSCLC (Fig. 8C) (Lee et al., 2015). Different types of EGFR-TKIs have been developed and some of them have been clinically approved, with response rates ranging from 50 to 80% (Rotow and Bivona, 2017) (Table 3).

3.5.1. First generation EGFR-TKIs

The first generation of EGFR-TKIs is constituted of compounds, such as gefitinib (Iressa, AstraZeneca) and erlotinib (Tarceva, Genentech/OSI Pharmaceuticals), that reversibly compete with ATP for the binding to the tyrosine kinase pocket of the receptor. These drugs demonstrated improved progression-free survival (PFS) and overall response rates (ORR) compared to standard chemotherapy in patients with mutant EGFR NSCLC (Mitsudomi et al., 2010; Zhou et al., 2015). Based on these data, gefitinib and erlotinib were approved as a first-line treatment for NSCLC patients harboring EGFR sensitizing mutations.

Unfortunately, not all the patients respond to these inhibitors, and those that respond to the treatment almost invariably relapse, as the tumors become resistant to EGFR-TKIs. Resistance to targeted therapy falls into two main categories: primary and acquired resistance.

3.5.1.1. Primary or intrinsic resistance to EGFR-TKIs

Primary resistance relates to tumors that fail to respond to the treatment. While exon 19 deletions and the L858R substitution are considered sensitizing mutations, NSCLCs that contain EGFR exon 20 insertions are typically resistant to most EGFR-TKIs, with the uncommon exception of the proximal A763_Y764insFQEA mutation (Vasconcelos et al., 2020). Recently, the novel EGFR-TKIs mobocertinib (Exkivity, Takeda Pharmaceuticals) received accelerated approval from the FDA based on a phase I/II non-randomized, open-label, multicohort clinical trial (NCT02716116) in patients with advanced or metastatic NSCLC carrying EGFR exon 20 insertion mutations (Friedlaender et al., 2022; Riely et al., 2021).

Table 3. Comparison of EGFR-TKIs.

	Drugs	Approval	Clinical trial	CNS penetration	EGFR mutations	EGFR binding
First generation	Gefitinib (ZD1839)	FDA/EMA approved	NCT01203917	Low	Del19/L858R	Reversible Competitive
	Erlotinib (CP358774)	FDA/EMA approved	NCT00446225	Low		
	Icotinib (BPI2009H)	Approved in China	NCT01040780	Low		
Second generation	Afatinib (BIBW2992)	FDA/EMA approved	NCT01466660	Low	Del19/L858R/ T790M	Irreversible Covalent
	Dacomitinib (PF299804)	FDA/EMA approved	NCT01774721	Low		
Third generation	WZ4002	Preclinical	NA	NA	Del19/L858R/ T790M	Irreversible Covalent
	Rociletinib (CO1686)	Rejected	NCT01526928	Low		
	Osimertinib (AZD9291)	FDA/EMA approved	NCT02151981 NCT02296125	High		
	Lazertinib (YH25448)	Approved in South Korea	NCT03046992	High		
	Olmutinib (HM61713)	Approved in South Korea*	NCT01588145	NA		
	Avitinib (AC0010)	Phase I/II (Active)	NCT02330367	Low		
	Nazartinib (EGF816)	Phase I/II (Active)	NCT02108964	NA		
	Mavelertinib (PF06747775)	Phase I/II (Terminated)	NCT02349633	NA		
	Naquotinib (ASP8273)	Phase III (Terminated)	NCT02588261	NA		
	Almonertinib (HS10296)	Approved in China	NCT02981108	Effective in patient with BM	Del19/L858R/ G719X/L861Q/ T790M	
	Alflutinib (AST2818)	Approved in China	NCT03127449	Effective in patient with BM		
Fourth generation	EAI001	Preclinical	NA	NA	L858R/T790M/ C797S	Reversible Allosteric
	EAI045	Preclinical	NA	NA		
	JBJ-09-063	Preclinical	NA	NA		
	BLU945	Phase I/II (Recruiting)	NCT04862780	NA	Del19/L858R/ T790M/C797S	Unknown
	BBT176	Phase I/II (Recruiting)	NCT04820023	NA		
	TQB3804	Phase I (Unknown)	NCT04128085	NA		

*Stopped in 2016 because of two cases of toxic epidermal necrolysis with one of them being fatal. CNS: central nervous system; BM: brain metastases, Del19: exon 19 deletion; NA: not available, EMA: European Medicines Agency, FDA: Food and Drug Administration of the United States.

Primary resistance can also be due to the presence of other concurrent aberrations, such as in the case of a deletion of BIM (also known as BCL2L11), which has been shown to decrease TKIs response in patients harboring EGFR sensitizing mutations (Morgillo et al., 2016; Rotow and Bivona, 2017).

3.5.1.2. Acquired resistance

Acquired resistance arises in patients after an initial period of drug response, and it can be classified as on-target or off-target. On-target resistance is caused by secondary EGFR mutations that interfere with the binding of the inhibitor to the receptor. The most common mechanism of acquired resistance to first generation, reversible EGFR-TKIs is the emergence of the T790M gatekeeper mutation in the ATP binding pocket of EGFR, occurring in 50-60% of the patients that relapse after an initial response to gefitinib or erlotinib (Pao et al., 2005; Yu et al., 2013). It has been shown that this mutation prevents drug binding by increasing the affinity of the mutant receptor for ATP, thus reducing the potency of ATP competitive EGFR-TKIs (Fig. 8D) (Yun et al., 2008). In other patients, off-target resistance results from events that can bypass EGFR signaling, including amplification of other RTKs, downstream activation of certain components of the pathway or histological/phenotypic transformation (Rotow and Bivona, 2017).

3.5.1.2.1. Activation of other RTKs

MET amplification is the most common bypass signaling pathways activated in the setting of acquired resistance to first generation EGFR-TKIs (Wang et al., 2019). Mechanistically, heterodimerization of MET and ErbB3 provides an alternative route to maintain the activation of MAPK, PI3K/AKT or STAT3 pathways, bypassing the EGFR inhibition conferred by TKIs (Engelman et al., 2007). Early phase clinical trials have demonstrated benefit when combining capmatinib to gefitinib for overcoming MET-mediated acquired resistance (Wu et al., 2018).

In addition to MET amplification, up-regulation of other RTKs, such as HER2 (ErbB2) and insulin-like growth factor 1 receptor (IGF-1R), has been shown to mediate resistance to targeted therapies in NSCLC. Amplification of HER2 has been reported in 12% of tumor samples obtained from patients who have acquired resistance to first generation EGFR-TKIs, and that lack EGFR-T790M mutations (Takezawa et al., 2012). It is believed that HER2 amplification can mediate EGFR-TKI resistance through alternative activation of the MAPK and PI3K signaling pathways.

3.5.1.2.2. Aberrations in downstream signaling pathways

Alterations involving downstream components of the EGFR pathway can also result in resistance to EGFR-TKIs. These include mutations in KRAS, BRAF or loss of phosphatase and tensin homolog (PTEN), and have been reported at a lower frequency in early studies of acquired resistance to first generation EGFR inhibitors (Westover et al., 2018). Loss of PTEN has been shown to contribute to erlotinib and gefitinib resistance in EGFR-mutant NSCLC. PTEN is a tumor suppressor gene that negatively regulates the PI3K/AKT/mTOR signaling pathway by converting phosphatidylinositol 3,4,5-triphosphate (PIP-3) back to phosphatidylinositol 4,5-bisphosphate (PIP2). The loss of PTEN function increases the level of PIP-3 and leads to AKT hyperactivation (Huang and Fu, 2015).

3.5.1.2.3. Lineage transformation

Histological transformation from NSCLC to SCLC occurs in up to 10% of patients with EGFR-mutant lung cancers as a mechanism of resistance to first generation EGFR-TKIs (Piotrowska et al., 2018; Sequist et al., 2011). The main common molecular features of all SCLC transformed cases are persistence of the initial EGFR mutation in combination with *RB1* and *TP53* loss (Passaro et al., 2021). Tumors that undergo SCLC transformation become less dependent on EGFR activity and lose sensitivity to EGFR-TKIs. As with other SCLCs, standard chemotherapy combining etoposide and platinum is the most common therapeutic option for those patients. This phenotypic transformation is not exclusive of EGFR mutant NSCLC and it has also been reported in a series of cases of ALK-positive NSCLC with acquired resistance to ALK inhibitors (Lin et al., 2017).

Another form of lineage plasticity that has been associated with acquired resistance to first generation EGFR-TKIs is epithelial-to mesenchymal transition (EMT), which occurs in 1-2% of the cases (Byers et al., 2013; Sequist et al., 2011). EMT is a highly conserved process that allows the acquisition of a mesenchymal phenotype by epithelial cells, resulting in a loss of polarity and an increase in motility. It is molecularly characterized by the down-regulation of epithelial markers, such as E-cadherin, and the up-regulation of mesenchymal markers, such as N-cadherin and vimentin (Singh and Settleman, 2010). EMT is orchestrated by multiple transcription factors, including SNAIL, SLUG, TWIST1/2 and zinc-finger E-box-binding (ZEB). For example, TWIST1 overexpression has been reported to promote EMT in erlotinib resistant NSCLC cell lines (Yochum et al., 2019).

3.5.2. Second generation EGFR-TKIs

Second generation inhibitors, including afatinib (Giotrif, Boehringer Ingelheim Pharmaceutical) and dacomitinib (Vizimpro, Pfizer), were originally designed in the hope to overcome the T790M-mediated resistance. Unlike gefitinib and erlotinib, these agents are irreversible inhibitors that also target HER2 and HER4, and have shown promising activity against the EGFR-T790M mutation in preclinical models. However, despite encouraging *in vitro* data, in patients these EGFR-TKIs failed to prevent the emergence of this mutation, due to non-selective inhibition of wild-type EGFR (Katakami et al., 2013; Miller et al., 2012). These drugs were approved for metastatic NSCLCs harboring non-resistant EGFR mutations and they can be used as an alternative to first generation inhibitors.

3.5.3. Third generation EGFR-TKIs

Given the limited efficacy of second generation EGFR-TKIs in overcoming T790M resistance in NSCLC patients, a number of third generation inhibitors were developed. These agents can form an irreversible covalent bond with the Cysteine-797 residue in the ATP binding site of EGFR and showed potent activity against the EGFR-T790M mutation, while poorly inhibiting the wild-type receptor. The first third generation EGFR-TKIs reported was WZ4002, identified by screening of a library of irreversible kinase inhibitor specifically targeting the T790M mutant EGFR. This drug was found 100-fold less potent against wild-type EGFR and 30

to 100 fold more potent against EGFR-T790M (Zhou et al., 2009). While WZ4002 did not progress into clinical trials due to legal issues between Novartis and the Dana-Farber Cancer Institute regarding the rights to develop this drug, other EGFR-TKIs with similar characteristics were developed and taken into early phase clinical trials. Among these, rociletinib (Xegafri, Clovis Oncology) and osimertinib (Tagrisso, AstraZeneca) were the first compounds to show significant clinical activity in EGFR mutated NSCLC patients who had relapsed after erlotinib, gefitinib or afatinib treatment (Jänne et al., 2015a; Sequist et al., 2015). While rociletinib was finally discarded because of side toxicity (it also inhibits the insulin receptor, provoking hyperglycemia) and lower efficacy (Van Der Steen et al., 2016), the FDA approved osimertinib in 2015 for the treatment of patients with metastatic EGFR-T790M mutant NSCLCs. The phase I/II AURA trial evaluated the pharmacokinetic profile, safety and efficacy of osimertinib in EGFR mutant patients progressing to prior EGFR-TKIs therapy, and initially demonstrated promising results with response rates over 70% and median PFS of 9.6 months when only T790M positive patients were considered (Jänne et al., 2015b). The Phase III AURA3 study further examined osimertinib compared to platinum/pemetrexed chemotherapy in 419 T790M-positive advanced NSCLC patients, who had progressed to first generation EGFR-TKIs. In this trial, osimertinib was associated with higher response rates and median PFS (mPFS) compared to chemotherapy (ORR 71% vs. 31%; mPFS 10.1 vs. 4.4 months), thus establishing osimertinib as the standard of care in this setting. Median PFS benefit was also seen in patients with brain metastases (8.5 vs. 4.2 months) (Mok et al., 2017). In addition, the phase III FLAURA study was designed to compare osimertinib to first generation EGFR-TKIs in front-line therapy for EGFR-mutated, treatment-naïve NSCLC patients. In this trial, osimertinib demonstrated prolonged PFS compared to gefitinib or erlotinib (18.9 months vs. 10.2 months) (Ramalingam et al., 2020; Soria et al., 2018). Based on these results, osimertinib was approved as first-line therapy for NSCLC patients harboring EGFR activating mutations. Other third generation EGFR-TKIs, such as almonertinib (Ameile, Hansoh Pharma) and lazertinib (Leclaza, Yuhan and Janssen Biotech) have been approved in China and South Korea, respectively. These inhibitors have demonstrated good efficacy and safety in patients with EGFR-T790M positive NSCLC (Ahn et al., 2019; Yang et al., 2020).

3.5.3.1. Resistance to third generation EGFR inhibitors

Despite its efficacy, resistance to osimertinib inevitably develops after approximately 11 months in the second-line and 19 months in the first-line setting. Similar to earlier generations, resistance to third generation EGFR-TKIs is classified as primary and acquired. EGFR-C797S mutation represent the most commonly occurring on-target resistance mechanism on osimertinib treatment (Oxnard et al., 2018; Thress et al., 2015). This mutation replaces the residue covalently bound by osimertinib, thus dramatically reducing the efficiency of kinase inhibition. The C797S mutation is seen in approximately 10-20% of NSCLC patients at disease progression to second-line osimertinib (Oxnard et al., 2018; Papadimitrakopoulou et al., 2018), and can emerge in cis or in trans with the EGFR-T790M mutation. The coexistence of C797S and T790M on the same allele (in cis) confers resistance to all generations of EGFR-TKIs, while when these mutations are on different alleles (in trans), the tumors retain sensitivity to the combination of first and third generation EGFR inhibitors (Niederst et al., 2015). In the front-line setting, C797S occur in 7% of patients and has been shown to emerge in the absence of T790M, in which case the tumor remain sensitive to first generation EGFR-TKIs (Ramalingam et al., 2018). In addition to C797S, other, less frequent, on-target EGFR mutations that interfere with the drug binding have also been reported as mechanism of resistance to osimertinib, including G796R/S/, L792H, L718Q and G724S substitutions (Yang et al., 2018; Zhang et al., 2018). It is worth noting that loss of T790M has been observed after osimertinib treatment. This suggests that T790M-positive clones could co-exist with other resistance mechanisms, which is associated with poor responses and shorter PFS (Lin et al., 2018; Passaro et al., 2021).

As described above for first-generation EGFR-TKIs, off-target mechanisms of resistance to osimertinib are frequently observed in patients, with MET amplification being the most common (15%) (Passaro et al., 2021). Histological transformation to SCLC has also been reported in some cases (Piotrowska et al., 2018; Weng et al., 2019).

3.5.4. Fourth generation EGFR-TKIs

In order to overcome C797S-mediated acquired resistance, fourth generation EGFR-TKIs were developed. EAI045 is an allosteric, non-ATP competitive inhibitor targeting both the

T790M and C797S mutations. It binds to the allosteric sites on EGFR, created by the displacement of the regulatory C-helix in the inactive conformation of the kinase. Due to EGFR dimerization, EAI045 is not effective as a single agent, and it requires the co-administration of the anti-EGFR antibody, cetuximab (Erbitux, Merck) (Jia et al., 2016). Recently, JBJ-09-063 was reported as a new EGFR allosteric inhibitor that is effective as a single agent in models harboring EGFR L858R/T790M/C797S mutations (To et al., 2022). Beside allosteric inhibitors, several fourth generation EFR-TKIs that can covalently bind to EGFR have been reported, such as UPR1444, which potently and irreversibly inhibits the EGFR-L858R/T790M/C797S through the formation of a sulfonamide bond with the catalytic residue Lys745 (Ferlenghi et al., 2021). It is also worth noting that several clinical trials are currently ongoing to evaluate the clinical efficacy and the safety profiles of fourth generation EGFR inhibitors, including BLU945 (Spigel et al., 2021), BBT176 (Lim et al., 2021) and TQB3804 (Liu et al., 2019).

3.5.5. EGFR degraders

Fourth generation EGFR-TKIs appear to be potent and effective against the EGFR-C797S mutations, and to have strong antitumor activity in preclinical models. Besides TKIs, other types of inhibitors have been designed that could improve the treatment of patients with EGFR-mutant NSCLC. In particular, the discovery of EGFR degraders holds great promise. A key focus of targeted protein degradation is the development of proteolysis targeting chimeras (PROTACs). PROTACs are heterobifunctional small-molecule degraders, typically consisting of two linked moieties with one binding the protein of interest and the other binding an E3 ligase. PROTACs recruit the E3 ligase to targeted protein, leading to its selective ubiquitination and degradation by the proteasome (Burslem and Crews, 2020; Dale et al., 2021). Multiple selective EGFR degraders were developed, which can selectively inhibit the proliferation of EGFR-mutant NSCLC cells (Qu et al., 2021; Zhang et al., 2020; Zhao et al., 2020). Recently, Du et al. described a novel EGFR-based PROTAC, HJM-561, that potently inhibits the proliferation of tumor cells harboring the EGFR-C797S mutation (Du et al., 2022).

4. Drug tolerant persister states

Despite the high response rates of EGFR-TKIs, acquired resistance almost invariably occurs. In some cases, this is due to the selection of pre-existing cells harboring well defined genetic resistance mechanisms, such as the T790M mutation or MET amplification, that enable growth during the treatment (Bhang et al., 2015; Hata et al., 2016; Turke et al., 2010). As an alternative mechanism, it has been shown that resistance can also arise from drug-tolerant persister (DTP) cells, sometimes referred to as minimal residual disease, which can survive during treatment, when the large majority of the cancer cells dies. In the presence of the drug, these subpopulations display a slow proliferative rate with altered cellular metabolism, and can survive prolonged treatment through epigenetic adaptations (Marine et al., 2020; Shen et al., 2020).

4.1. Persisters in NSCLC

The term of DTP comes from the field of microbiology, where it has been shown that a small fraction of dormant bacteria has the capacity to survive in the presence of antibiotics despite the fact that they do not contain a genetic mechanism of resistance. These drug tolerant cells are able to resume growth and re-establish a drug sensitive population upon drug withdrawal, indicative of a transient and not inheritable resistance mechanism (Fisher et al., 2017). Similar to bacterial persisters, cancer cells can enter a reversible drug-tolerant state when exposed to anticancer therapy (Russo et al., 2021). The first report describing the transient acquisition of drug-tolerant state in NSCLC came in 2010 from Jeffrey Settleman's laboratory. In this study, Sharma and colleagues used the well established EGFR-mutant NSCLC cell line PC9 and observed that a small fraction (range 0.3-5%) of quiescent cells remained viable after 9 days of treatment with the first generation EGFR-TKIs erlotinib. These DTPs can resume normal proliferation in the presence of gefitinib or erlotinib, and give rise to a second population of cells termed drug-tolerant expanded persisters (DTEP) (Fig. 9). Of note, the authors showed that after drug withdrawal these populations remained resistant to EGFR-TKIs for up to 90 cycles of cell divisions, before reverting to a drug-sensitive state (Sharma et al., 2010). After this initial description, other studies showed that DTPs can function as a reservoir from which

heterogeneous mechanisms of acquired resistance can arise (Hata et al., 2016; Ramirez et al., 2016). Similar subpopulations of drug-tolerant cells have also been identified in other types of cancer, including colon cancer (Touil et al., 2014), melanoma (Sun et al., 2014a) and glioblastoma (Liau et al., 2017).

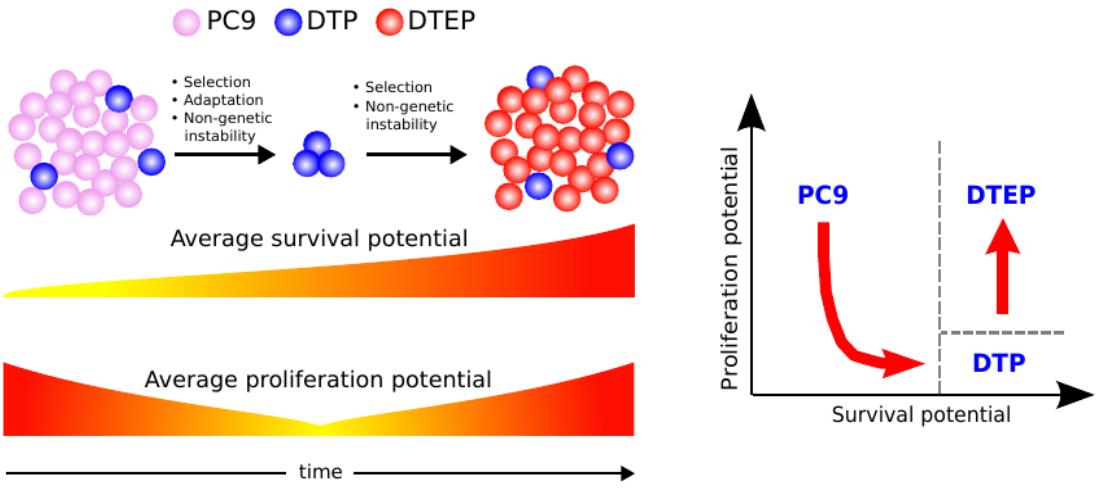


Figure 9. Schematic representation for the acquisition of drug-tolerant phenotype in NSCLC as described by Sharma and colleagues. PC9 cells consist of a heterogeneous mix of subpopulations with different sensitivity to EGFR-TKIs. The majority of the population is eliminated in response to EGFR-TKIs, but a small fraction of cells can survive and constitute a population of drug-tolerant persisters (DTPs). Under continuous treatment, DTPs can resume growth and generate a population of drug-tolerant expanded persisters (DTEP). From Chisholm et al., Cancer Res. 2015.

4.2. Molecular characteristics of drug tolerance in NSCLC

Several mechanisms have been associated with the ability of drug-tolerant cells to withstand EGFR-TKIs treatment, including chromatin remodeling, activation of bypass pathways and altered cellular metabolism. In this section, I will summarize the main characteristics of drug-tolerant state described in EGFR-mutant NSCLC. For a more general discussion of how DTP cells can escape cell death in different types of cancer, please refer to our recent review (Swayden et al., 2020).

4.2.1. Epigenetic modifications

To gain insight into the underlying molecular mechanism of the drug-tolerant phenotype, Sharma and colleagues analyzed the gene expression profiles of DTP and DTEP, and found that

these populations are characterized by altered chromatin states. In particular, the authors showed that these cells were able to survive in the presence of EGFR-TKIs by up-regulating the histone demethylase KDM5A, while they could be selectively ablated by the histone deacetylase (HDAC) inhibitor trichostatin A. The study also suggested a role for the insulin-like growth factor 1 receptor (IGF-1R) in the emergence of DTP populations: activation of this receptor was shown to drive drug tolerance by increasing the expression of KDM5A, leading to overall repressive changes in chromatin structure (Sharma et al., 2010).

In line with these findings, Guler et al. showed that EGFR-TKIs induce increased expression of the long interspersed repeat element 1 (LINE-1), an active retrotransposable element that can propagate and insert randomly throughout the genome, resulting in genome instability. The authors showed that DTPs exhibit a repressive chromatin state by increased methylation of lysines 9 and 27 of histone H3 (H3K9 and H3K27), particularly on the LINE-1 locus, thus decreasing DNA damage in this cell sub-population. Treatment with histone methyltransferase inhibitors, such as tazemetostat (Tazverik, Epizyme), which specifically reduces the global levels of H3K27me3 by inhibiting the activity of enhancer of zeste homolog 2 (EZH2), increase the chromatin accessibility and results in the ablation of DTP cells through derepression of LINE-1 elements (Guler et al., 2017).

4.2.2. Reactivation of EGFR signaling and up-regulation of other pathways

Reactivation of ERK1/2 has been identified as a resistance mechanism to EGFR-TKIs in NSCLC. Ercan and colleagues found that this activation is caused by either an amplification in chromosome 22 harboring the mitogen-activated protein kinase 1 (MAPK1) or by down-regulation of negative regulators of ERK signaling, including the dual specificity phosphatase 6 (DUSP6) (Ercan et al., 2012). A combination of EGFR and MEK inhibitors effectively prevents reactivation of ERK1/2 and delay the emergence of DTP cells (Tricker et al., 2015).

The tyrosine kinase receptor AXL has been reported to mediate NSCLC tolerance in response to osimertinib treatment. Taniguchi et al. showed that AXL can be activated through suppression of a negative feedback loop involving SPRY4, resulting in AXL heterodimerization with EGFR or HER3. Consistent with these findings, inhibition of AXL could restore the sensitivity to osimertinib and prevent the emergence of DTP cells (Taniguchi et al., 2019). Of

note, AXL was also shown to drive EMT transition in EGFR-mutant NSCLC models with acquired resistance to erlotinib (Zhang et al., 2012).

Shah and colleagues found that aurora kinase A (AURKA) plays a major role in the emergence of DTP cells induced by EGFR-TKIs treatment. AURKA activation, which is triggered by its regulator TPX2, prevents osimertinib-induced apoptosis by increasing BIM phosphorylation. The authors showed that targeting AURKA and EGFR was efficient to reduce the proportion of DTP in vitro and in vivo (Shah et al., 2019). A phase I/II clinical trials are currently ongoing to evaluate the efficacy of this combination in patients with advanced EGFR-mutant NSCLC (NCT04085315/NCT05017025). In addition to AURKA, Tanaka et al. reported a role for AURKB in the emergence of DTP cells in NSCLC. Mechanistically, the authors found that AURKB inhibition overcomes resistance to osimertinib by enhancing BIM and PUMA-mediated apoptosis (Tanaka et al., 2021).

In another recent study, it was shown that NSCLC cells survive EGFR and MEK dual inhibition by entering a senescence-like state, which is accompanied by up-regulation of the YAP/TEAD pathway. YAP is a transcriptional coactivator that shuttles between the cytoplasm and the nucleus, where it interacts with the transcriptional factor TEAD and regulates the expression of genes that promote cell growth and survival. The authors found that YAP and TEAD cooperate with the EMT transcription factor SLUG to repress the expression of the pro-apoptotic factor BMF, allowing the cells to escape apoptosis and survive. They also developed a new TEAD inhibitor less toxic compared to previous compounds and they showed that it can enhance apoptosis and prevent the emergence of DTP cells in response to EGFR and MEK dual inhibition (Kurppa et al., 2020). Similarly, in a preprint by Trever Bivona's laboratory, it was reported using a humanized mouse model that YAP activation can decrease sensitivity to osimertinib treatment and induce an immunosuppressive microenvironment to support tumor growth. In addition, the auhors found that YAP activation is regulated by the focal adhesion kinase (FAK). These findings highlight the importance of the FAK-YAP signaling axis in promoting drug tolerance in PC9 cells treated with osimertinib (Haderk et al., 2021).

The Wnt/β-catenin signaling has also been associated with the maintenance of DTP cells in response to EGFR-TKIs treatment. One study showed that EGFR inhibition results in the

activation of β-catenin signaling in a Notch3-dependent manner, leading to the survival of a subpopulation of DTP cells with stem cell-like properties (Arasada et al., 2018). Consistent with these data, Maynard and colleagues used single-cell RNA sequencing to analyze tumor biopsies of EGFR-mutant NSCLC patients treated with osimertinib, and showed that DTP cells are characterized by an alveolar-regenerative signature, which was related to activation of Wnt/β-catenin signaling (Maynard et al., 2020).

4.2.3. Metabolic reprogramming

Remodeling cellular metabolism, including the ability to maintain the redox balance under nutrient-deprived conditions and other stresses, is one of the hallmarks of cancer (Hanahan, 2022). Compared to normal cells, which rely primarily on mitochondrial oxidative phosphorylation (OXPHOS) to generate ATP for energy, cancer cells generally depend on aerobic glycolysis. This phenomenon, also known as the Warburg effect, represents the most common feature of metabolic reprogramming observed in cancer cells, and it is characterized by increased glucose uptake via glycolysis, rather than mitochondrial oxidative phosphorylation, regardless of oxygen availability and mitochondrial activity (Vander Heiden et al., 2009). While aerobic glycolysis appears to occur in rapidly proliferating cancer cells, it has been shown that slowly-cycling DTP cells depend more on mitochondrial respiration for their energy production (Shen et al., 2020). This metabolic shift to OXPHOS in DTP cells results in increased levels of reactive oxygen species (ROS). Thus, DTP cells require a robust antioxidant process to protect themselves from oxidative stress (Mikubo et al., 2021). Raha et al. demonstrated that aldehyde dehydrogenase (ALDH) is required for the maintenance of a persistent cell population in NSCLC. They found that ALDH protects DTP cells from ROS-mediated toxicity and that pharmacologic inhibition of ALDH activity leads to accumulation of ROS to toxic levels, causing DNA damage and cell death within the drug-tolerant subpopulation (Raha et al., 2014). Consistent with an increased susceptibility to oxidative stress in DTP cells, another study reported that these cells rely on the expression of the glutathione peroxidase 4 (GPX4) in EGFR-mutant NSCLC, as well as in other different types of cancer. The authors found that two GPX4 inhibitors (RSL3 and ML210) were selectively lethal to DTP cells and that GPX4 inhibitor-mediated cell death was accompanied by accumulation of lipid hydroperoxides and could be

rescued by the lipophilic antioxidants ferrostatin-1 and liproxstatin-1, suggesting a ferroptotic mechanism (Hangauer et al., 2017). Using a DNA barcoding strategy, Oren et al. recently showed that in NSCLC PC9 cells treated with osimertinib two populations of cycling and non-cycling DTP cells can be identified, which display different transcriptional and metabolic profiles. Of note, cycling DTP cells were associated with increased expression of antioxidant genes and a metabolic shift to fatty acid oxidation (Oren et al., 2021).

4.3. Origin of DTP cells: Darwinian selection or Lamarckian induction?

While it's becoming more and more evident that acquisition of drug resistance doesn't rely only on genetic mechanisms, a major question that is still unanswered concerns how tolerant/persister cells originate. Certain cell populations could be enriched because of intrinsic epigenetic properties that favor growth in the presence of the drug, through a non-genetic Darwinian selection (Fig. 10). For example, Shaffer et al. found that small populations of BRAF-mutated melanoma are primed to become tolerant by transiently expressing high levels of EGFR, AXL or the nerve growth factor receptor. They showed that these cells exhibit profound transcriptional heterogeneity at the single cell level that predicts which cells will eventually resist drug therapy (Shaffer et al., 2017). In another study from the same group, the authors performed a high-throughput CRISPR/Cas9 genetic screen to identify modulators of cell fate in the context of resistance to BRAF inhibition in melanoma cells and they found that inactivation of different factors, including DOT1L, LATS2 and BRD2, can modify the proportion of cells primed to become DTP (Torre et al., 2021). In a recent preprint, the Raj's laboratory describes a strategy combining DNA barcoding and single-cell RNA-Sequencing (scRNA-Seq) that was used to show that melanoma DTP cells can adopt different transcriptional and functional profiles in response to targeted therapy. By comparing the transcriptional profile of individual drug selected cells with their barcodes across twin replicates derived from the same population, the authors concluded that for some subpopulations the DTP phenotype is intrinsically predetermined before the onset of the treatment, suggesting that DTP cells are pre-existing and selected upon drug exposure (Goyal et al., 2021).

According to a different model, the tolerant/persister state could be directly induced by the treatment (Fig. 10). This process, defined as Lamarckian induction (Pisco et al., 2013; Su et al.,

2017), implies that these cells arise more randomly, through a cell fate decision that can be potentially influenced by stochastic fluctuations of gene expression (Moris et al., 2016; Raj and van Oudenaarden, 2008) or a particular phase of the cell cycle, as it has been shown for embryonic stem cells (Pauklin and Vallier, 2013). Consistent with this type of scenario, Kurppa et al. reported that NSCLC DTP cells surviving EGFR-MEK dual inhibition don't derive from pre-existing primed clones, but they arise randomly from the mass population of untreated cells (Kurppa et al., 2020). Moreover, gene signatures found in NSCLC DTP cells treated with osimertinib were not present in untreated cells, implying that the induction of the drug-tolerant phenotype is, at least in part, an adaptive process (Oren et al., 2021). Rambow et al. identified distinct drug tolerant transcriptional states that emerge upon BRAF/MEK inhibition in melanoma patient derived xenografts (PDXs). The authors showed that DTP cancer cells displaying a neural crest stem cell-like profile are more likely to give rise to fully resistant clones, indicating that the type of transcriptional changes induced by the treatment can affect the long term fate of DTP cells (Rambow et al., 2018). Recent studies on colorectal and breast cancer cells suggest that any cancer cells has the ability to enter a drug tolerant persister state in response to chemotherapy by adopting a state that resembles diapause, a highly conserved developmental mechanism used by embryos to survive unfavorable environmental conditions (Dhimolea et al., 2021; Rehman et al., 2021).

It is important to note that Darwinian selection and Lamarckian induction are not mutually exclusive, since certain subpopulations can be selected based on a particular pattern of gene expression, followed by additional changes induced by the drug. For example, the study by Schaffer et al. discussed above identified rare cell subpopulations displaying particular transcriptional profiles that are primed to survive in the presence of the drug. The treatment induces the acquisition of a more stable resistant phenotype through an epigenetic reprogramming, possibly engendered by inhibition of SOX10-mediated differentiation and induction of AP1 and TEAD transcription factors (Shaffer et al., 2017). A better understanding of the mechanisms underlying the survival of certain cell populations during treatment could provide new strategies to target residual cancer cells, a necessary step to improve clinical efficacy and prevent tumor relapse.

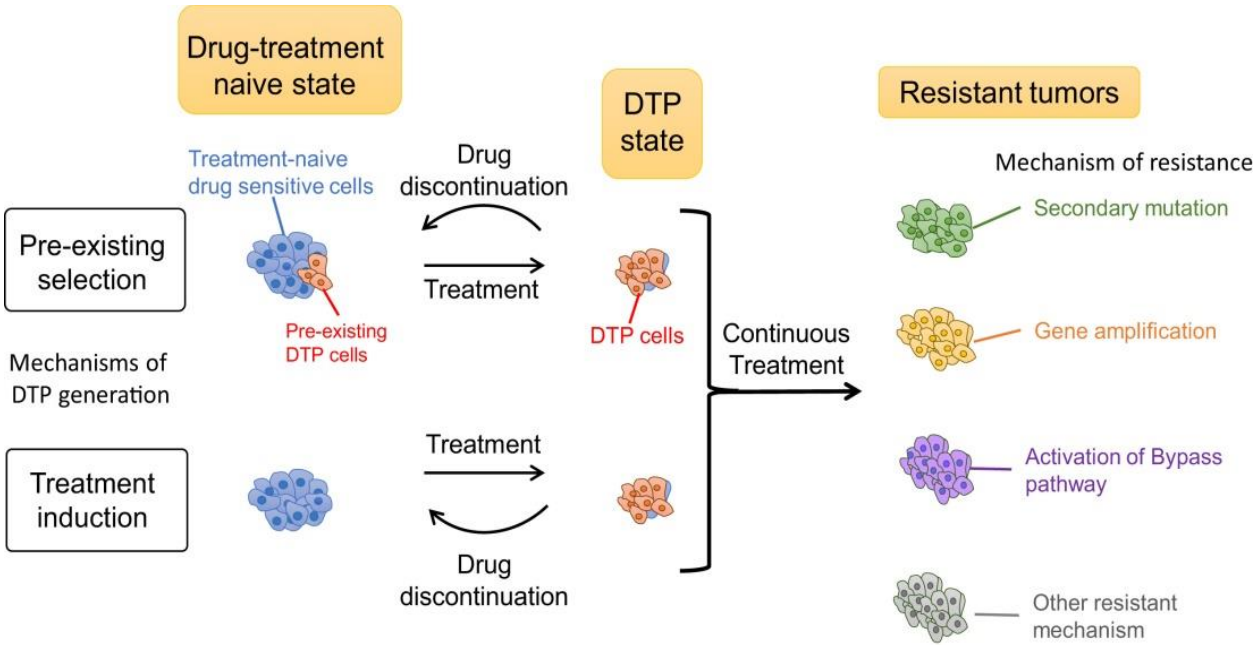


Figure 10. Models for the emergence of drug-tolerant cells. In the pre-existing selection model (top), DTP cells are selected in response to treatment because of some intrinsic properties through non-genetic Darwinian selection. Alternatively, in the drug induced model, also known as Lamarckian induction (bottom), DTP cells can originate more randomly as a direct effect of the treatment. These cells can evolve over time to acquire various genetic or non-genetic mechanisms of resistance. From Mikubo et al., J. Thorac. Oncol. 2021.

5. Cellular barcoding

An elegant strategy to investigate how intratumor heterogeneity can drive drug resistance relies on genetic labeling of individual clones within a mass population of cancer cells. Initially used for lineage tracing and fate mapping in developmental biology, cellular barcoding has been subsequently applied to stem cell research and in modeling tumor heterogeneity in cancer.

5.1. Brief history

Cell labeling was first developed in 1929 by Walter Vogt, who used non toxic vital dyes implanted into a small piece of agar placed on the cells of interests. Through this approach, Vogt was able to label groups of cells in Xenopus early embryos and follow the fate of their progeny during gastrulation. To prevent the diffusion of vital staining to neighboring cells, lipophilic dyes, such as DiI or DiO (water-insoluble) and horseradish peroxidase (too large to pass through gap junctions), were later used to directly label and track the cells. However, this method is

limited by the dilution and the loss of the marker after multiple rounds of cell division (Buckingham and Meilhac, 2011).

The discovery and the isolation of green fluorescent protein (GFP) from *Aequorea victoria* in 1962 has completely transformed cell tracking. Reporter genes encoding fluorescent proteins can be used as inherited genetic markers that enable permanent tracking of cells over time and space. Similar strategies are based on the LacZ gene encoding β-galactosidase, which can be visualized by adding the X-galactose substrate that turns blue after cleavage (Kretzschmar and Watt, 2012). A major drawback of these approaches is the limited number of available proteins, which makes it impossible to discriminate the different cells forming complex populations. To circumvent these limitations, the insertion of new DNA sequences in the genome can be used as genetic barcodes to investigate higher levels of heterogeneity. This type of strategy was first developed by Walsh and Cepko in 1992 to label progenitor cells of the cerebral cortex through retroviral integration of a transgene in their genome (Walsh and Cepko, 1992). To further increase barcode complexity, other approaches were devised based on the insertion of short highly variable sequences of DNA in the genome of the cells. For example, in 2008 Ton Schumacher's laboratory generated a retroviral library containing around 5000 distinct barcodes formed by a semi-random stretch of 98 bp of noncoding DNA. The authors used this library to label and trace different subpopulations of mouse T-cells. To "read" the barcodes, they amplified by PCR from genomic DNA (gDNA) the region of the retrovirus containing the variable sequences, and the amplicons were labeled with fluorescent dyes for hybridization on a barcode microarray. For the first time, the term cellular barcoding was used to describe a high-throughput fate-mapping screen of individually labeled cells (Schepers et al., 2008).

5.2. Strategies for cellular barcoding

The basic principle underlying cellular barcoding involves the tagging of individual cells of interest with unique and heritable labels. The ideal way to label a sample would be to manually assign individual barcodes to each cell. This method guarantees that the cells are uniquely labeled. However, one-by one labeling is labor intensive and can only be used in certain conditions, when the number of cells is limited (Kebschull and Zador, 2018). The next

paragraphs provide an overview (summarized in Table 4) of the different barcoding strategies that have been used in various biological models, both *in vivo* and *in vitro*.

5.2.1. Cre-mediated recombination

The Cre-loxP system is one of the most widely used approach for *in vivo* barcoding. This strategy is based on an inducible and tissue-specific Cre recombinase that acts on a locus in which genes encoding different fluorescent proteins are flanked by loxP sequences (Fig. 11). One of the first systems was Brainbow, in which Cre-recombinase can excise or flip the genes of four fluorescent proteins in a stochastic manner, resulting in a mosaic of differently colored cells that can be analyzed by imaging (Livet et al., 2007). While in this study the expression of the reporter was restricted to neuronal cell types (the cassette was placed under the control of regulatory elements from the Thy1 gene that are highly expressed in a variety of nerve cells), Hans Clevers' laboratory developed a universal rainbow strain, named R26R-confetti, in which the Brainbow cassette was cloned in the Rosa26 locus under the expression of the ubiquitous CAGG promoter, which lies immediately upstream of a floxed NeoR-cassette that prevents the expression of the reporter. The authors crossed these mice with a strain in which an inducible Cre was expressed under the control of the Lgr5 promoter, specific of intestinal stem cells. Upon Cre activation, the NeoR cassette is excised and the array of fluorescent genes is recombined, allowing the expression of reporters of different colors in the stem cells and their progeny. Using this system, the authors showed that homeostasis of Lgr5-positive stem cells within the intestinal crypt is regulated by neutral competition between cells (Snippert et al., 2010). To overcome the limitation due to the small number of fluorescent proteins, the Polylox system used as genetic labels short DNA sequences instead of entire genes. More specifically, the Polylox cassette is composed of 10 loxP sites interspaced with unique sequences that can be randomly excised or flipped upon Cre-induced recombination. This system allows the generation of approximately one million distinct barcodes, which can be analyzed by sequencing. The Polylox approach has been applied to investigate the fate of hematopoietic stem cells (HSCs). The authors found that most HSC clones can originate multiple hematopoietic lineages, suggesting an alternative route to the classical hierarchical model of hematopoiesis, where HSCs produce different types of progenitor cells that differentiate into all major blood cell types (Pei et al., 2017).

A major drawback of the Cre-loxP based approach is that the size of the targeted array generally decreases over time, resulting in a reduction of barcoding diversity. This is mainly due to the fact that Cre recombinase favors excision over flipping (Fig. 11). An additional limitation is that barcode target arrays are long and repetitive because of the low diversity of loxP sites and the nature of their minimum spacing requirements (Kebschull and Zador, 2018).

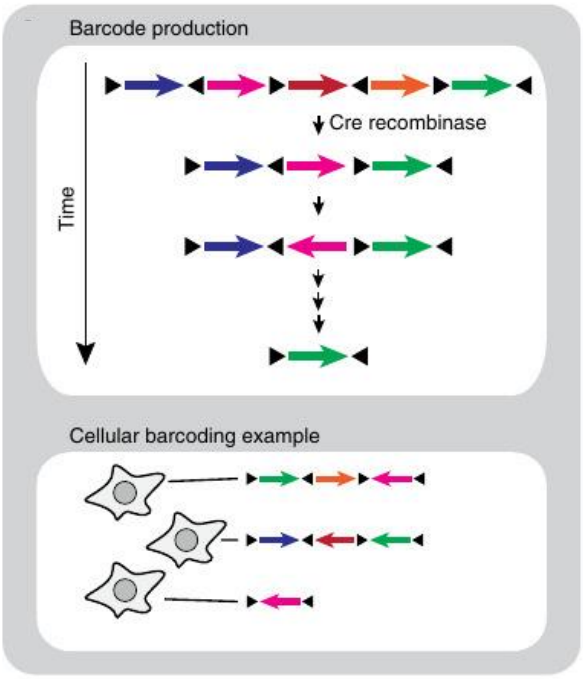
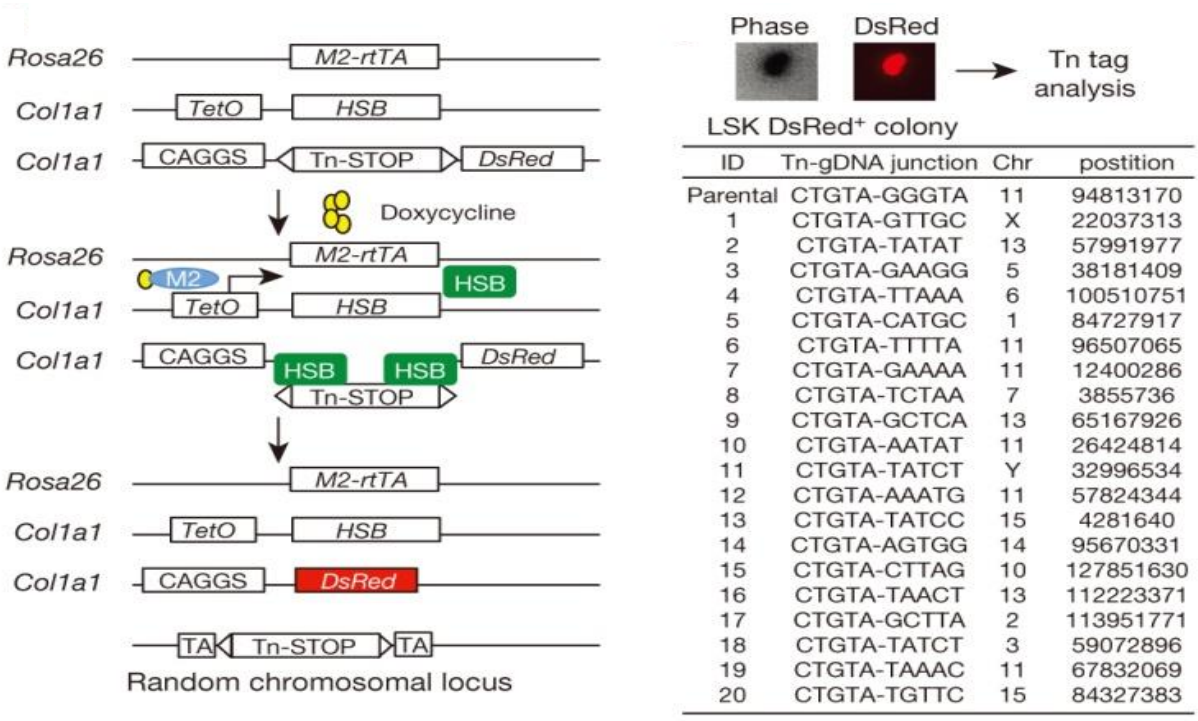


Figure 11. Cre-loxP recombination for in vivo barcoding. This approach relies on the action of Cre recombinase to rearrange an array of target sites (colored arrows) flanked by loxP sites (black triangles) to generate barcode diversity. Because Cre intrinsically favors excision over flipping, the target array shrinks in size over time, thus resulting in reduced diversity. From Kebschull and Zador, Nat. Med. 2018.

5.2.2. Transposon-based approaches

In 2014, Sun et al. developed a strategy for *in vivo* genetic labeling based on inducible expression of a transposase that triggers the mobilization of a DNA transposon. This system involves a doxycycline-dependent transactivator (rtTA-M2) located in the Rosa26 locus, a hyperactive Sleeping Beauty (HSB) transposase driven by a doxycycline inducible promoter and one non-mutagenic, HSB-responsive transposon element (Tn). In mice carrying these three cassettes (named M2/HSB/Tn), doxycycline administration results in HSB expression and subsequent Tn mobilization to a random genomic location. To monitor DNA transposition, a red

fluorescent protein is expressed upon transposon mobilization by the concurrent removal of a stop codon (Fig. 12). As a result, each cell undergoing transposition will carry a single and distinct insertion site that can serve as a stable genetic tag. The authors used this strategy to study the clonal dynamics of native haematopoiesis and they showed that, while HSCs are responsible for the reconstitution of the different blood lineages after transplantation, under steady state conditions the renewal of the hematopoietic system depends mostly on long-lived progenitors (Sun et al., 2014b). While powerful, this approach suffers from the difficulty to recover the transposon insertion sites.



ID	Tn-gDNA junction	Chr	postition
Parental	CTGTA-GGGTA	11	94813170
1	CTGTA-GTTGC	X	22037313
2	CTGTA-TATAT	13	57991977
3	CTGTA-GAAGG	5	38181409
4	CTGTA-TTAAA	6	100510751
5	CTGTA-CATGC	1	84727917
6	CTGTA-TTTTA	11	96507065
7	CTGTA-GAAAA	11	12400286
8	CTGTA-TCTAA	7	3855736
9	CTGTA-GCTCA	13	65167926
10	CTGTA-AATAT	11	26424814
11	CTGTA-TATCT	Y	32996534
12	CTGTA-AAATG	11	57824344
13	CTGTA-TATCC	15	4281640
14	CTGTA-AGTGG	14	95670331
15	CTGTA-CTTAG	10	127851630
16	CTGTA-TAACT	13	112223371
17	CTGTA-GCTTA	2	113951771
18	CTGTA-TATCT	3	59072896
19	CTGTA-TAAAC	11	67832069
20	CTGTA-TGTTC	15	84327383

Figure 12. Cellular barcoding through a Sleeping Beauty transposon system. Upon doxycycline treatment, the HSB transposase is expressed and induces the mobilization of a transposon element (Tn) to random chromosomal locations, which serve as stable genetic barcodes. From Sun et al., Nature. 2014.

5.2.3. CRISPR-mediated approaches

Originally an adaptive immune system in prokaryotes, CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) has been engineered as a new powerful tool for genome editing. This system is composed of the Cas9 nuclease from S. Pyogenes and a short RNA sequence, the single-guide RNA (sgRNA). When co-expressed in the cells, Cas9 and the sgRNA form a complex that specifically recognizes the targeted DNA sequence through Watson-Crick

pairing and promotes its cleavage. CRISPR/Cas9-induced double-stranded breaks in the genome can be repaired by two endogenous mechanisms: non-homologous end joining (NHEJ) or homology directed repair (HDR) (Hsu et al., 2014). NHEJ can introduce unique small insertions or deletions (indels) at the cleavage site, which can be used as permanent genomic barcodes to label cells. GESTALT (genome editing of synthetic target arrays for lineage tracing) was one of the first approaches to perform *in vivo* barcoding using CRISPR technology. The authors generated an array of 10 different target sequences that was integrated into the genome of zebrafish embryos. By injecting the fertilized eggs with Cas9 and 10 different sgRNAs with perfect match to the target sequences, random indels were generated that could serve as unique barcodes to label and trace cell lineages during organogenesis in zebrafish embryos (McKenna et al., 2016). Another similar method, named ScarTrace, was developed by the van Oudenaarden laboratory and it is based on a sgRNA directed against eight tandem repeated copies of GFP fused to the ubiquitously expressed gene encoding the histone variant H2A.F/Z (Junker et al., 2016). In this study, the authors take advantage of an available transgenic zebrafish line (H2A.F/Z:GFP), in which a fusion between H2A.F/Z and GFP was shown to efficiently drive expression of the fluorescent protein in the developing embryo and into adulthood (Pauls et al., 2001). The injection of sgRNA and Cas9 protein into the embryos led to the generation of scars or mutations that can be used as barcodes. In ScarTrace, loss of GFP serves as a direct visual confirmation of efficient scar formation. These scars can be then detected by targeted sequencing of GFP. Variations of these methods were designed to be used in combination with scRNA-Seq analysis (Alemany et al., 2018; Raj et al., 2018). However, large deletions and loss of barcode information from one or more target sites have been observed with these methods (Baron and van Oudenaarden, 2019). LINNAEUS (lineage tracing by nuclease-activated editing of ubiquitous sequences) has been developed to mitigate these effects by using a Cas9 target array composed of 16 copies of red fluorescent protein (RFP) across the genome of zebrafish (Fig. 13A) (Spanjaard et al., 2018).

While powerful, these methods are only suitable for fast-developing organisms, because Ca9 editing occurs during a short period after injection and it is restricted to the first cell divisions in the embryo. This limitation represented a particular challenge to design genetic barcoding in mammals, where high diversity is needed to label large cell populations over long

periods (Baron and van Oudenaarden, 2019; VanHorn and Morris, 2021). Self-targeting approaches, such as homing CRISPR barcodes, were developed that rely on an engineered sgRNA that targets, instead of a target array, the spacer sequence of its own DNA locus (Kalhor et al., 2017, 2018). Canonical CRISPR normally involves a sgRNA and Cas9 protein to target a DNA sequence that matches the sgRNA. For a DNA locus to be edited by a CRISPR/Cas9 there are two requirements: its sequence has to be complementary to the sequence of the sgRNA and it has to contain a protospacer adjacent motif (PAM) recognized by Cas9. The PAM is not present on the sgRNA sequence, therefore the sgRNA locus cannot be targeted by the Cas9/sgRNA complex (Hsu et al., 2014). In a paper published in Science in 2018, Kalhor et al. generated a modified sgRNA, referred to as homing sgRNA (hgRNA), containing a PAM sequence, thus enabling Cas9 to target the expression cassette encoding the hgRNA. After a double-strand DNA break is introduced and repaired via NHEJ, the resulting de novo mutated hgRNA locus should continue to be transcribed as a mutated version of the original hgRNA and participate in another cycle of self-targeting mutagenesis. Multiple cycles of transcription followed by cleavage and error-prone repair can occur, resulting in a continuous generation of a unique barcode that enables tracking of cells over time and space. For *in vivo* barcoding, Kalhor and colleagues created a mouse strain, MARC1 (mouse for actively recording cells 1), in which 60 hgRNAs are integrated across the genome. When MARC1 mice are crossed with a strain that constitutively expresses Cas9, barcoding in all hgRNA expressing sites can be initiated (Fig. 13B). As a proof of concept, the authors reconstructed the early lineage-tree of trophectoderm, primitive endoderm and epiblastin during mouse embryogenesis (Kalhor et al., 2018). A Similar approach, named mSCRIBE (mammalian synthetic cellular recorders integrating biological events), was reported, in which a self targeting guide (stgRNA) was used to labeled the cells *in vitro* before implanting them into the mice (Perli et al., 2016). While powerful, these methods do not allow single cell resolution analyses. As an alternative approach, Chan et al. developed a molecular recorder combined with scRNA-Seq readout for mammalian cell lineage tracing. This approach is based on PiggyBac transposase induced integration of a mCherry gene containing, in its 3' UTR, a cassette with three different Cas9 target sequences and their corresponding sgRNAs (Fig. 13C) (Chan et al., 2019).

The establishment of breeding lines from the mouse models described above is not possible due to the high number of random transgene insertions, thus new embryonic manipulations are required for each new experiment (VanHorn and Morris, 2021). To overcome this limitation, Bowling et al. established a new mouse line named CARLIN (CRISPR array repair lineage tracing). In this system, a series of 10 sgRNAs under the control of a U6 promoter are located upstream of a GFP gene driven by the constitutive CAG promoter and containing, in its 3' UTR, the sgRNA target sequences. This cassette was integrated into the Col1a1 locus of mouse embryonic stem cells (ESC) that also express an enhanced reverse tetracycline transactivator from the ubiquitous Rosa26 promoter. They obtained a mouse strain that was subsequently crossed with another line containing an inducible Cas9. Upon doxycycline administration, the authors identified approximately 44,000 distinct barcodes in different cells, which could be analyzed by scRNA-Seq. They used this strategy to demonstrate that, after depletion induced by 5-fluorouracil, only a small number of mouse HSCs were responsible for the full reconstitution of the hematopoietic system (Bowling et al., 2020).

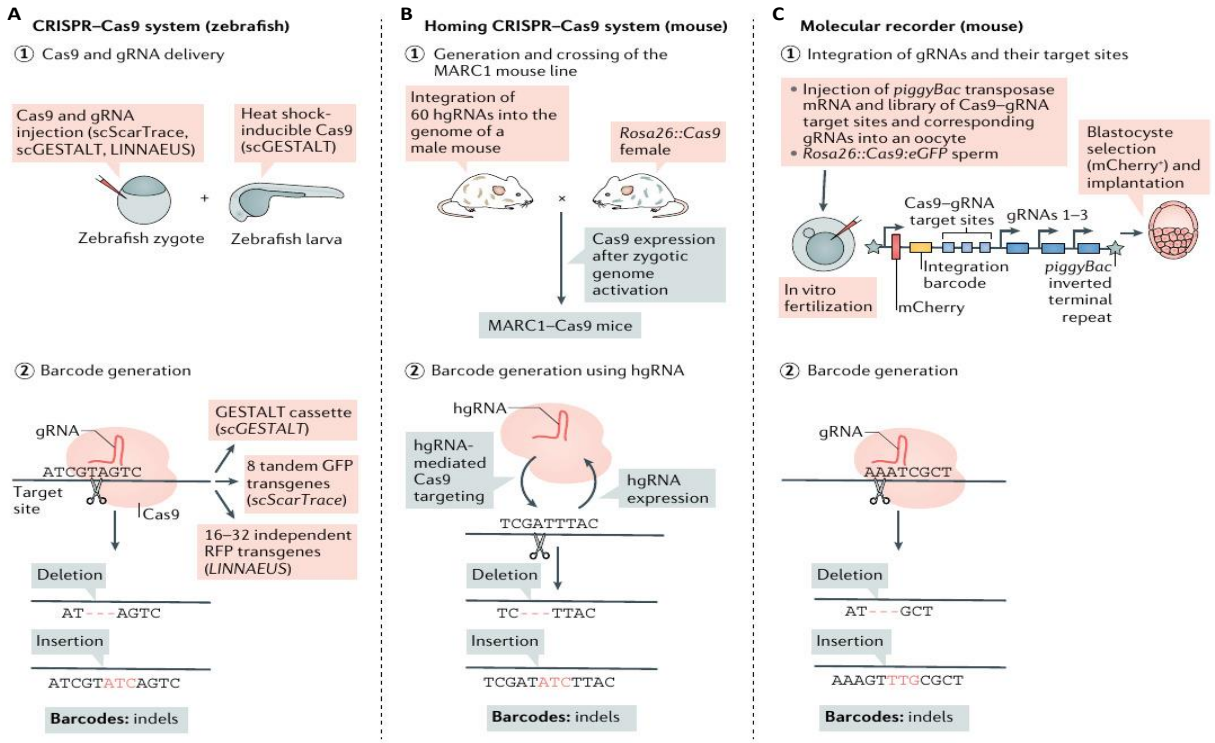


Figure 13. CRISPR-based in vivo barcoding. (A) scGESTALT, scScarTrace and LINNAEUS were used to perform lineage tracing in zebrafish embryos. In these systems, Cas9 and gRNAs directed against

a specific target array are injected into the zygote, which allow the generation of DNA barcodes during early embryogenesis. (B) CRISPR-homing is an attractive strategy to label a large population of cells in mammals, by using an engineered guide RNA that targets its own genomic spacer sequence. In the MARC1 mouse line, 60 hgRNAs are integrated across the genome, which can in theory produce more than 10^{74} different barcodes. (C) As an alternate approach to increase barcode diversity, piggyBac transposase has been used to integrate an array containing Cas9 target sites and encoding the gRNAs targeting them in multiple genomic loci. This system is also compatible with scRNA-Seq analysis. From Baron and van Oudenaarden, Nat. Rev. Mol. Cell Biol. 2019.

Table 4: Overview of the currently available techniques for cellular barcoding.

	Model	**Method**	**Name**	**Redout**	**Application**
In vivo	Mouse	Cre-loxP	Brainbow	Microscopy	Development
			Confetti	Microscopy	
			Polylox	NGS	Hematopoiesis
		Transposase	M2/HSB/Tn	NGS/scRNA-Seq	
	Zebrafish	CRISPR/Cas9	GESTALT	NGS	Development
			scGESTALT	NGS/scRNA-Seq	
			ScarTrace	NGS	Hematopoiesis
			scScarTrace	NGS/scRNA-Seq	
			MEMOIR	FISH	Lineage reprogramming
			LINNAEUS	NGS/scRNA-Seq	Development
	Mouse		mSCRIBE	NGS	Inflammation
			MARC1	NGS	Development
			Molecular recorder	NGS/scRNA-Seq	
			CARLIN	NGS/scRNA-Seq	Hematopoiesis
In vitro	Cell line	Lentiviral	ClonTracer	NGS	Cancer drug resistance
			TraCe-seq	NGS/scRNA-Seq	
			Watermelon	NGS/scRNA-Seq	
			CellTag	NGS/scRNA-Seq	Lineage reprogramming
			LARRY	NGS/scRNA-Seq	Hematopoiesis
		CRISPR/Cas9	CRISPR-barcoding	NGS	Cancer drug resistance

LARRY: lineage and RNA recovery; M2: doxycycline-dependent transactivator; HSB: hyperactive Sleeping Beauty transposase; Tn: transposon; GESTALT: genome editing of synthetic target arrays for lineage tracing, mSCRIBE: Mammalian Synthetic Cellular Recorders Integrating Biological Events;

MEMOIR: memory by engineered mutagenesis with optical *in situ* readout; LINNAEUS: lineage tracing by nuclease-activated editing of ubiquitous sequences; MARC1: mouse for actively recording cells 1; CARLIN: CRISPR array repair lineage tracing; FISH: fluorescence *in situ* hybridization.

5.2.4. Lentiviral barcodes to investigate intratumor heterogeneity

The strategies discussed in the previous paragraphs have been mainly used to investigate the fate of cell lineages during embryonic development and hematopoiesis. As a more suitable approach to study the clonal dynamics of cancer cells, lentiviral vectors can be used to efficiently deliver thousands to millions barcodes that are integrated into the genome of the cells, and thus heritable. The barcode correspond to a highly complex stretch of nucleotides that is inserted in a particular region of the vector. The vectors, each containing a different barcode, are pooled to form a library that can be used to transduce the cells of interest. In the case of completely degenerated sequences, the possible combinations are equivalent to 4^N, where N stands for the number of nucleotides. For example, a 20 bp sequence can yield 4^{20} (~10^{12}) unique barcodes. To avoid the generation of aberrant sequences, such as long repeats or highly unbalanced proportions of G/C or A/T, semi-random pools of DNA barcodes can also be designed, in which certain positions are constrained to one or more specific nucleotides. To ensure that each cell contains only one copy of the vector, and hence one barcode, the lentiviral library is transduced at low multiplicity of infection (MOI), followed by selection of the infected cells, generally using an antibiotic, for which the vector encodes a resistance gene. The barcodes can be "read" by amplifying by PCR the corresponding sequence of the vector from the gDNA of the cells, followed by high-throughput sequencing of the amplicon (Fig. 14) (Bramlett et al., 2020; Kebschull and Zador, 2018).

Short hairpin RNA (shRNA) lentiviral libraries have been used in genome-wide functional screen for almost 20 years (Berns et al., 2004; Paddison et al., 2004). With the advent of CRISPR/Cas9 technology, different strategies based on sgRNA libraries have been developed (Bock et al., 2022; Wang et al., 2014). These screens are designed to identify genes whose perturbation can affect the fitness of the cells by detecting the significantly depleted or enriched shRNAs or sgRNAs, which can also function as barcodes, in response to a selective pressure, *e.g.* treatment with an anticancer drug. With the recognition that intratumor heterogeneity plays an essential role in cancer progression and response to therapy, similar screens have been

designed, in which the viral vector doesn't affect the cell phenotype, but it is only designed to label different cell subpopulations. While random integration of the barcodes into the genome could conceivably alter the behavior of the transduced cells, several studies demonstrated that lentiviral insertion does not cause significant phenotypic changes in the cells (Milone and O'Doherty, 2018; Naik et al., 2014).

The viral barcoding strategy has been extensively used over the last few years to perform lineage tracing and to investigate drug resistance in a wide variety of cancers. For example, Bhang et al. developed a high complexity barcode library, named ClonTracer, to individually label several thousand clones within a mass population of NSCLC cells. To assess whether acquired-resistance to EGFR-TKIs is driven by the emergence of pre-existing or *de novo* clones, they analyzed the barcode composition of different replicates of cells treated in the presence or the absence of erlotinib. They reasoned that, if pre-existing resistant cells are selected during the treatment, a large fraction of shared enriched barcodes should be identified in various replicates. On the other hand, if resistant cells arise *de novo* during the treatment, distinct barcodes are expected to emerge across replicates. The authors found that 40% of the barcodes were shared in multiple culture replicates, implying that EGFR-TKIs resistant clones can be present before the onset of the treatment (Bhang et al., 2015). A similar strategy using a different NSCLC cell line was used by the Engelman laboratory to demonstrate that resistant cells can either derive from rare pre-existing clones or from DTP populations that are capable of surviving during the treatment. These cells can then function as a reservoir for the acquisition of *de novo* mutations that make them fully resistant (Hata et al., 2016). As another example of lentiviral barcode libraries, 'Connie Eaves' laboratory used this approach to investigate the tumor initiating capacity of human breast cancer cells after serial passages in immunodeficient mice. The authors found that only 1/10 to 1/10,000 cells were able to form tumors upon xenotransplantation (Nguyen et al., 2014). In another study from the same group, it was shown that the carcinogenic process in human normal epithelia requires the acquisition of multiple driver mutations, while only KRAS-G12D was crucial for efficient tumor formation. Through lentiviral barcoding experiments, the authors found that heterogeneity occurred very rapidly after the initial transformation event. Within two weeks, they observed dramatic change in the numbers and sizes of clones generated from the initial transformed cells, and noticed the first appearance of

many new clones when cells from the primary tumors were transplanted into secondary recipient mice (Nguyen et al., 2015). As another example of the use of lentiviral barcode libraries to investigate the mechanisms driving intratumoral heterogeneity, Lan et al. showed that glioblastoma cells grafted in the brain of immunodeficient mice display variations in their proliferation and apoptosis rate, as well as their response to the chemotherapeutic agent temozolomide (Temodar, Merck). The authors also found that, when tumor cells derived from a primary tumor undergo serial transplantation, the number of unique surviving clones diminish with each passage. These findings provide supporting evidence that tumorigenesis in glioblastoma is driven by a relatively small population of CSCs (Lan et al., 2017).

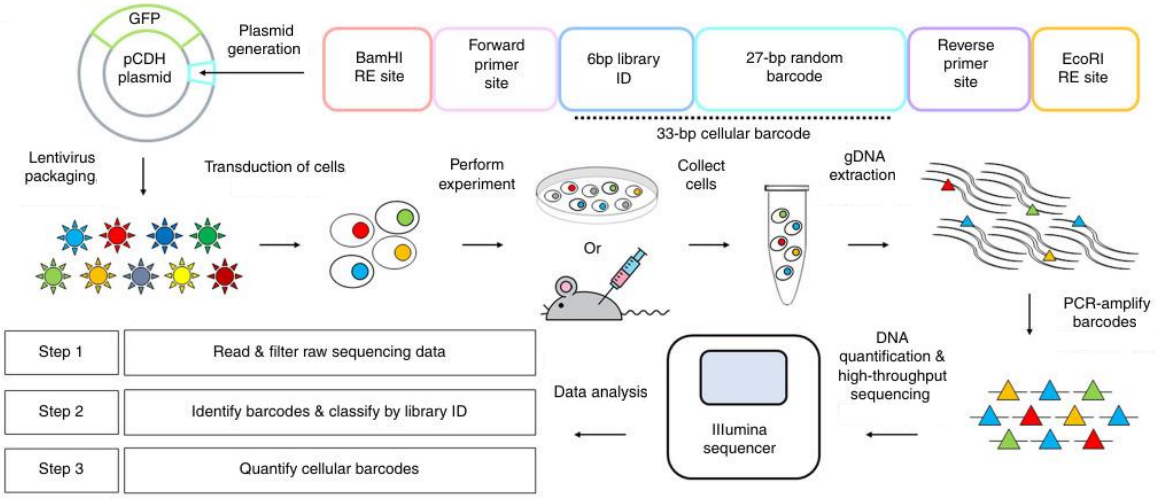


Figure 14. Lentiviral-based in vitro barcoding. A large pool of randomly synthesized oligonucleotides is delivered into the genome using lentiviral vector. The cell lines of interest are transduced with the virus at low multiplicity of infection to label each cell with one unique barcode. The barcoded cells are subjected to selective pressure, *e.g.* treatment with anticancer drug, and the selected populations are harvested at the end of the experiment. The barcodes are then amplified by PCR and quantified by next-generation sequencing. From Bramlett et al., Nat. Protoc. 2020.

Recent studies described lentiviral barcode libraries that are compatible with scRNA-Seq analysis. The 10X Genomics Chromium is one of the most frequently used scRNA-Seq platforms, which provides transcriptional profiling of thousands of individual cells. This system performs rapid droplet-based encapsulation of single cells using a gel bead in emulsion (GEM) approach. With this method, each GEM contains the reverse transcription mix and an oligonucleotide that consists of a unique 10x barcode, a unique molecular identifiers (UMI), sequencing adapters and an oligo(dT).

To enable the simultaneous detection of the barcode and the transcriptome of individual cells, the lentiviral vectors are designed to contain the highly variable DNA sequence in the 3' UTR of a constitutively expressed transgene. The barcodes are thus expressed and they can be identified by 3' scRNA-Seq. In the first examples of this type of strategy, the CellTag and LARRY (lineage and RNA recovery) systems, the DNA barcode is located in the 3' UTR of the GFP mRNA. These libraries were used to investigate the fate of reprogrammed fibroblasts (Biddy et al., 2018) and the different hematopoietic lineages (Weinreb et al., 2020). This strategy has also been employed to investigate drug resistance in cancer cells. Oren and colleagues took advantage of scRNA-Seq and lineage tracing with DNA barcodes to characterize NSCLC DTP cells at single-cell resolution. After two weeks of osimertinib treatment, they identified two main populations of surviving cells: non-cycling DTPs and DTPs that have re-entered cell cycle to divide and form colonies despite drug pressure. To characterize the molecular mechanisms associated with cycling and non-cycling DTPs, the authors developed a system called Watermelon, allowing back-tracing and transcriptional profiling of each cell in the population before and after drug addiction. They showed that these two populations arise from different cell lineages with distinct transcriptional and metabolic programs. Of note, non-cycling DTPs were characterized by the expression of genes associated with cholesterol homeostasis, interferon-α and Notch-signaling. By contrast, the cycling persistent state was characterized by increased expression of the transcription factor NRF2 and decreased levels of reactive oxygen species (ROS). The addition of the ROS scavenger N-acetylcysteine (NAC) was sufficient to increase the proportion of cycling DTPs, consistent with a role of the redox balance in regulating the proliferative ability of these cells. The authors also showed that a switch towards a fatty acid oxidation (FAO) can contribute to the cycling persister phenotype, and that inhibition of FAO using the compound etomoxir reduced the proliferative capacity of DTPs in the presence of osimertinib (Oren et al., 2021).

Chang and colleagues recently developed a similar approach combining DNA barcoding and scRNA-Seq, called TraCe-seq, to compare the effects of conventional EGFR-TKIs with those of GNE-641, a dual EGFR inhibitor-degrader. They found that GNE-641 was less effective than erlotinib and osimertinib in inhibiting NSCLC cell growth, as well as in reducing the absolute number and the diversity of TraCe-seq barcodes, despite similar levels of MAPK

pathway suppression. scRNA-Seq analysis revealed that GNE-641 resistant clones exhibited reduced expression of genes involved in protein processing in the endoplasmic reticulum (ER). The authors showed that the EGFR protein itself plays a crucial role in mediating full cellular efficacy of EGFR-TKIs, as its expression increases ER stress and subsequent pro-death signaling. Consistent with these findings, combination of GNE-641 with low concentrations of ER stress inducers, such as tunicamycin or thapsigargin, strongly enhanced the cytotoxic effects and led to the complete elimination of residual cells. This study uncovered an essential role of the ER protein processing pathway in the response to EGFR targeted therapies (Chang et al., 2022).

5.2.5. CRISPR-barcoding

As an alternative strategy of randomly integrating lentiviral libraries, our laboratory developed the CRISPR-barcoding strategy, in which a DNA barcode can be introduced at a specific genome location through CRISPR/Cas9-induced HDR. In HDR, a donor DNA co-introduced into the cells functions as a template for precise repair: through appropriate design of the donor DNA, this mechanism can be used to generate a wide range of genetic modifications, including specific point mutations or the insertion of an entire gene. Depending on the extent of the desired modification, a single-stranded DNA oligonucleotide (ssODN) or a double-stranded DNA targeting construct can be used as donor DNA for HDR. Dr Grumolato's group used CRISPR-barcoding approach to introduce a short stretch of degenerated nucleotides in the safe harbor locus adeno-associated virus integration site 1 (AAVS1). As illustrated in Figure 15, the cells are co-transfected with a AAVS1 targeting CRISPR/Cas9 plasmid and a ssODN containing nine degenerate nucleotides flanked by homology arms for HDR. The barcodes can be analyzed by amplifying the corresponding region of the AAVS1 locus by PCR from gDNA and deep sequencing of the amplicons. In the proof-of-concept paper, CRISPR-barcoding was used to investigate the tumorigenic potential of different subpopulations of breast cancer cells upon injection in immunodeficient mice and to compare the effects of gefitinib treatment on the clonal architecture of NSCLC cells (Fig. 15) (Guernet et al., 2016).

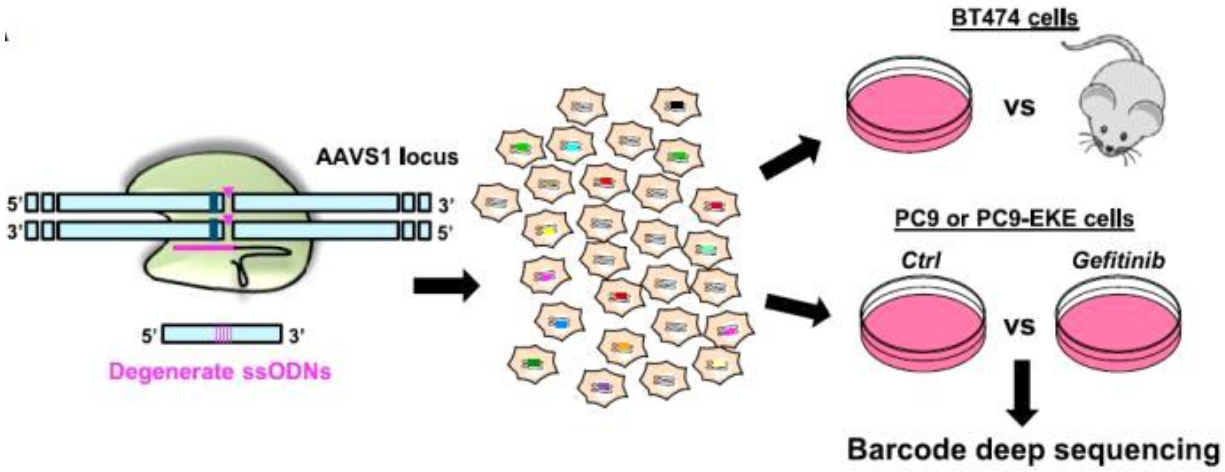


Figure 15. CRISPR-barcoding as a strategy for in vitro barcode delivery. Using CRISPR/Cas9 technology, the barcode sequences are inserted into the AAVS1 locus. From Guernet et al., Mol. Cell. 2016.

Objective of the thesis

A major consequence of intratumor heterogeneity is cancer capacity to adapt to therapy through the emergence of subclonal cell populations, displaying lower sensitivity to the treatment. This is for example the case of NSCLC harboring EGFR activating mutations. While it's becoming more and more evident that individual cells within a population can respond differently to the same treatment, there is still a debate on how this variability in drug response originates. To answer to this question, my Ph.D. project was aimed at comparing the effects of different types of anticancer drugs on the evolution of thousands of individual clones within a mass population of NSCLC cells using highly complex genetic barcodes. This analysis provided a database of barcode profiles for a wide array of known compounds, which can be used as specific signatures to investigate the mechanism of action of new anticancer molecules. Finally, as a complementary approach to investigate the response of different subpopulation of cancer cells, we devised a strategy, named Barcode-Tracker, to identify and isolate clones of interest based on the recognition of a specific genetic barcode. These cells are sorted directly from the original mass population, thus providing the means to isolate NSCLC cell subpopulations resistant, tolerant or sensitive to anticancer treatment, to which they were never exposed. A better understanding of the underlying causes of the heterogeneous drug sensitivity of a mass population of cancer cells should help improve treatment efficacy. The specific aims of my Ph.D. project were:

Aim 1: Elucidate the global effects of anticancer treatment on the clonal architecture of a mass population of NSCLC cells.

Aim 2: Take advantage of intratumor heterogeneity and the predetermined drug response identified in certain cell subpopulations to compare the mechanism of action of different anticancer agents.

Aim 3: Isolate and characterize specific cell subpopulations displaying an intrinsic low or high sensitivity to treatment.

Results

1. Distinct predetermined fates emerge from a heterogeneous cell population in response to anti-cancer therapy

The existence of distinct phenotypic states, such as sensitive or tolerant, is a characteristic of the different cells that form a cancer population. A major question is whether the variability in drug response is driven by intrinsic properties of individual cell subpopulations, or it is a stochastic phenomenon. In other words, are there cells intrinsically primed to survive in the presence of the treatment, or do cells randomly adapt to the drug and emerge?

1.1. Diverse sensitivity to anti-cancer drugs by randomly isolated NSCLC clones

To investigate whether different subpopulations of cancer cells show intrinsically distinct sensitivity to treatment, we derived by limiting dilution various clones (PC9-C1/C2/C3/C4/C5) from a batch of the EGFR-mutant NSCLC PC9 cell line that was never exposed to any drug. In parallel, we generated DTP cells by treating PC9 cells for 3 weeks with 1 μM osimertinib (PC9-OT), 2 μM gefitinib (PC9-GT), 100 nM pemetrexed (PC9-PT) or 500 nM cisplatin (PC9-CT) (Fig. 16A). We then examined the sensitivities of these clones to EGFR-TKIs using CellTiter-Glo. As shown in figure 16B, PC9-C2 displayed higher sensitivity to osimertinib and gefitinib compared to the parental population (PC9-Par), while PC9-C4 showed a profile similar to that of PC9-OT and PC9-GT. We then tested the response of these clones in the presence of chemotherapeutic agents and we found that PC9-C2 was less sensitive to both cisplatin and pemetrexed (Fig. 16B). Similar results were obtained after longer treatment of the cells, as shown by the colony formation experiments illustrated in Figure 16C. To investigate whether the response of the clones was consistent over time, we compared their sensitivity to osimertinib after about 60 passages in culture (PC9-C2_1 and PC9-C4_1) or a freeze-thaw cycle (PC9-C2_2 and PC9-C4_2). As shown in Figure 16D, the response of PC9-C2 and PC9-C4 was not affected by long term culture or freeze-down, indicating that these cells display an intrinsically defined and stable sensitivity to targeted therapy. To confirm the heterogeneous response to EGFR-TKIs and chemotherapy of NSCLC cells from the same mass population, similar experiments were performed using other clones (Fig. 17A-B). Taken together, these data demonstrate that the

tolerant and sensitive phenotypes can constitute an intrinsic feature of certain cell subpopulations that is already present before the onset of the treatment.

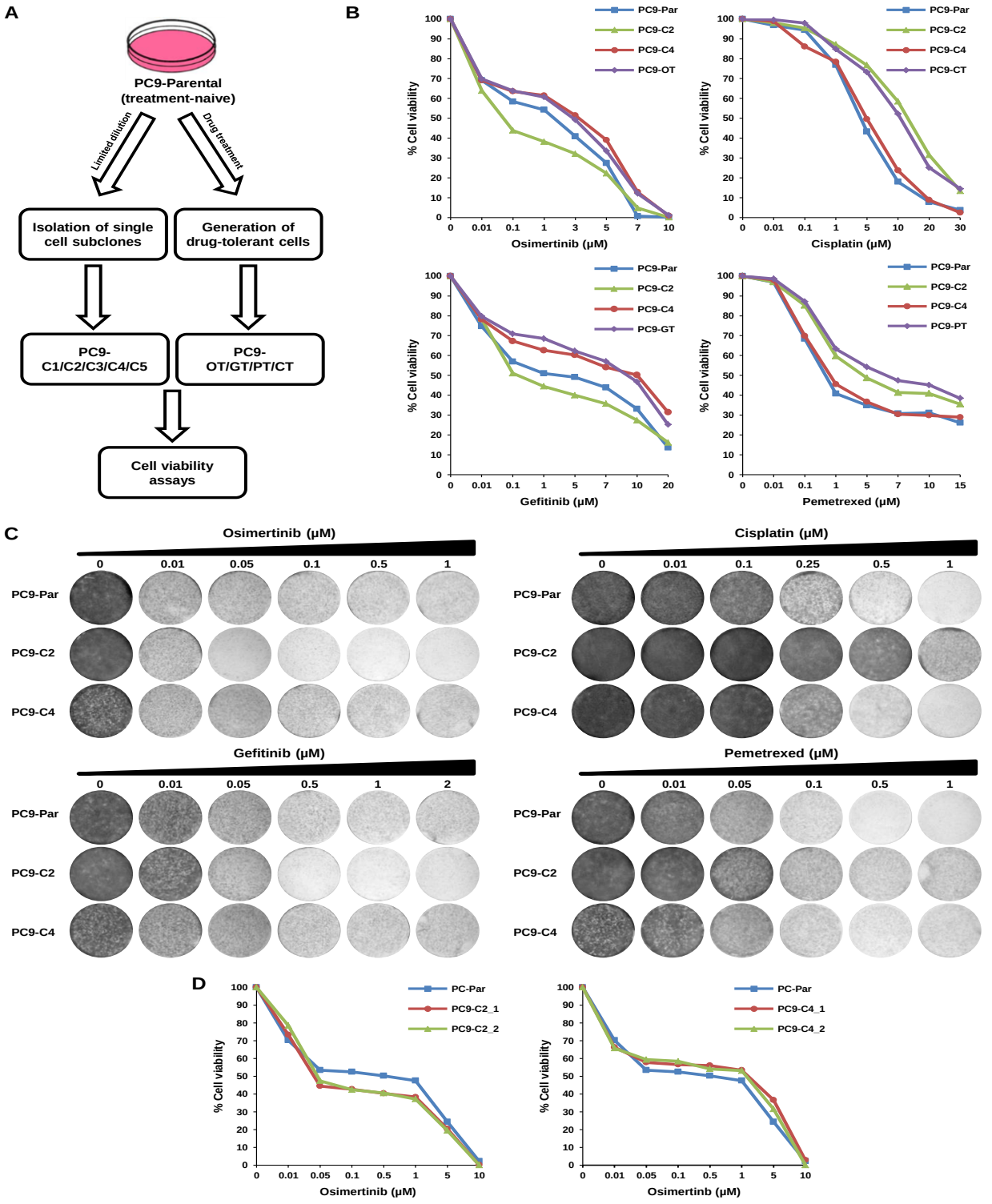


Figure 16. Different sensitivities of PC9-derived clones to anti-cancer drugs. (A) Schematic diagram showing the isolation of PC9-derived clones (PC9-C1/C2/C3/C4/C5) and the generation of PC9 drug-tolerant persister cells grown for several weeks in the presence of osimertinib (PC9-OT) gefitinib (PC9-GT), cisplatin (PC9-CT) or pemetrexed (PC9-PT). (B) PC9 parental (PC9-Par), tolerant and two different clones (PC9-C2 and PC9-C4) were treated for three days with the indicated concentrations of osimertinib, gefitinib, cisplatin or pemetrexed, and their viability was measured by CellTiter-Glo assay (n=3). (C) The cells were treated for 6 days with different concentrations of the indicated drugs, followed by fixation and crystal violet staining (n=3). (D) PC9-C2 and PC9-C4 clones were kept in culture for over 60 passages (PC9-C2_1 and PC9-C4_1) or freshly thawed from a frozen stocks (PC9-C2_2 and PC9-C4_2), then treated for three days with the indicated concentrations of osimertinib. Cell viability was measured using CellTiter-Glo (n=1).

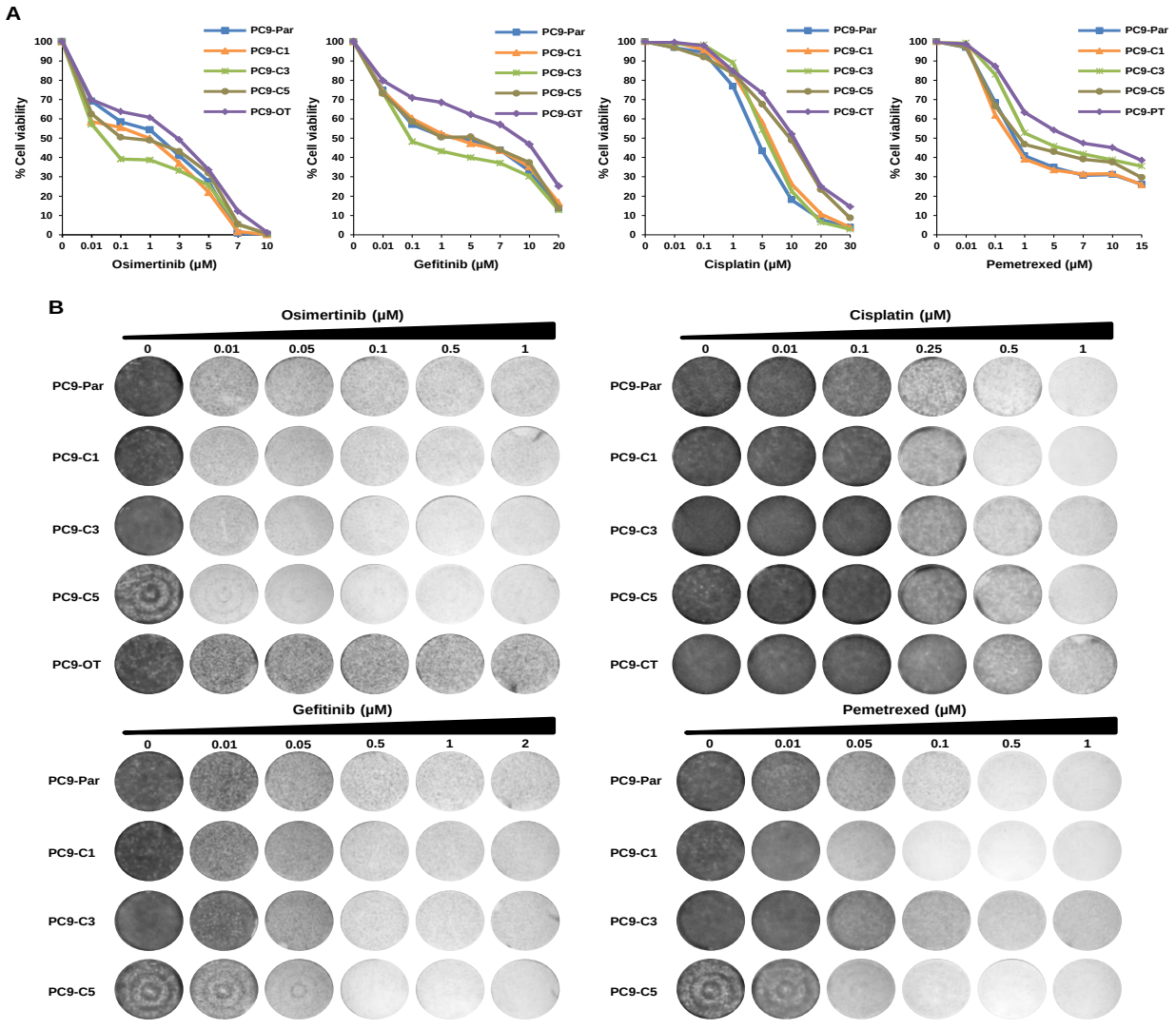


Figure 17. Cell viability assays showing the response of PC9-C1/C3/C5 to anti-cancer therapy. (A) PC9 parental (PC9-Par), tolerant and three different clones (PC9-C1, PC9-C3 and PC9-C5) were treated for three days with the indicated concentrations of osimertinib, gefitinib, cisplatin or pemetrexed, and their viability was measured by CellTiter-Glo assay (n=3). (B) The cells were treated for 6 days with different concentrations of the indicated drugs, followed by fixation and crystal violet staining (n=3).

1.2. Functional characterization of the PC9-derived clones

Next, we asked whether the differences in the sensitivity to osimertinib of PC9 derived clones are associated with specific changes in the EGFR-MAPK signaling pathway. Western blot analysis revealed that the phosphorylation of MAPK and AKT were inhibited by osimertinib, while the expression of pAKT was not affected in PC9-OT cell, consistent with their tolerance to EGFR-TKIs. Of note, in the absence of osimertinib, PC9-C2 displayed increased phosphorylation of ERK and AKT as compared to PC9-C4 and parental PC9 cells, indicating a stronger MAPK basal activation in this clone (Fig. 18A). To confirm this observation, we measured the expression of a well-established MAPK target gene, the dual specificity phosphatase 6 (DUSP6) (Pratilas et al., 2009). DUSP6 belongs to the family of MAPK phosphatases that negatively regulates EGFR signaling and its downstream effector RAS, in part by inducing the dephosphorylation and inactivation of ERK1/2 (Caunt and Keyse, 2013). PC9-OT, PC9-C2, PC9-C4 and parental PC9 cells were treated in the presence or the absence of osimertinib for 72h, and the levels of DUSP6 were assessed by q-RT-PCR. As shown in Figure 18B, in the absence of osimertinib the expression of DUSP6 was more than three fold higher in PC9-C2 compared to all the other cells, consistent with higher MAPK basal activation in this clone, which may explain its increased sensitivity to EGFR inhibition.

To gain insight into the different susceptibilities to chemotherapy observed in our PC9-derived clones, we assessed the effects of pemetrexed on the response to DNA damage of these cells. We found that, compared to PC9-C4, PC9-OT and parental PC9, PC9-C2 and PC9-PT cells have reduced levels of phosphorylated histone H2AX (γH2AX) in the presence of pemetrexed (Fig. 18C). These data suggest that PC9-C2 and PC9-PT cells are capable of maintaining low levels of DNA damage in response to genotoxic stress, which could explain their decreased sensitivity to chemotherapy.

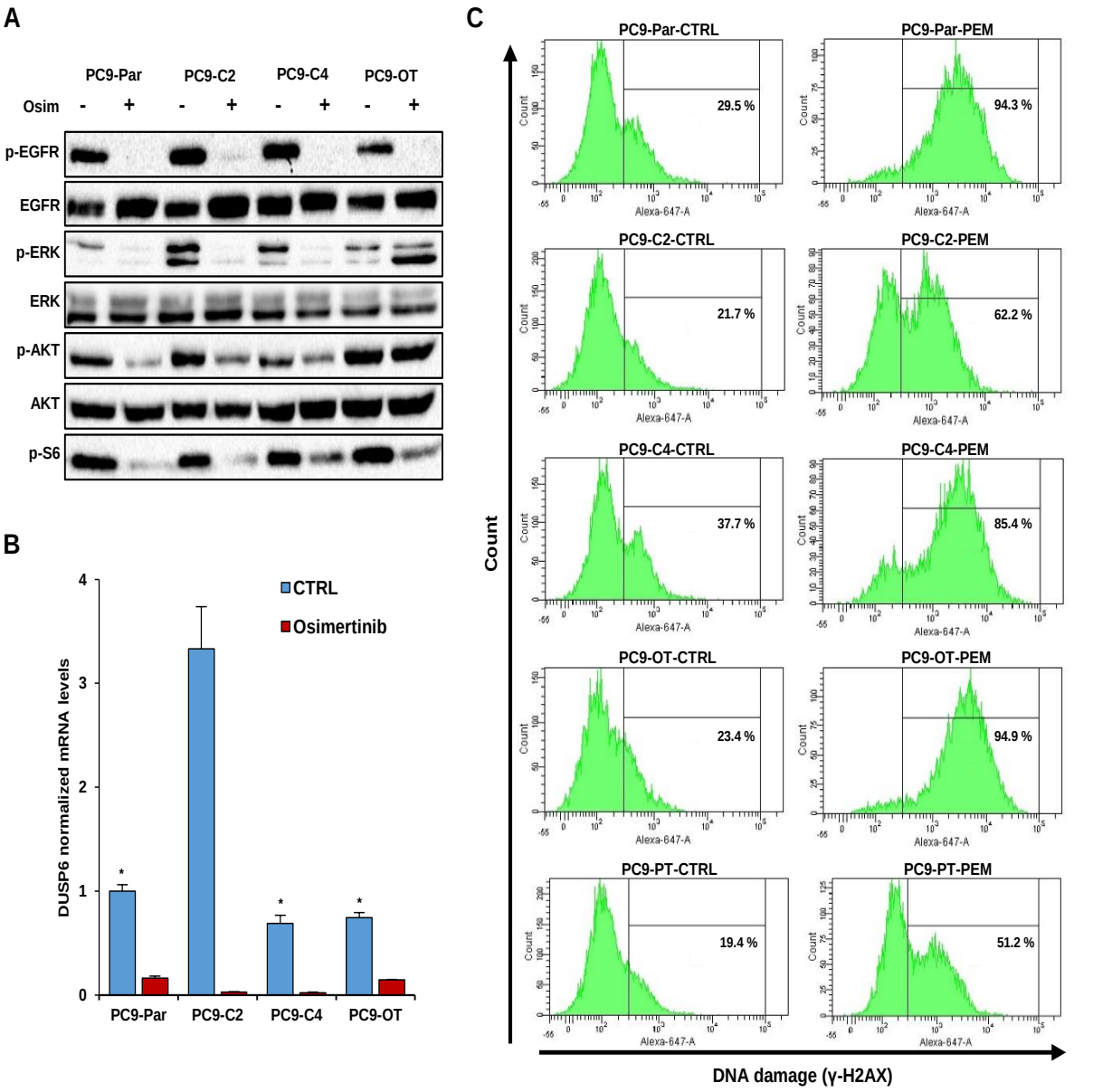


Figure 18. Characterization of PC9-derived clones. (A) PC9-Par, PC9-C2, PC9-C4 and PC9-OT cells were treated for 48 h in the presence or the absence of osimertinib (100 nM), and immunoblot was performed using the indicated antibodies (n=1). (B) The mRNA levels of DUSP6 in the same cells treated for 72h with or without osimertinib (100 nM) were assessed by q-PCR. The mean ± SEM of one representative of three independent experiments is represented. *p-value < 0.05 (Mann-Whitney test). (C) The cells were treated for 48h in the presence or the absence of pemetrexed (100 nM) and the levels of γH2AX were measured by FACS (n=1).

1.3. Highly complex CRISPR-barcodes reveal predetermined phenotypes in response to different treatments

Our data on PC9 clones argue that the response to treatment of individual cancer cells can be predetermined. To test this hypothesis with a different strategy, we used CRISPR-barcoding to label different subpopulations of PC9 cells (Guernet et al., 2016). Through insertion of a highly complex DNA barcode in the AAVS1 locus, we compared the effects of both targeted therapy and chemotherapeutic agents on the relative proportion of several thousand clones. Five replicates of PC9 barcoded cells were treated with or without the EGFR-TKIs gefitinib or osimertinib, the MEK inhibitor trametinib or the chemotherapeutic drug pemetrexed. After 2 weeks of treatment, the cells were collected, the barcodes were amplified from gDNA by PCR and analyzed by NGS (Fig. 19A). While certain subpopulations were differently over/under-represented across replicates, we found that a subset of barcodes was consistently enriched in at least one condition, implying that certain cells are intrinsically less sensitive to a given drug (Fig. 19B). These findings were confirmed by hierarchical clustering and by Spearman's correlation analysis (Fig. 19C-D). Indeed, we observed high correlation levels between replicates of the same condition, indicating that certain subpopulations consistently display the same response to the drug. Of note, we also found high correlation between cells treated with the two EGFR-TKIs and, to a lesser extent, trametinib, consistent with the fact that these drugs target the same signaling pathway. On the contrary, cells treated with pemetrexed showed low correlation with those treated with EGFR or MEK inhibitors, indicating that distinct pre-existing subpopulations were enriched in response to different treatments. These data imply that certain NSCLC cells display a specific and predefined response to the treatment, which may result in their capacity to drive resistance to therapy.

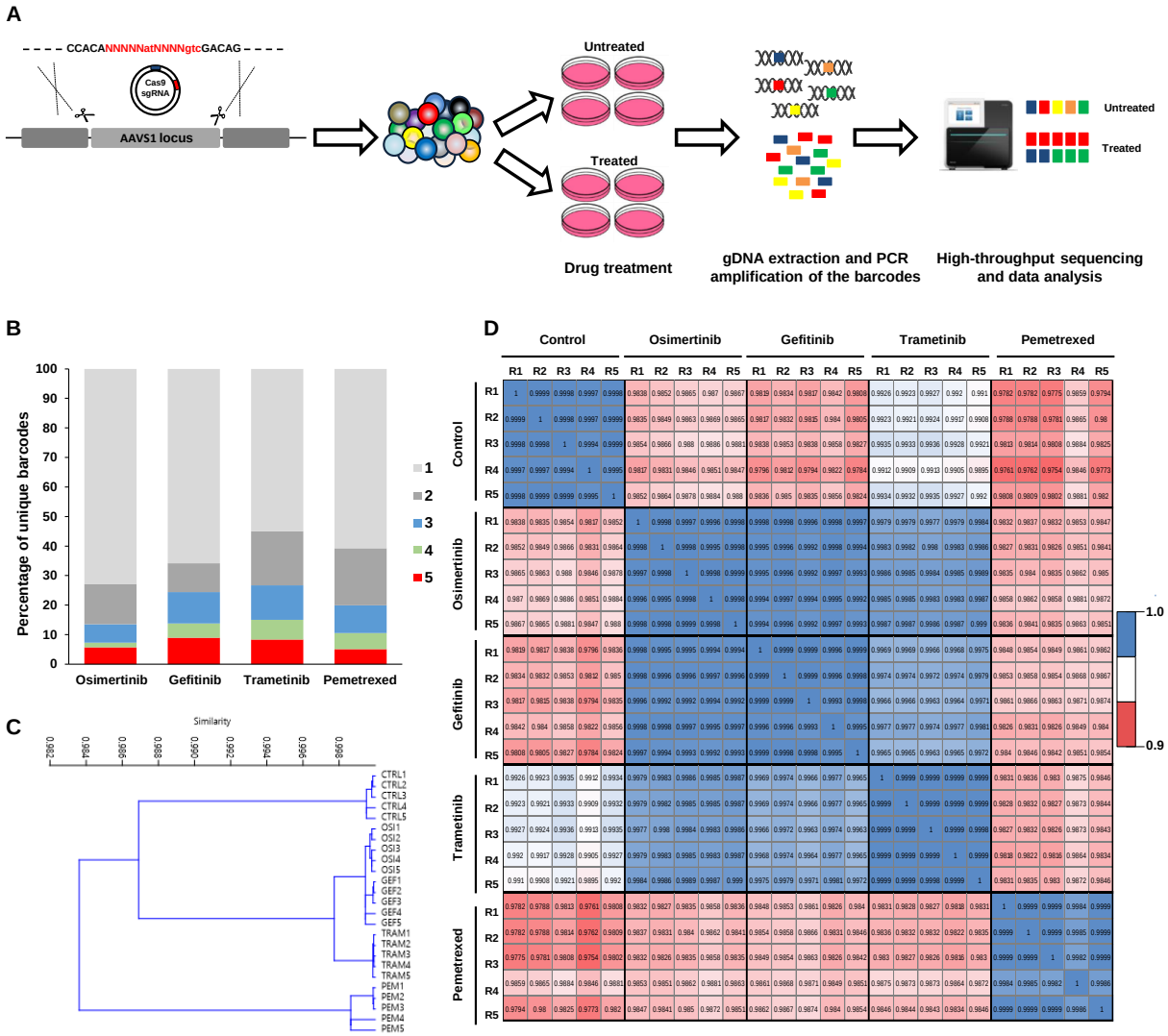


Figure 19. Clonal dynamics of PC9 cells in response to anti-cancer treatment using highly complex CRISPR-barcodes. (A) A partially degenerate nucleotide sequence is inserted in the AAVS1 locus through CRISPR/Cas9 technology, to individually label several thousands of cells within the mass population. Different replicates of these cells were then treated for two weeks in the presence or the absence of osimertinib (100 nM), gefitinib (2 µM), trametinib (30 nM) or pemetrexed (100 nM), followed by analysis of the barcode distribution by deep sequencing. (B) Percentage of enriched barcodes in each replicate following indicated treatment in PC9 barcoded cells. (C-D) Barcode hierarchical clustering tree (C) and Spearman's pairwise correlation analysis (D) of positively selected barcodes generated from the same experiment (CTRL: control; OSI: osimertinib; GEF: gefitinib, TRAM: trametinib PEM: pemetrexed).

1.4. Potential involvement of epigenetic modifications on drug response

To assess the role of epigenetics in the heterogeneous response of NSCLC cells, we evaluated the effects of the histone deacetylase inhibitor Trichostatin A (TSA) on drug-induced PC9 clonal evolution. We reasoned that a pretreatment with TSA should have a global effect on histone marks, resulting in epigenetic reprogramming. Control and TSA-pretreated cells were then grown for two weeks in the presence or the absence of osimertinib or pemetrexed, followed by NGS analysis of the barcodes (Fig. 20A). For TSA pretreatment, we used a duration and a concentration that didn't inhibit cell growth, while strongly increasing H3K27 acetylation (Fig. 20B-C). Analysis of the barcode distribution between the different conditions revealed that TSA pre-treatment modified clonal evolution induced by osimertinib, but not pemetrexed (Fig. 20D-E). In particular, in TSA-pretreated cells, we found a 10% decrease in the number of osimertinib consistently enriched subpopulations across replicates, and a 15% increase in depleted clones (data not shown), suggesting a global enhancement of EGFR-TKIs sensitivity. We also found that this type of reprogramming did not affect the response of cells to pemetrexed (Fig. 20D-E). Taken together, these findings indicate that histone acetylation plays a role in the predetermined tolerance to EGFR-TKIs, but not chemotherapy, of certain subpopulations of NSCLC cells.

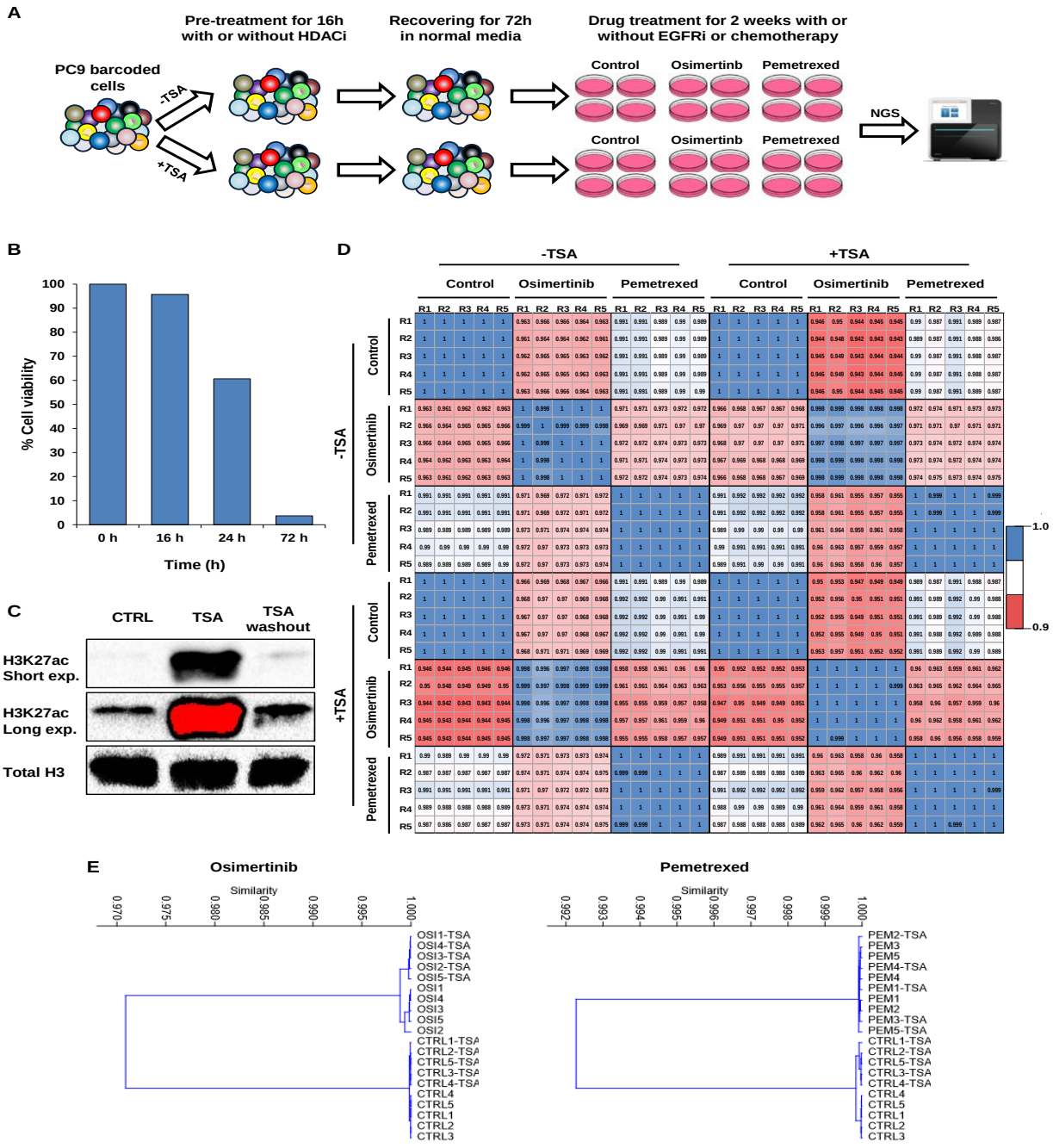


Figure 20. Effects of TSA pre-treatment on the clonal evolution of PC9 cells. (A) Schematic representation of the epigenetic reprogramming strategy. (B) PC9 cells were treated with 2 µM TSA for the indicated time points and cell viability was measured using CellTiter-Glow. (C) PC9 cells treated for 16h in the presence or the absence of TSA (2 µM), followed by 72h washout. Acetylation of histone H3K27 was assessed by western blot. (D-E) PC9 barcoded cells were treated as illustrated in A and the barcode distribution was assessed by NGS. Spearman's pairwise correlation analysis (D) and barcode hierarchical clustering tree (E) of positively selected barcodes are shown (CTRL: control; OSI: osimertinib; PEM: pemetrexed, TSA: trichostatin A).

2. High throughput screen to investigate the heterogeneous response of different subpopulations of cancer cells to anti-cancer therapies

Our CRISPR-barcoding data indicate that certain subpopulations of PC9 NSCLC cells show a consistent and predetermined response to treatment that depends on the compound used. In the second part of my thesis, we decided to use a different strategy for cell tracing. While CRISPR-barcoding can be a useful tool to trace cellular heterogeneity, it doesn't provide the high levels of resolution required for certain applications. Indeed, even after optimization of HDR efficiency, only part of the cell population contains highly complex barcodes (around 10-15% of the cells) in our PC9 cell experiments (Guernet et al., 2016). Most of the other cells display indels generated by NHEJ, which can be used as barcodes, even if the same sequence can be conceivably shared by different subpopulations. Finally, another limitation is related to the fact that, since genome editing can affect one or both alleles, some cells could contain two distinct barcodes. To overcome these potential drawbacks, we further investigated the heterogeneous response to treatment of NSCLC cells using a highly complex lentiviral barcode library that we designed and generated.

2.1. Generation of a highly complex barcode library to label PC9 cells

To trace the fate of different subpopulations of PC9 cells, we designed a lentiviral library containing a semi-random 38 base-pair long sequence functioning as a DNA barcode. This sequence lies in the 3' UTR of the puromycin resistance gene to enable transcription of the barcodes (Fig. 21), thus making the library compatible with single-cell analysis of gene expression. We generated a library containing approximately one million barcodes and used it to transduce 2 x 10^6 PC9 cells with a multiplicity of infection (MOI) of less than 0.3, to ensure that most of the transduced cells contain a unique barcode. After infection, we cultured the cells for one week in the presence of puromycin to eliminate untransduced cells. After selection, we randomly isolated 25,000 cells and rapidly expanded them to obtain a sufficiently large population that was frozen down in more than 20 vials stored in liquid nitrogen. This step was performed to decrease the complexity of the PC9 cell population, thus facilitating the analysis of the effects of a large number of different compounds. In a typical experiment, a vial of barcoded

cells is thawed and expanded for two-three passages. The cells are then plated in quadruplicate for each condition, and treated for nine days with a panel of drugs of interest. The cells are subsequently harvested and the gDNA is extracted. The sequence of the vector containing the barcode is amplified by PCR from each gDNA sample and the amplicons are sequenced on an Illumina MiniSeq. As mentioned above, our library is compatible with scRNA-Seq analysis, through association of the gene expression profile of individual cells to their unique barcode (Fig. 21).

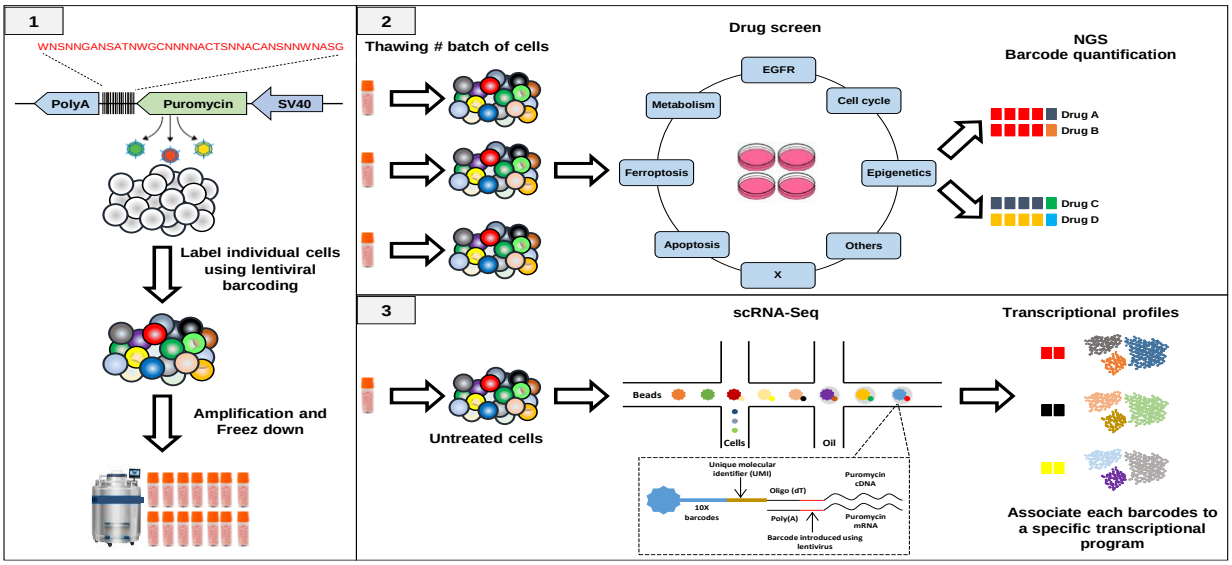


Figure 21. Lentiviral barcode strategy to investigate the variability of drug response within individual cancer cells. (1) The lentiviral library consists of semi-random 38-bp barcode sequence, inserted in the 3' UTR of the puromycin resistance gene. PC9 NSCLC cells were transduced with the library at low multiplicity of infection to label each cell with one unique barcode. 25,000 cells were then isolated, amplified and cryopreserved in liquid nitrogen. (2) Thawed vials are plated in different replicates and treated with the compounds of interest. After treatment, the cells are harvested and gDNA is extracted. The amplicons containing the barcodes are amplified by PCR and subjected to NGS to determine the relative abundance of the barcodes in each sample. (3) This system is compatible with scRNA-Seq analysis.

2.2. The barcode profile of different compounds is consistent across experiments

To assess whether the response of the different subpopulations is consistent over time, we performed two independent experiments from two different vials of frozen PC9 barcoded cells. The cells were thawed, expanded in culture and then treated in the presence or the absence of osimertinib, trametinib or pemetrexed (4 replicates/conditions). After treatment, gDNA was

extracted and the barcode distribution was analyzed by NGS (Fig. 22A). We found that the effects of the drugs on the barcode profile was consistent not only between replicates of the same experiment, but also across the two experiments. As shown in Figure 22B, Spearman's ranking of positively selected barcodes showed high degree of correlation between cells treated with osimertinib in the two experiments, indicating that the response of these cells to osimertinib are stable over time. Similar results were obtained from cells treated with trametinib and pemetrexed (Fig. 22B). These data are in line with our previous findings and indicate that some subpopulations display a specific and predetermined response to the treatment.

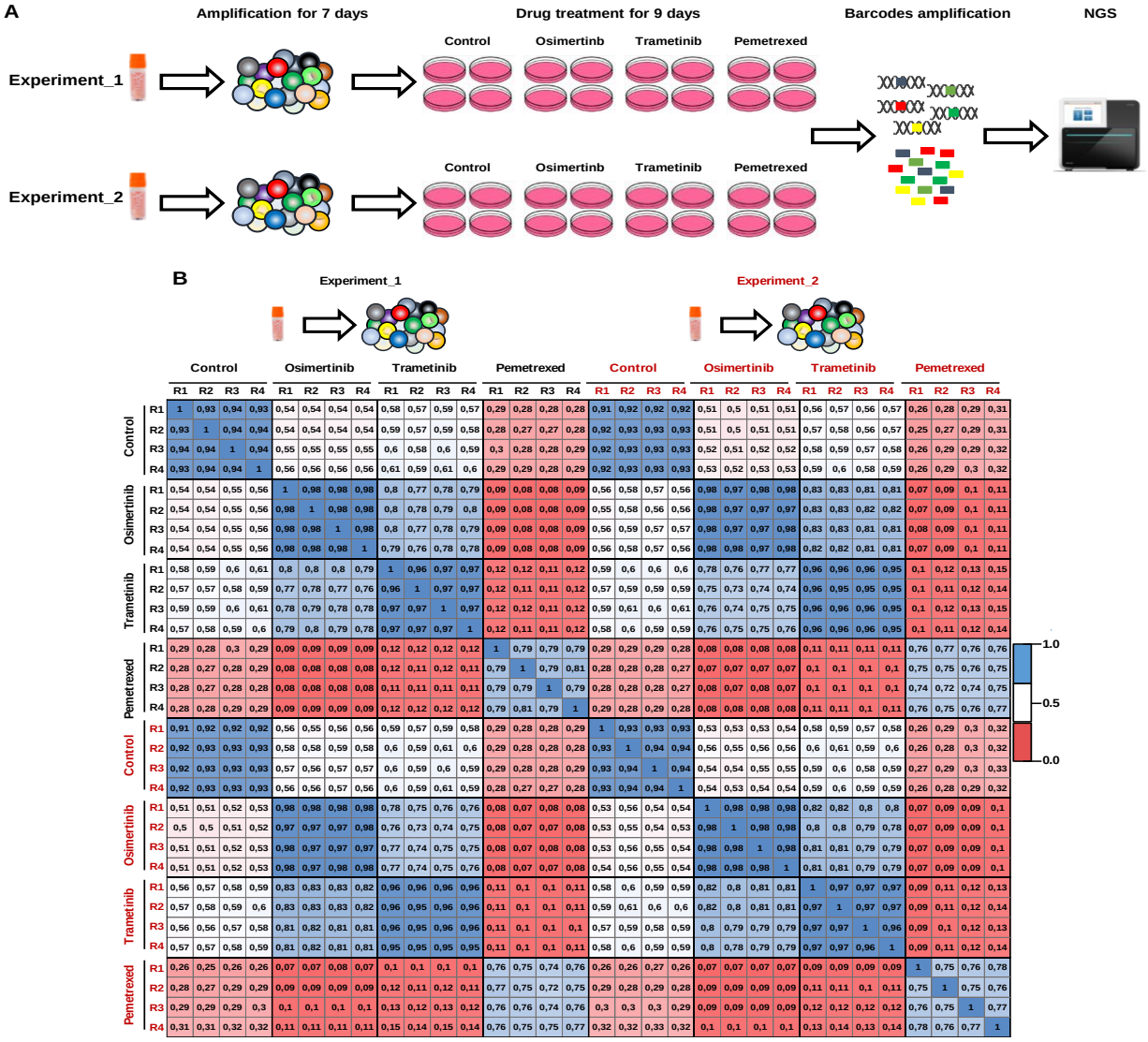


Figure 22. The drug response of the cell population is stable over time. (A) Two different vials of barcoded PC9 cells were thawed, amplified for one week, and treated with or without osimertinib

(100 nM), trametinib (30 nM) or pemetrexed (100 nM) for 9 days. (B) Spearman's pairwise correlation analysis of positively selected barcodes revealed a high degree of correlation between samples treated with the same drug across the two experiments.

2.3. Distinct pre-existing subpopulations emerge in response to different treatments

To further investigate the response of cancer cells to different anti-cancer drugs, we screened the PC9 barcoded cells against a custom library of compounds that were selected to target a wide range of pathways. The library consisted of 87 small molecules, including several kinase inhibitors and other compounds used in the clinic (Table 5).

Table 5. Classification of the drugs used in the screen based on their mechanism of action.

Chemotherapy	**EGFR**	**Epigenetics**	**Multi-kinase**	**RAF**
Pemetrexed Methotrexate Pralatrexate 5-Fluorouracil Carboplatin Oxaliplatin Paclitaxel Cisplatin Doxorubicin Mitomycin C	Osimertinib WZ4002 Rociletinib Lazertinib Mavelertinib Gefitinib Gefitinib-PROTAC	Vorinostat Trichostatin A Sodium lbutyrate Sodium phenylbutyrate Tazemetostat 5-Azacytidine	Sorafenib Regorafenib Sunitinib Lenvatinib Cabozantinib	LY3009120 Vemurafenib RAF265 CCT196969
Antibiotics	**Autophagy**	**JAK/STAT**	**AMPK**	**ER stress**
Bafilomycin A1 Trimethoprim Nigericin	Chloroquine Spautin-1 SBI-0206965	Napabucasin Stattic V AZD1480	Ibudilast Phenformin HCL Dorsomorphin	CCT020312 Tunicamycin Thapsigargin
Ferroptosis	**Apoptosis**	**MEK**	**Proteasome**	**Gap junction**
Erastin Ferrostatin-1 Deferoxamine	S63845 Navitoclax	U0126 Trametinib	MG-132 Bortezomib	Mefloquine Probenecid
CDK	**MNK**	**ALK**	**ROCK**	**Wnt**
THZ1 Palbociclib	eFT-508 CGP57380	Crizotinib TAE684	LIMKi 3 Y-27632	XAV-939 LGK-974
mTORC	**Mitochondrial biogenesis**	**Adrenergic receptor**	**MDM2 antagonist**	**ERK1/RasGAP/ RSK**
Temsirolimus	Levofloxacin	L-755,507	Nutlin-3	Pluripotin (Sc-1)
Notch	**CaM-Kki**	**FGFR**	**GPX4**	**PI3K**
IMR-1	STO-609	Infigratinib	ML-210	AZD8186
BRCA	**AKT**	**RAS**	**Antioxidant**	**SHP2**
Olaparib	Capivasertib	Sotorasib	N-acetylcysteine	SHP099 HCL
Src	**AURKA**	**IGFR**	**NFKB**	
Saracatinib	Alisertib	BMS-536924	JSH-23	

The screen was performed using 9 different vials of PC9 cells, yielding around 505 samples, including 40 control samples. After treatment, the cells were collected, gDNA was extracted, and the amplified barcodes from the 505 samples were sequenced over 11 NGS runs. In collaboration with Dr. Vera Pancaldi (Centre de Recherches en Cancérologie de Toulouse), the barcode sequences were isolated after sequencing and processed for analysis. We observed high correlation between samples treated with the same compound, indicating the emergence of pre-existing populations in response to each drug (Fig. 23). We also found low correlation between cells treated with inhibitors targeting different pathways, indicating that distinct pre-existing subpopulations were enriched in response to different treatments.

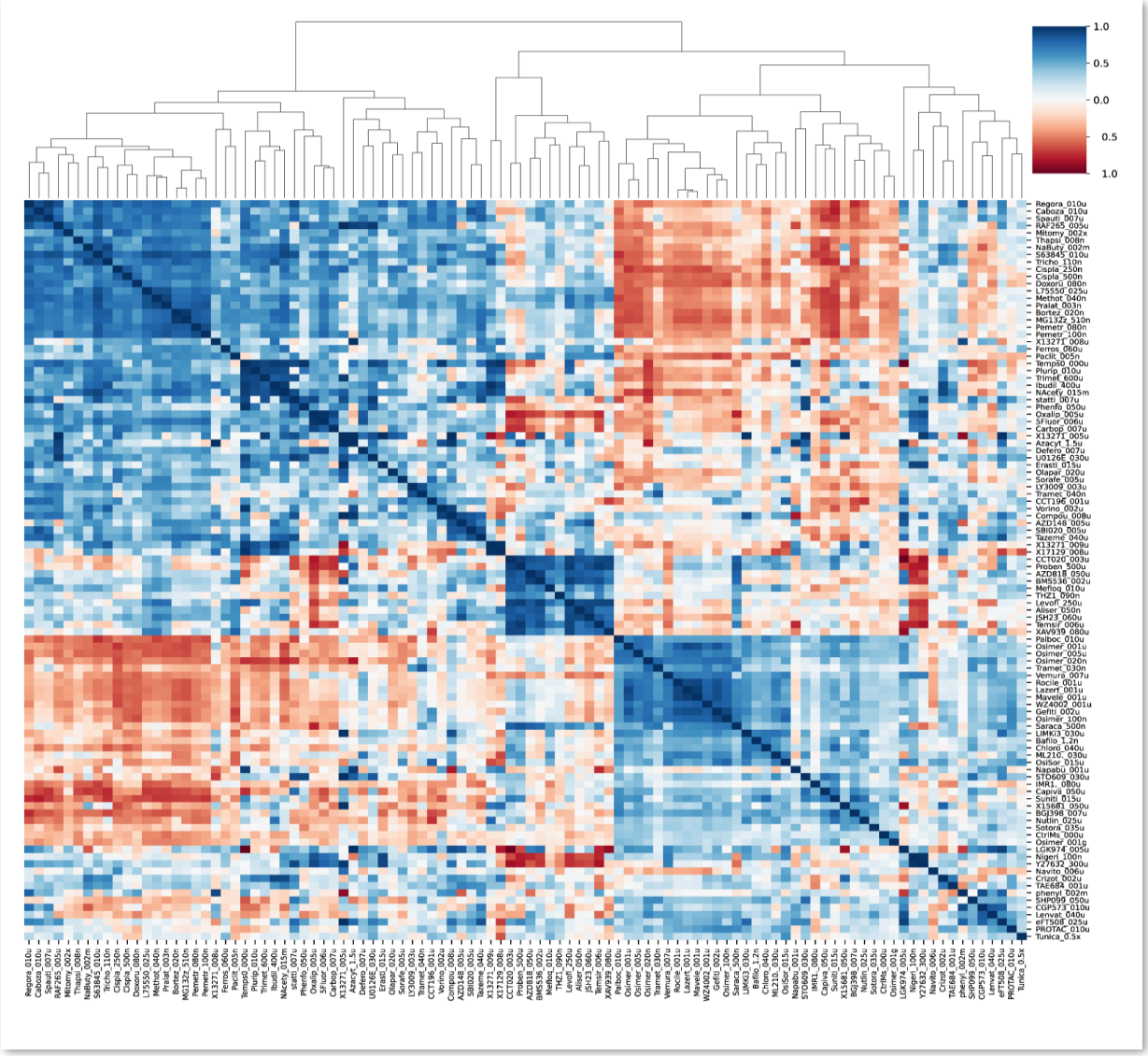


Figure 23. Drug-drug correlation clustered heatmap. The PC9 barcoded cells were treated with a panel of 87 anti-cancer drugs for 9 days. After treatment, the barcodes were amplified from gDNA and sequenced on Illumine MiniSeq. The barcode sequences were compared between different conditions. Each column in the heatmap represents one condition and shows the Pearson correlation to all other conditions (including itself), with blue for correlation = 1, red for correlation = -1 and white for 0 or insignificant correlation.

Our data showed that drugs targeting similar pathways were clustered together. As shown in Figure 24, we observed high correlation between cells treated with different generations of EGFR-TKIs, indicating that these drugs select for the same pre-existing subpopulations. We noticed that the correlation between samples treated with first and third generations EGFR-TKIs was higher as compared to cells treated with the EGFR-degraders gefitinib-based PROTACs, since these drugs inhibit EGFR using different mechanisms. Consistent with our previous data, high correlation was also found between cells treated with EGFR-TKIs and those treated with drugs targeting downstream components of the pathway, such as the MEK. To further confirm this finding, we treated the cells with the Src inhibitor saracatinib. Src is a non-receptor tyrosine kinase that can phosphorylates EGFR at the tyrosine 845 residue, which promotes signaling downstream of the receptor (Ortiz et al., 2021). We observed high degree of correlation between cells treated with saracatinib and the different EGFR inhibitors, consistent with the major role of Src in EGFR signaling. Of note, we found a lower correlation with inhibitors of PI3K and mTOR, consistent with previous data indicating that EGFR promotes PC9 cell growth mainly through activation of the MAPK pathway (Fig. 24) (Tricker et al., 2015).

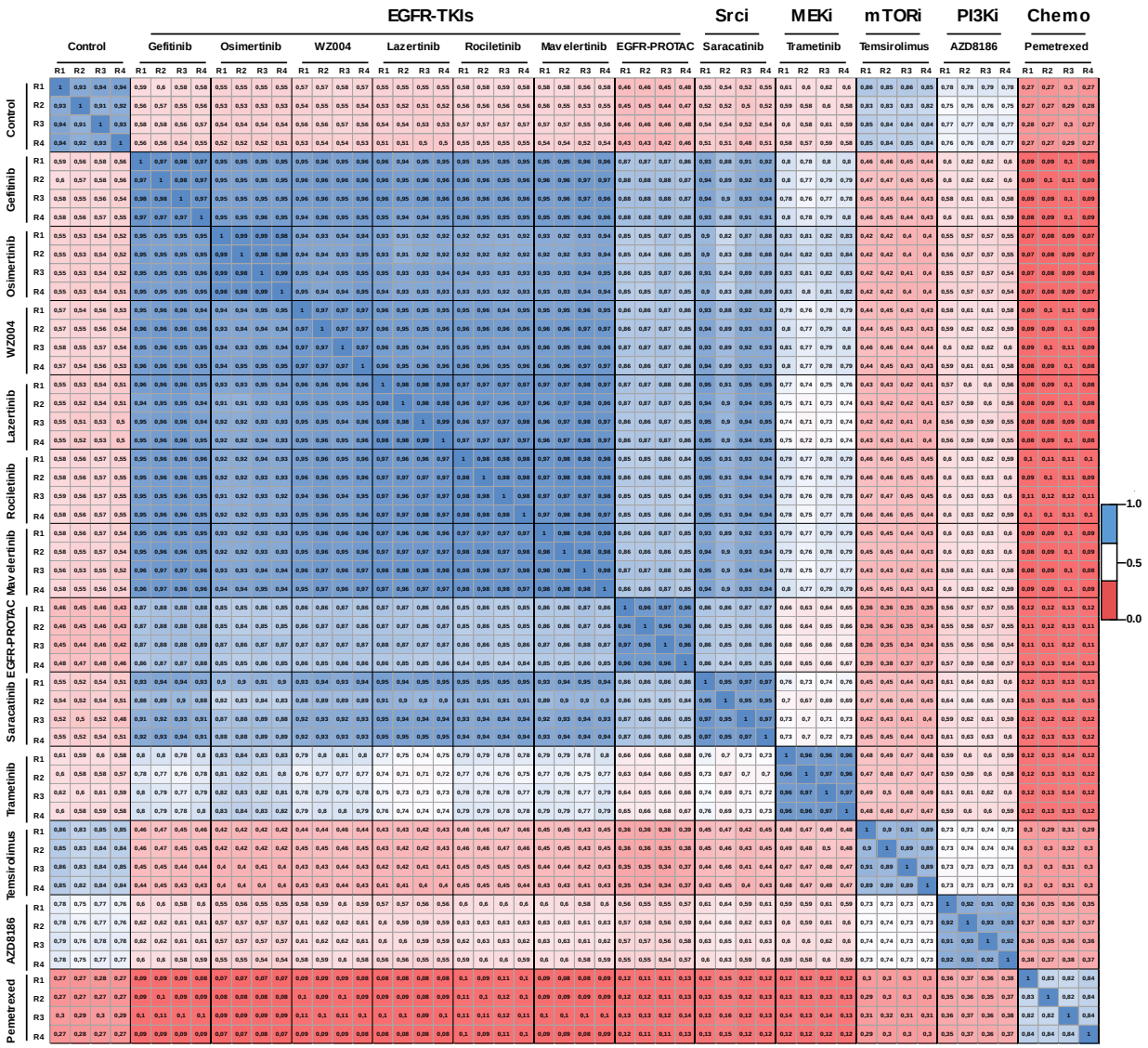


Figure 24. Effect of different drugs targeting the EGFR and its downstream signaling pathways on the clonal evolution of PC9 NSCLC. PC9 barcoded cells were treated in the presence or the absence of gefitinib (2 μM), osimertinib (100 nM), WZ4002 (1 μM), Rociletinib (1 μM), Lazertinib (1 μM), Mavelertinib (1 μM), EGFR-PROTAC (10 μM), Saracatinib (500 nM), trametinib (30 nM), AZD8186 (50 μM), temsirolimus (6 μM) or pemetrexed (100 nM) for 9 days. Spearman's pairwise correlation analysis of positively selected barcodes is represented.

2.4. Investigate the clonal dynamics of PC9 cells in response to different concentration of EGFR-TKIs

We next asked whether changing the drug dose may affect the responsiveness of the different subpopulations of cancer cells. To test this hypothesis, we treated PC9 cells containing

the lentiviral barcode library with different concentrations of osimertinib, ranging from 20 nM to 5 µM. After 9 days of treatment, the cells were collected and gDNA was extracted from each condition. The barcodes were then amplified and sequenced. As shown in Figure 25, cells cultured in the presence of 20 nM osimertinib, which is sufficient to efficiently inhibit EGFR phosphorylation in PC9 cells (data not shown), showed high correlation with cells treated with concentrations of 100 nM and 1 µM. Conversely, we found lower correlation in the presence of 5 µM osimertinib (Fig. 25), probably reflecting a decreased specificity of the treatment at high drug concentrations.

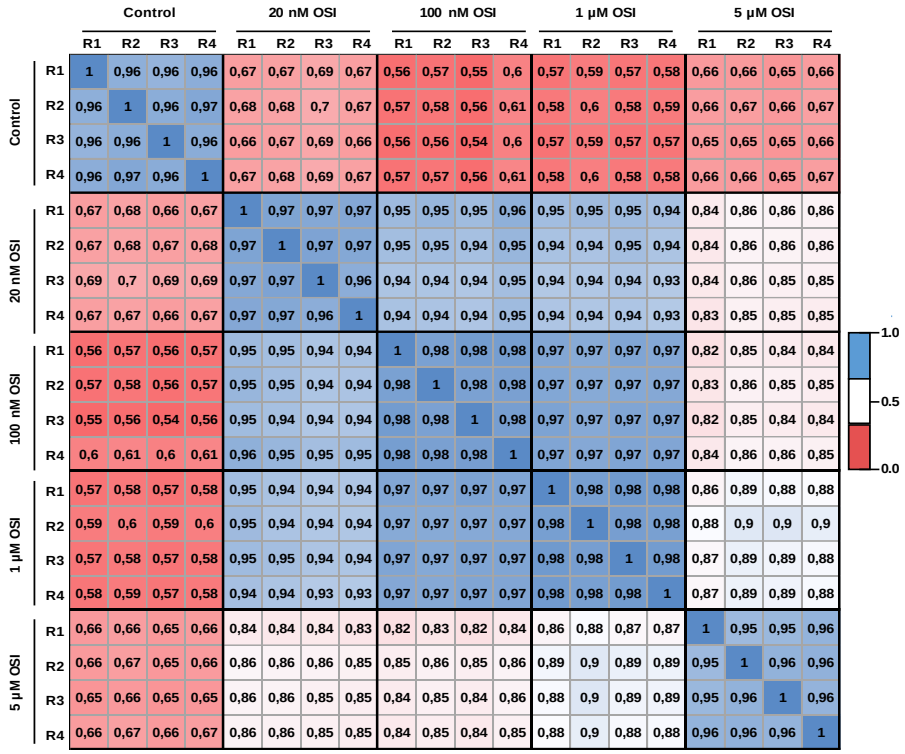


Figure 25. Clonal dynamics of PC9 NSCLC in response to different concentrations of osimertinib. PC9 cells were treated in the presence or the absence of 20 nM, 100 nM, 1 µM or 5 µM osimertinib for 9 days. Spearman's pairwise correlation analysis of positively selected barcodes was generated.

2.5. Comparison of our dataset with other drug similarity scores

In this study, we used the barcode profiles to compare the effects of different compounds and showed similarity between drugs targeting the same pathway. To compare these data with those obtained with other types of screen, we used DrugSimDB, a database that integrates

several drug similarity scores based on various criteria. We filtered our correlation matrix for drug pairs with a correlation coefficient > 0.4 or < -0.4, resulting in all significant interactions between the 87 drugs. Out of the possible interactions between them, 82 pairs are annotated as interacting in DrugSimDB. We found that our dataset correlates best with the drug similarities based their chemical structure or their shared targets (Fig. 26A-B). We next generated a network of 27 drugs in common between the DrugSimDB and our dataset (Fig. 26C-D).These data indicate that drug similarities generated using our strategy are consistent with previously published dataset.

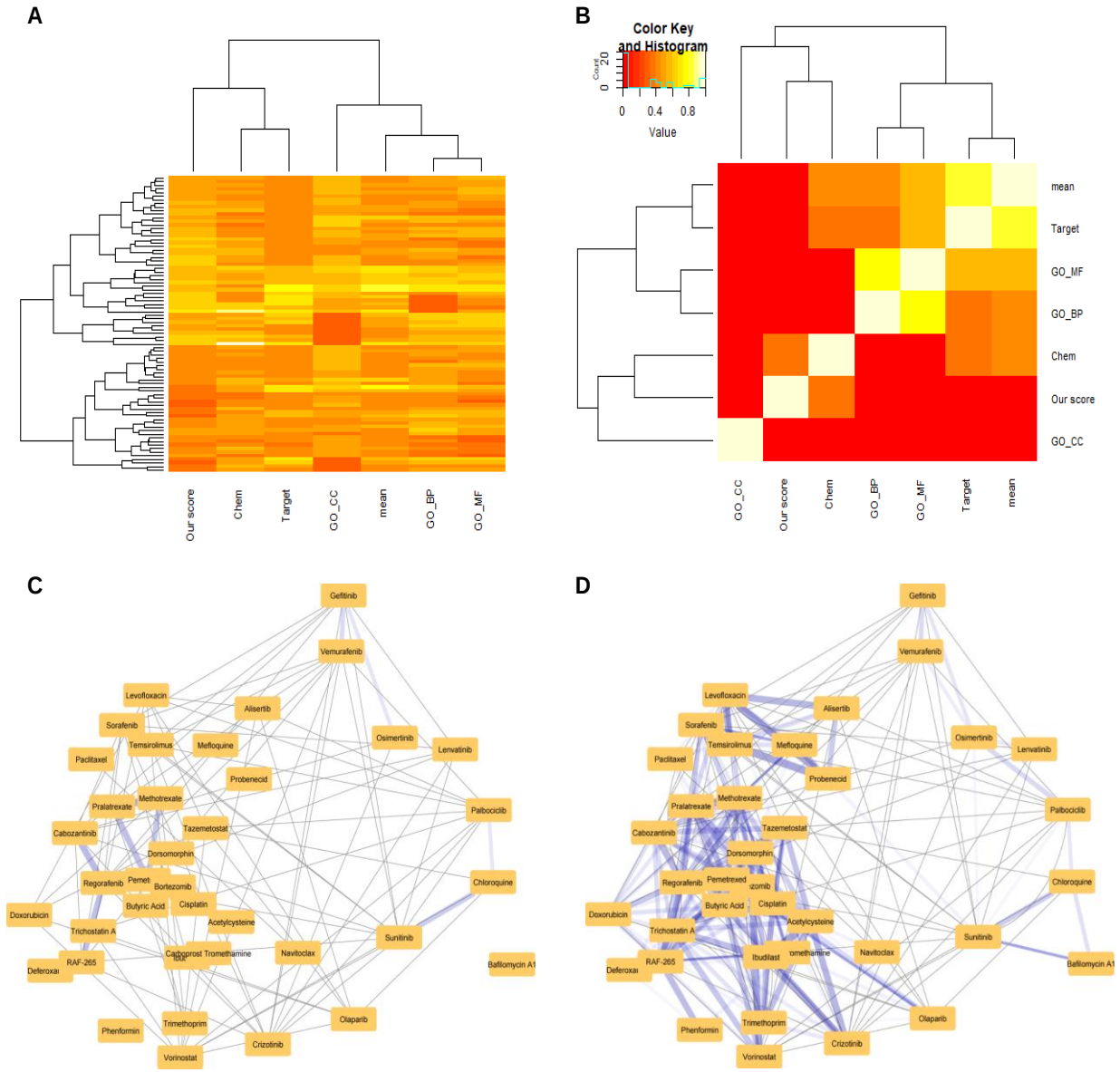


Figure 26. Comparison between the drugs used in our screen with other drug similarity scores. (A) Clustered heatmap of estimated similarity for 82 different drug pairs. The heatmap shows the similarity based on our correlation score and the different DrugSimDB scores (similarity based on Chemical structure, Targets, Gene Ontology Cellular component, Molecular Function and Biological process of induced pathways, mean of all scores). Columns show similarity scores, rows are drug pairs. (B) Correlation matrix of similarity measures based on (A) calculated over the 82 drug pairs. Our similarity correlates best with chemical similarity of drugs (Pearson R=0.35, p=0.001), and similarity based on shared targets (Pearson R= 0.23, p=0.04). (C) Network of 27 drugs in common between DrugSimDB and our own set (converted to DrugBank ID) showing DrugSimDB edges (grey) and edges in common between DrugSimDB and our own network filtered at score > 0.6. (D) Union of the DrugSimDB network and our own filtered at score > 0.6 showing in blue our score (thicker and less transparent lines show higher similarity score).

2.6. Comparison of the barcode profiles to identify the mechanism of action of new anti-cancer drugs

Our results showed that drugs targeting the same pathway display similar barcode patterns (Fig. 27). We reasoned that this type of analysis could then be used to investigate the mechanism of action of new anti-cancer agents. To assess the feasibility of this approach, we selected a new molecule capable of inhibiting the growth of PC9 cells. We tested a panel of 77 new compounds generated by Laura Holzhauer and Anke Deckers at the Karlsruhe Institute of Technology (Germany) on the viability of PC9 cells. To identify a drug that selectively inhibit the growth of NSCLC cells, we assessed in parallel the effects of these drugs in the neuroblastoma cell line SKNAS. The cells were treated for six days in the presence or the absence of the different compounds (5 µM), then fixed and stained with crystal violet. We found that treatment with the drug X-13271 resulted in a profound inhibition of proliferation of PC9 cells, which was less pronounced in SKNAS cells (Fig. 28). Because of its specific effect on PC9 cell growth, we selected this compound and generated a specific barcode signature as described above for the other molecules. We then compared the profile of X-13271 using the database generated from the 87 known drugs tested. A preliminary analysis suggests that X-13271 displays a barcode pattern close to that of the antioxidant compound N-acetylcysteine and to the proteasome inhibitors bortezomib and MG-132 (data not shown). Future experiments will be performed to validate these findings.

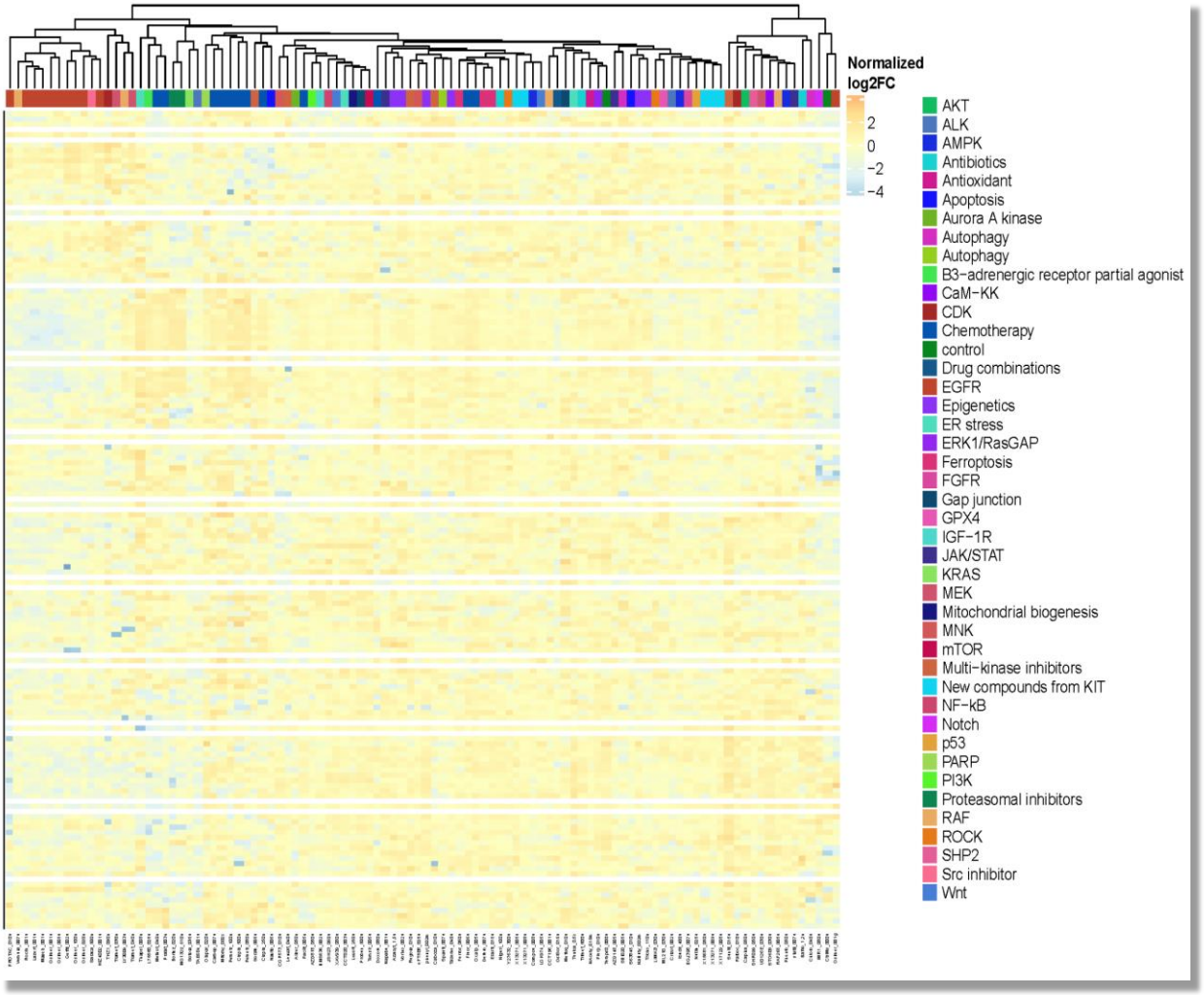


Figure 27. Annotated drug clustering based on barcodes signatures. Heatmap showing that barcode profiles generated from the 87 compounds used in our screen. Drugs targeting the same pathway were clustered together. Each column represents a condition and rows represent barcodes.

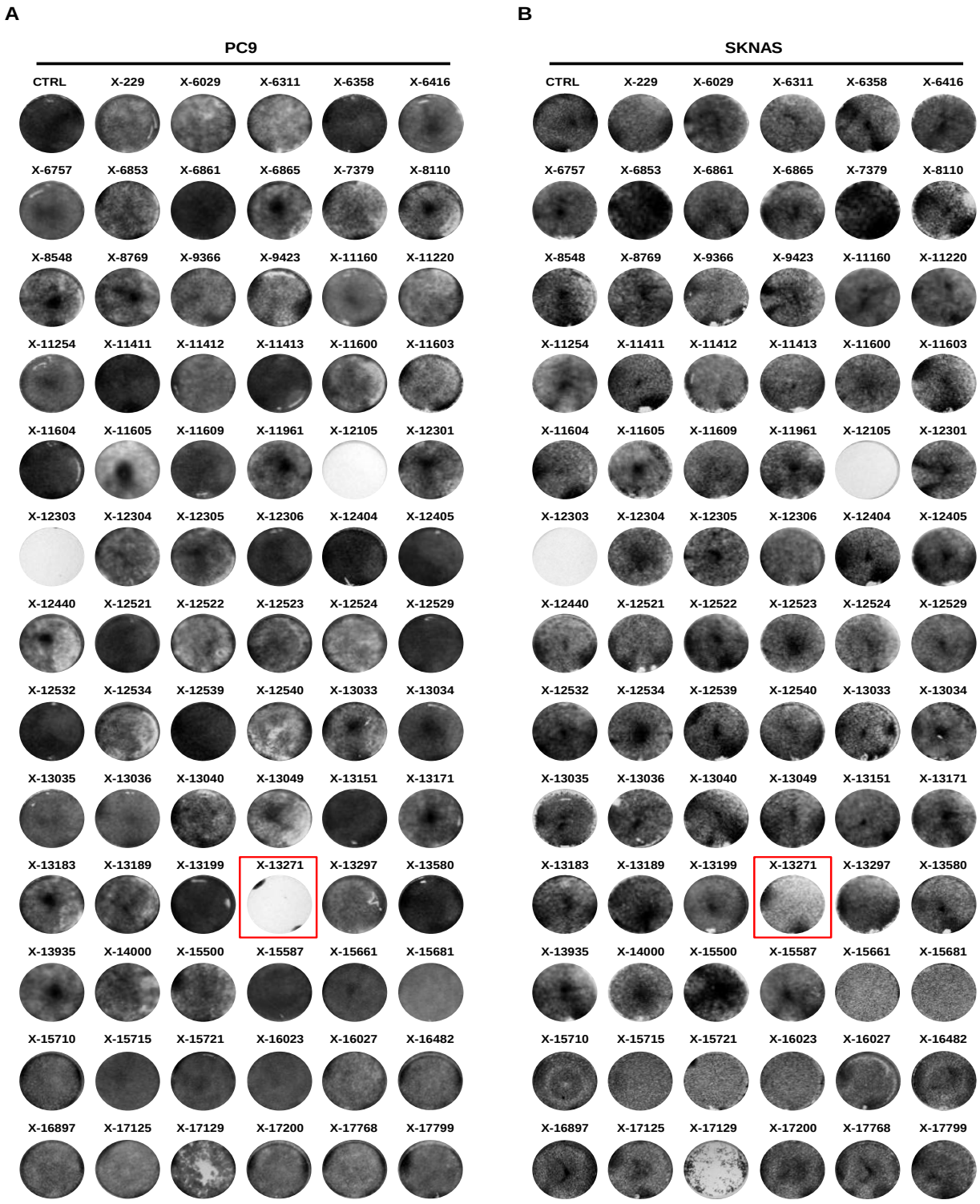


Figure 28. Screen of novel small molecules on the growth of PC9 and SKNAS cells. The PC9 (NSCLC, A) and SKNAS (neuroblastoma, B) cells were treated with a panel of newly synthesized compounds (5 µM) for six days. The colony formation assay showed that the drug X-13271 exerted a stronger effect on cell growth in PC9 cells as compared to SKNAS.

3. Isolation and characterization of individual clonal lineages from a heterogeneous cell population

Cellular barcoding is a powerful tool to track the fate and monitor the clonal dynamics of individual subclones and their progeny, enabling a better understanding of how tumors circumvent treatment. However, this approach remains rather descriptive and it doesn't enable in depth characterization of cell subpopulations of interest to identify the mechanisms responsible for their predetermined response to treatment. To address this limitation, we devised Barcode-Tracker, a novel strategy to isolate clonal populations of cells based on the recognition of a specific genetic barcode. This approach combines CRISPR/Cas9 technology with viral barcoding, and relies on fluorescent reporters driven by barcode-containing promoters that can be recognized by a nuclease-dead form of Cas9 (dCas9). Since dCas9 retains the capacity to form a complex with sgRNAs and specifically interact with the targeted DNA, it can be coupled to a transactivator or a repressor to function as a programmable transcription factor to induce (dCas9-activator) or inhibit (dCas9-repressor) the expression of a DNA sequence of interest (Gilbert et al., 2013). By inserting the barcodes in the promoter of a reporter gene and generating an sgRNA specifically targeting a barcode of interest, the Barcode-Tracker approach should provide the means to identify and isolate for functional analysis rare clones displaying a particular predetermined phenotype (resistant, tolerant or sensitive) directly from the treatment-naïve population.

3.1. Barcode-Tracker workflow

The Barcode-Tracker approach is based on a library of lentiviral vectors in which the barcodes are inserted in proximity of the sequence encoding a fluorescent reporter gene. The library is then transduced at low multiplicity of infection into cells constitutively expressing a dCas9-activator. After the functional screen, the relative abundance of each barcode is analyzed by deep sequencing, enabling the identification of subpopulations behaving in a particular way, *e.g.* highly sensitive or tolerant in the presence of a certain drug. A sgRNA that is specific to the barcode of the clone of interest is then generated and introduced in the original, untreated population of library transduced cells. In complex with the dCas9-activator, the sgRNA will

induce the expression of the fluorescent reporter only in the cells containing the barcode of interest, which can then be isolated by fluorescence-activated cell sorting (FACS) (Fig. 29). Since the cells can be directly isolated from the unselected mass population, their phenotype is not modified by the drug, and it can be analyzed to investigate the mechanism(s) underlying the intrinsic capacity of the cells to behave in a particular manner during the treatment, thus mimicking a therapy-naïve tumor population.

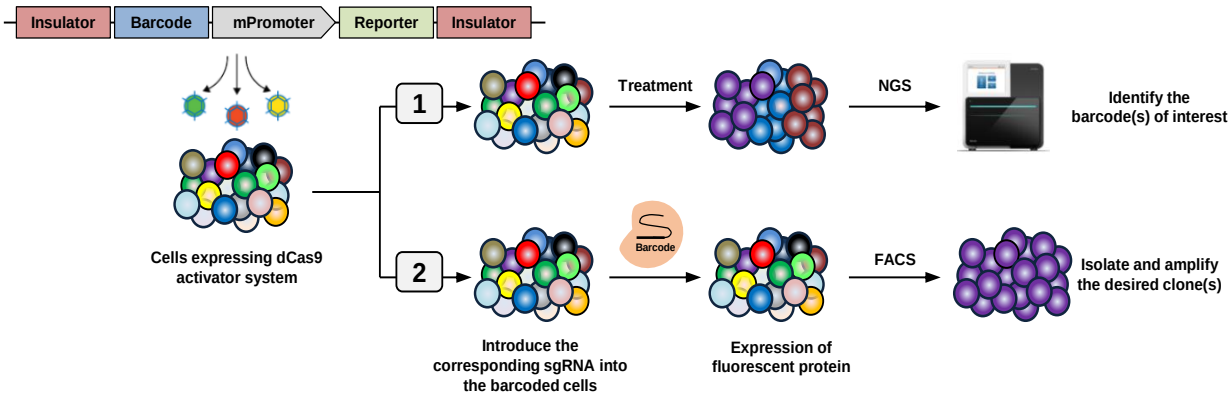


Figure 29. Diagram describing the Barcode-Tracker approach. A heterogeneous population of cancer cells containing a dCas9-activator system and the Barcode-Tracker library are amplified, and a part of the cells is preserved for subsequent lineage isolation. (1) The cells are treated and the barcodes of enriched and depleted clones are identified by NGS. (2) A sgRNA complementary to a barcode of interest is generated and introduced into the untreated mass population of barcoded cells. The sgRNA/dCas9 complex lead to the expression of the fluorescent reporter only in the cells containing the barcode of interest, which can be sorted by FACS.

3.2. Components of the Barcode-Tracker approach

3.2.1. dCas9-activator

Cas9 is an RNA-guided endonuclease that can be directed to a specific genomic region through complementarity between the Cas9-associated sgRNA and the targeted sequence. Cas9 generates a precise double-stranded break (DSB) upstream of the protospacer-adjacent motif (PAM) through a process mediated by its HNH and RuvC catalytic domains (Jinek et al., 2012) (Fig. 30A). To enable RNA guided transcriptional regulation without modifying the genomic sequence, a variant of Cas9 have been generated through mutations of amino acids critical for DNA catalysis within the RuvC and HNH nuclease domains. This catalytically inactive Cas9 (dCas9) has been engineered to transcriptionally induce a genomic sequence of interest by direct

fusion of a transactivator, such as VP64 (tetrameric repeat of the minimal activation domain of herpes simplex virus VP16) to the C terminus of dCas9. When combined with a sgRNA targeting a sequence near the promoter, the dCas9-VP64 can upregulate transcription of the gene of interest (Fig. 30B). Further improvement in transcriptional activation can be achieved by the addition of helper proteins. An example is the SAM (synergistic activation mediator) system developed by the Zhang laboratory and applied in high throughput genome-scale screens (Joung et al., 2017; Konermann et al., 2015). This system is composed of a chimeric sgRNA backbone with synthetic MS2 binding loops that enable the recruitment of two different activation domains, P65 (transactivating subunit of NF-kappa B) and HSF1 (activation domain of human heat-shock factor 1), to a dCas9-VP64 fusion protein (Fig. 30C). We used the SAM based CRISPR-activator technology to robustly activate the transcription of our Barcode-Tracker reporter construct.

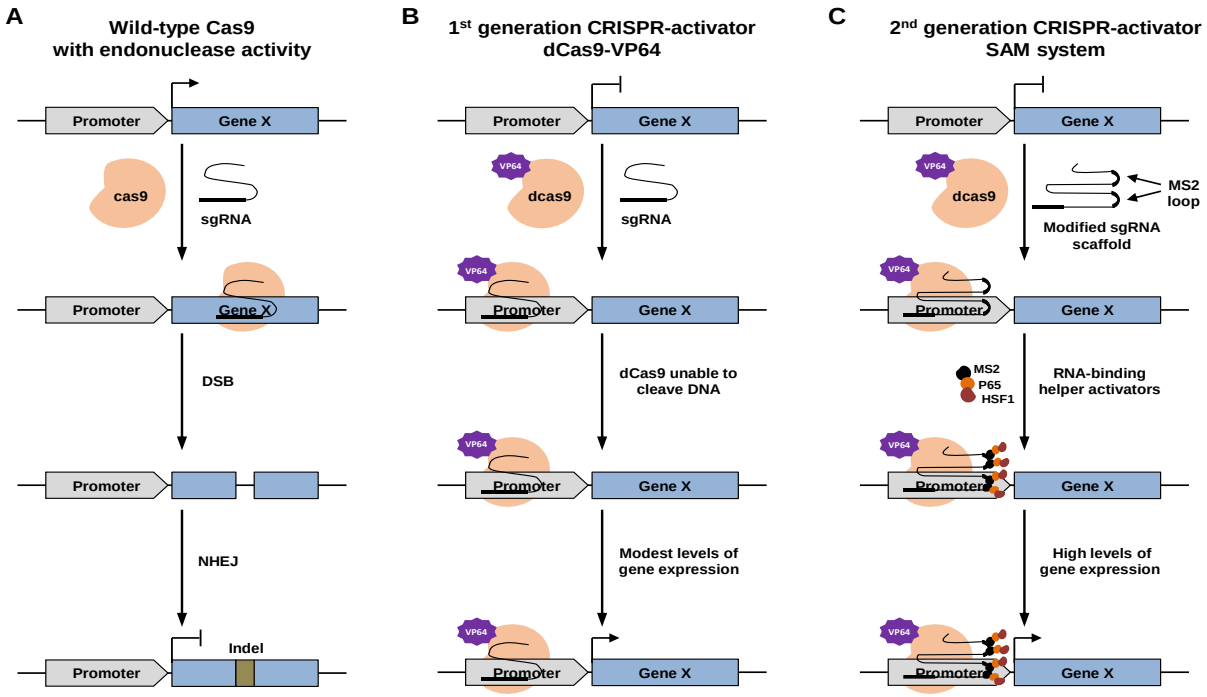


Figure 30. RNA-guided transcriptional activation. (A) Cas9 is directed to a genomic target region by a sgRNA and it functions as a site-specific endonuclease, leading to the formation of double strand breaks (DSB). (B) CRISPR activation is based on the fusion of dCas9 with the transactivator VP64. In combination with a sgRNA, this complex can recruit the transcriptional machinery to a specific region of the genome. This first generation of CRISPR activators exhibits modest levels of gene activation. (C) SAM is based on modified sgRNAs containing aptamers that bind to MS2 proteins, enabling the recruitment of additional transactivators (HS1 and P65). The SAM system provides higher expression levels compared to other CRISPR activators.

3.2.2. Barcode-Tracker lentiviral vector

The Barcode-Tracker construct contains GFP and mCherry coding sequences inserted in opposite direction under the control of two cytomegalovirus minimal promoters (mCMV) and separated by a BbsI site for barcodes insertions. This cassette has been cloned into the VIRHD-EP backbone (Grumolato et al., 2010) containing the puromycin resistance marker. To avoid expression due to enhancers potentially located in the proximity of the integration site of the lentivirus, the cassette is flanked by insulator sequences (Kyrchanova et al., 2008). The vector also contains several unique restriction sites, which can be used to replace any component of the cassette (Fig. 31A). Since the reporter is driven by a basally-inactive minimal promoter (mCMV), the expression of GFP and mCherry occurs only upon the binding of the dCas9-activator to the barcode sequence (Fig. 31B).

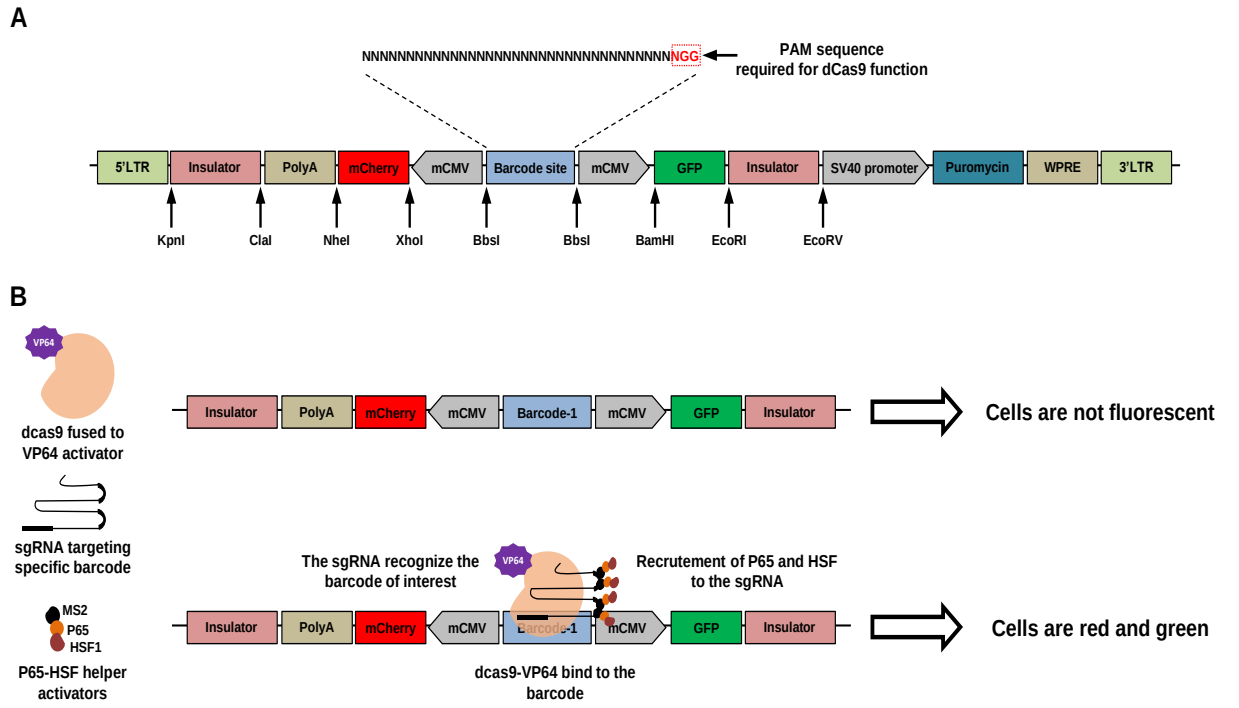


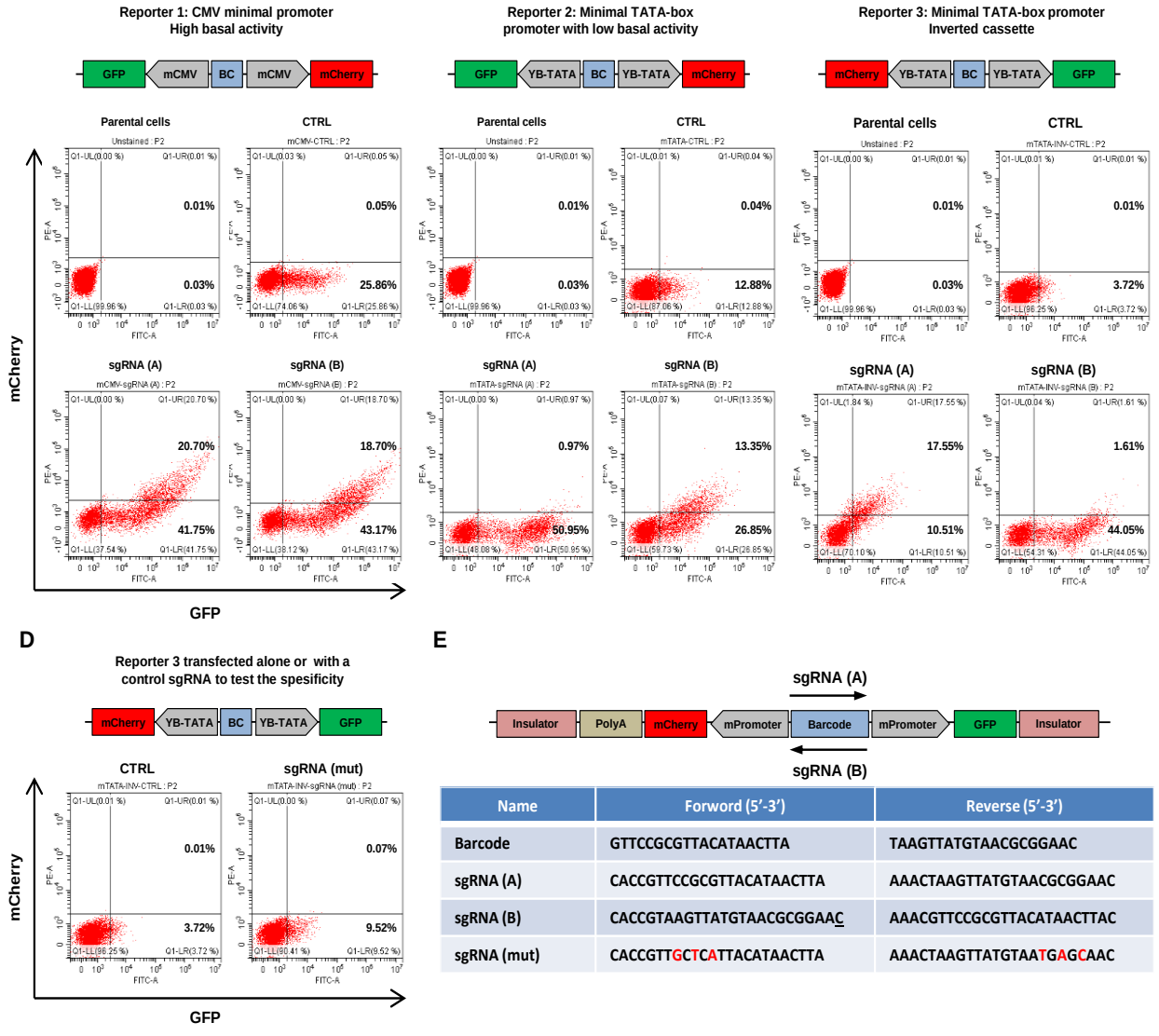
Figure 31. Diagram illustrating the Barcode-Tracker reporter based on the dCas9-SAM system (A) Schematic representation of the Barcode-Tracker lentiviral vector. (B) When co-expressed with a barcode specific sgRNA, the dCas9 activator can drive the expression of GFP and mCherry, enabling isolation of the clonal population of interest. LTR: long terminal repeat; PolyA: polyadenylation signal; mCMV: minimal CMV promoter; WPRE: woodchuck hepatitis virus posttranscriptional regulatory element; N: any nucleotide (A, C, G or T), PAM: protospacer-adjacent motif.

3.3. Optimization of the Barcode-Tracker approach

To test our system, we inserted a published sgRNA target sequence (barcode) from the firefly luciferase gene adjacent to a PAM site. We generated 293T cells stably containing the SAM system. The cells were first transduced with the dCa9-VP64 and selected with blasticidin. These cells were then infected with the MS2-P65-HSF construct carrying the helper activator proteins and the hygromycin selection marker. After transduction, the cells were transiently transfected with the Barcode-Tracker vector alone or in combination with a sgRNA targeting the barcode sequence. As shown in Figure 32A, the expression of both mCherry and GFP can be induced by the sgRNA. However, we noticed that the GFP is also expressed in cells transfected with the reporter alone in the absence of the sgRNA, indicating leakiness of the system (Fig. 32A). It has been reported that the CMV minimal promoter may exert high basal activity. To test this possibility, we generated a second construct by replacing the mCMV with a synthetic YB-TATA minimal promoter (mYB-TATA) with low basal activity (Ede et al., 2016). While this minimal promoter was able to reduce the leakiness, a small fraction of the cells still expressed GFP (12.8%) in the absence of the sgRNA (Fig. 32B). We noticed that GFP, but not mCherry, was expressed in basal conditions, suggesting that the leakiness could be due to the orientation of the fluorescent proteins. To test this hypothesis, we inverted the two reporter genes and we found that the new construct showed reduced leakiness of GFP (3% instead of 12.8% of GFP expressing cells; Fig. 32C). Based on these findings, we decided to use this configuration for further development of the system.

Our data showed that the Barcode-Tracker system can drive reporter expression by a sgRNA-dCas9 activator complex. We then wanted to test the specificity of the system by using a sgRNA containing few mismatches. After transfection of the Barcode-Tracker vector with the control sgRNA, we found a slight increase in the cells expressing the reporter (Fig. 32D). These results suggest that our transcriptional activation-based approach may lack specificity and display a certain degree of leakiness potentially difficult to fully prevent, thus making it difficult to isolate a sufficiently pure cell population of interest. Moreover, while we were in the process of testing different Barcode-Tracker systems, a similar method for specific lineage isolation, named CaTCH (CRISPRa tracing of clones in heterogeneous cell populations), was published in

Nature Biotechnology (Umkehrer et al., 2021). For these reasons, we decided to continue our project by developing a different and potentially more robust strategy, based on a toggle switch, a genetic circuit used in synthetic biology.



Name	Forword (5'-3')	Reverse (5'-3')
Barcode	GTTCCGCGTTACATAACTTA	TAAGTTATGTAACGCGGAAC
sgRNA (A)	CACCGTTCCGCGTTACATAACTTA	AAACTAAGTTATGTAACGCGGAAC
sgRNA (B)	CACCGTAAGTTATGTAACGCGGAAC	AAACGTTCCGCGTTACATAACTTAC
sgRNA (mut)	CACCGTTGCTCATTACATAACTTA	AAACTAAGTTATGTAATGAGCAAC

Figure 32. Optimization of the Barcode-Tracker system. (A-C) 293T cells containing the SAM dCas9 activator system were transiently transfected with 3 different versions of the Barcode-Tracker construct, alone (CTRL) or with a barcode-specific sgRNA (sgRNA-A/B). The levels of mCherry and GFP were measured by FACS 72h after transfection. (D) 293T cells expressing the SAM system were transfected with reporter 3 alone or in combination with a control sgRNA (sgRNA mut), containing three mismatches, to test the specificity of our system. (E) Schematic representation showing the orientation of the sgRNA upon binding to the barcode. FITC-A: GFP; PE-A: mCherry, BC: barcode.

3.4. A toggle switch strategy for the Barcode-Tracker system

The toggle switch (Gardner et al., 2000) is a synthetic biology circuit based on two repressible promoters arranged in a mutually inhibitory network, which results in two alternative steady states of gene expression (OFF and ON) without stable intermediate conditions (Deans et al., 2007; Slusarczyk et al., 2012). As illustrated in Figure 33, this type of circuit relies on the capacity of a gene (gene A in the figure) to turn OFF the activity of another gene (gene B) and *vice versa* (Fig. 33), and it has been used to tightly regulate gene expression in both prokaryotic and eukaryotic cells, including mammalian cells (Kis et al., 2015; Weber and Fussenegger, 2012). For example, Deans et al. developed a mammalian genetic circuit based on the Lac repressor to switch between two stable transgene expression states after transient administration of isopropyl-b-thiogalactopyranoside (Deans et al., 2007). To reduce leakiness and ensure efficient induction of the reporter, in the last part of my thesis we generated a toggle switch to drive the expression of either GFP or mCherry. In our system, the insertion of a barcode specific sgRNA, which binds to a dCas9-repressor protein, results in a change of state, where one fluorescent protein is turned OFF and the other is switched ON specifically in the cells containing the barcode.

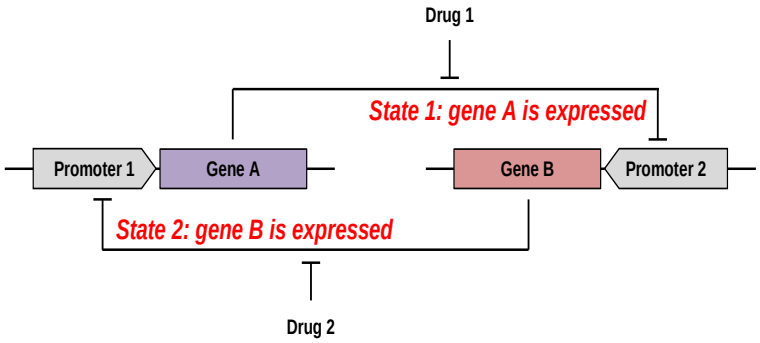


Figure 33. Diagram illustrating the concept of a genetic toggle switch. In the state 1, the gene A is expressed and it stably represses gene B through binding to its promoter. Acute inhibition of the activity of gene A induced by treatment with drug 1 restores the expression of gene B, which can bind to the promoter of gene A and block its expression. This results in a switch to state 2, in which gene B is expressed, while gene A is repressed. In these conditions, the circuit can be turned again to state 1, simply by adding drug 2, which can inhibit the activity of gene B.

3.5. Components of the Barcode-Tracker toggle switch

3.5.1. Tet repressor

The tetracycline (Tet)-inducible system is one of the most widely used strategy to regulate the expression of a transgene. It is based on the Tet repressor (tTR) derived from the Tn10 Tc operon of Escherichia coli that can bind to the Tet operator (TetO) sequence in the absence of tetracycline or one of its derivatives, such as doxycycline (Deuschle et al., 1995). To increase its repressor activity, we fused the tTR to a Krüppel-associated box (KRAB) domain. The resulting protein strongly inhibits the expression of a transgene containing TetO sites in its promoter, unless the cells are treated with tetracycline or doxycycline (Fig. 34A).

3.5.2. Cumate repressor

The cumate switch system has been successfully used for inducible expression in mammalian cells. The cumate repressor (CymR) can be used to block the transcription of a transgene by binding to cumate responsive elements (CuO) placed downstream of the initiation site of a constitutive promoter (Mullick et al., 2006). Addition of cumate provokes a conformational change in the CymR such that it can no longer bind to the CuO sequences, resulting in the expression of the transgene of interest (Fig. 34B).

3.5.3. dCas9-repressor

Similarly to the dCas9-activator strategy, CRISPR/Cas9 has been adapted to inhibit the transcription of a gene of interest by addition of repression domains, such as KRAB, to the dCas9. It has been reported that in some conditions dCas9-KRAB inhibition can be inefficient (Gilbert et al., 2013), so we decided to use a dCas9-KRAB fused to another transcriptional repressor, the methyl CpG binding protein 2 (MeCP2) (Yeo et al., 2018) (Fig. 34C).

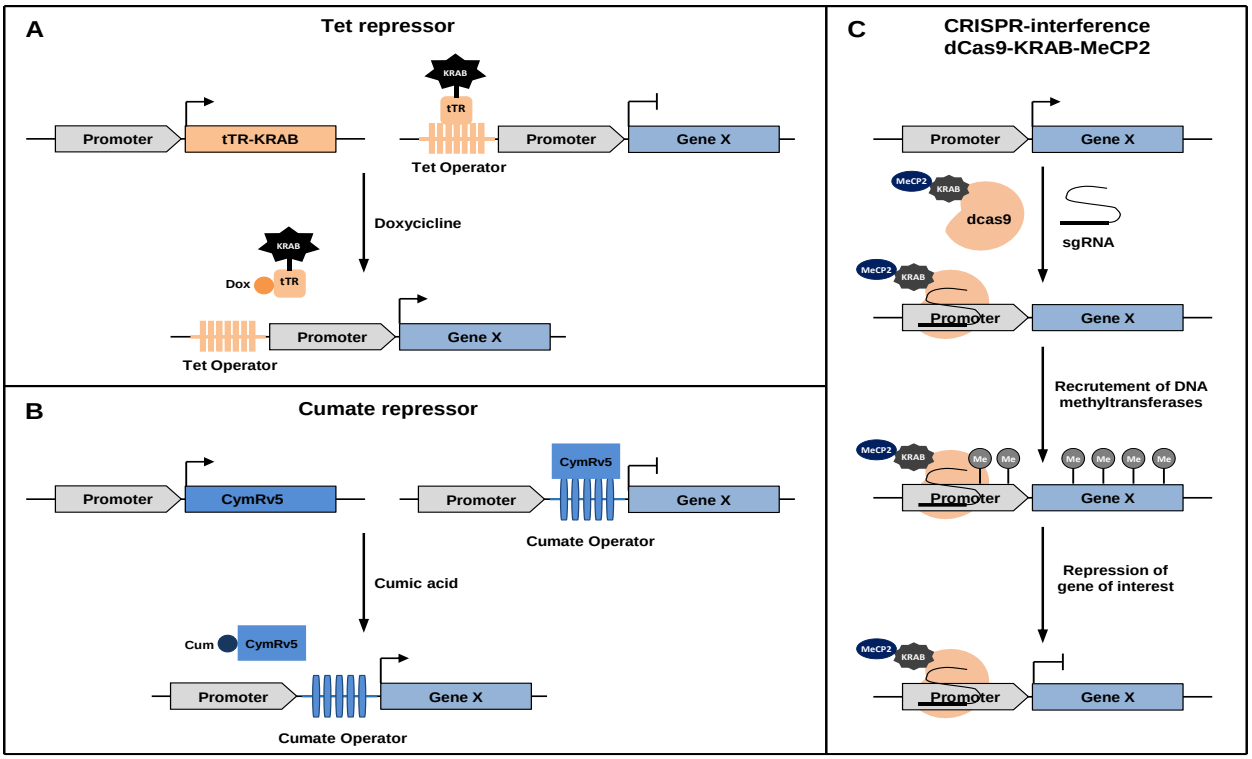


Figure 34. The basic components of our synthetic gene network. (A) In the absence of doxycycline (Dox), tTR-KRAB binds to Tet operator and suppresses the expression of the gene of interest. (B) The CymRv5 can bind to the cumate operator (CuO) and blocks transcription from the promoter. In the presence of cumate (Cum), CymRv5 does not bind to the CuO elements, thus allowing transgene expression. (C) CRISPR/dCas9 repression system. dCas9 fused to KRAB and MeCP2 repressors binds to a specific DNA sequence guided by a sgRNA, leading to transcriptional suppression of the targeted gene.

3.6. Overview of the Barcode-Tracker toggle switch constructs

Our system is based on two repressor cassettes (Fig. 35):

-Repressor 1 (R1): the barcode is located directly upstream of a CMV promoter containing seven repeats of the TetO sequence, driving the expression of mCherry, puromycin and cumate repressor (CymRv5), all separated by self-cleaving 2A peptides (P2A and T2A). The 2A peptides can induce ribosomal skipping during translation, allowing simultaneous expression of different transgenes from the same promoter. Downstream of CymRv5, we also included a GFP-targeting shRNA based on a mir-30 backbone sequence. The cassette is flanked by anti-repressor elements (ARE) (Kwaks et al., 2003).

-Repressor 2 (R2): a CMV promoter containing five CuO repeats drives the expression of GFP, zeomycin and tTR-KRAB, separated by self-cleaving 2A peptides.

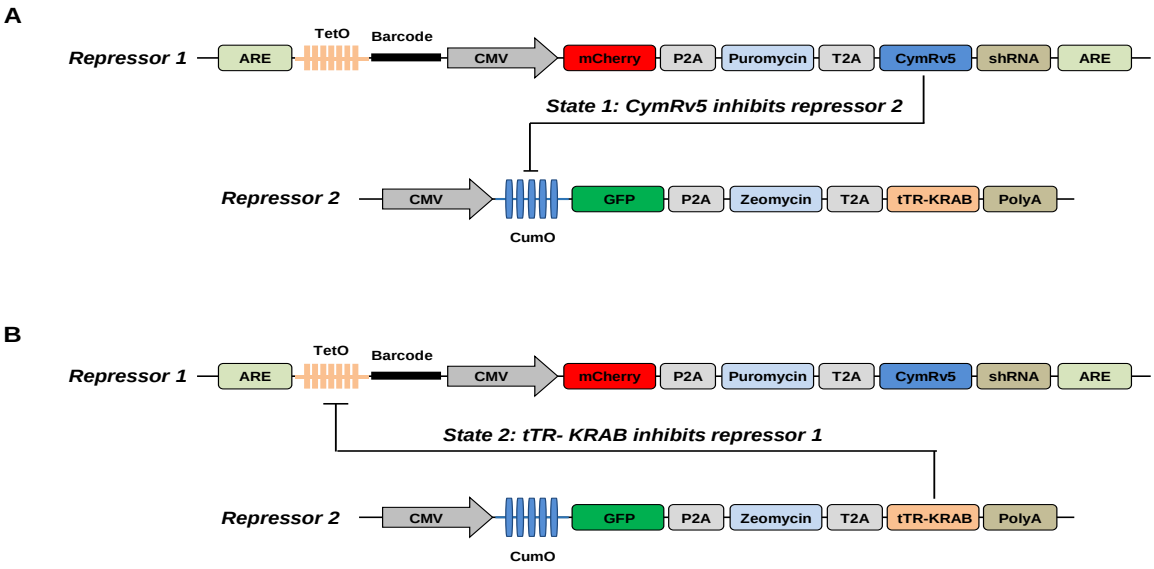


Figure 35. Schematic diagrams of the toggle switch constructs. (A) In the state 1, mCherry and CymRv5 are expressed, and CymRv5 binds to the CuO sites in Repressor 2. This causes transcriptional repression of GFP and tTR-KRAB. (B) In the state 2, GFP and tTR-KRAB are expressed, and tTR-KRAB binds to TetO sequences on Repressor 1, preventing the expression of mCherry and CymRv5.

3.7. Workflow of the Barcode-Tracker toggle switch

As shown in Figure 36A, in cells expressing GFP and tTR-KRAB, the R1 cassette is silenced (green cells). Conversely, expression of CymRv5/mCherry/shRNA results in the inhibition of the R2 cassette (red cells; Fig.36B). Since one cassette prevails on the other in a random manner, the mass population should be composed of a mix of red and green cells. The inhibitory activity of tTR-KRAB and CymRv5 can be controlled by administration or withdrawal of doxycycline and cumic acid, respectively, thus preventing permanent (irreversible) inhibition of our reporters. Adding both doxycycline and cumate to the media will prevent the binding of the tTR-KRAB and CymRv5 to their response elements, resulting in the expression of GFP and mCherry at the same time (Fig. 36C). Thanks to these inducible systems, R1 and R2 can be switched from an OFF to an ON state and *vice versa*. Before isolation of a specific clonal lineage, all cells will be turned red by adding doxycycline to the media (Fig. 36D). To isolate cells containing a barcode of interest, the corresponding specific sgRNA will be introduced in the cells, where it will form a complex with the dCas9- KRAB-MeCP2 repressor, thus blocking the expression of the R1 cassette (Fig.36E). As a result, the clones containing the barcode will turn green (Fig.36F), thus enabling FACS sorting of the subpopulation of interest.

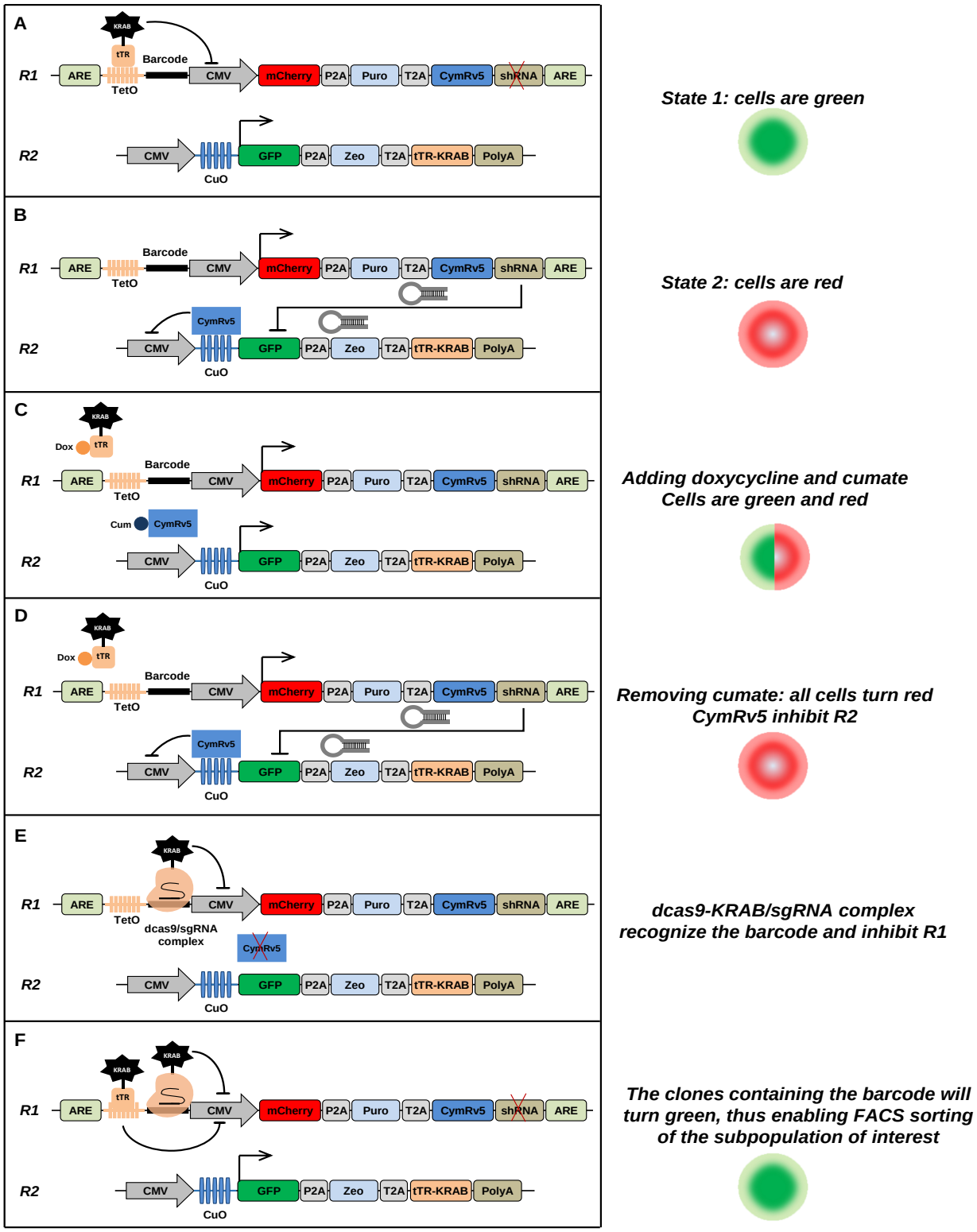


Figure 36. Schematic representation of the Barcode-Tracker-toggle switch strategy. See the text for details.

3.8. Optimization of the Barcode-Tracker toggle switch approach

To test the toggle switch system, we generated PC9 cells expressing R1, R2 and dCas9-KRAB-MeCP2 (Fig. 37). To prevent differential expression levels of our reporter caused by random integration of lentivirus, we inserted R2 in the AAVS1 locus using CRISPR/Cas9-mediated recombination. Because of the relative large size of the R2 cassette, the knock-in was performed in two steps. First, we designed a targeting vector containing a CMV promoter, five CuO elements, GFP and a polyA sequence, flanked by left and right homologous arms to integrate the cassette into the AAVS1 locus through homologous recombination. PC9 cells where co-transfected with the targeting vector and plasmids encoding Cas9 and two distinct sgRNAs recognizing two adjacent sequences of the AAVS1 locus. After transfection, we maintained the cells in culture for around 4 weeks to dilute out the targeting plasmid not integrated into the genome and present under an episomal form. The GFP positive cells were then sorted (Fig. 37A), amplified, and subjected to a second round of integration to introduce the other components of R2. As shown in Figure 37B, a second donor DNA vector containing two arms homologous to GFP and polyA signal was used to introduce the zeomycin resistant gene and the tTR-KRAB, separated by self-cleaving peptides into the AAVS1 site. The edited cells were then enriched by adding media containing zeomycin every 3 days. The knocked-in cells were transduced with a lentiviral vector encoding the dCas9-KRAB-MeCP2 repressor and selected with blasticidin (Fig. 37B). The cells were then transduced at low MOI with another lentiviral vector containing the R1 cassette, including the barcode, and selected with puromycin. For a preliminary assessment of the system, we used a R1 vector containing only one barcode (for the actual experiment, the cells would be transduced at low MOI with a barcode library generated in the R1 vector to differently label individual cells). We then treated the cells for 72 h in the presence or the absence of 1 µg/ml doxycycline, 30 µg/ml cumate or a combination of both (Fig. 37C). After treatment, we collected the cells and the fluorescent levels of GFP and mCherry were assessed by FACS. As shown in Figure 37D, we found that 84% of the cells expressed both GFP and mCherry in the presence of doxycycline and cumic acid, while a small fraction of the cells (15%) expressed only GFP. In the presence of cumate alone, around 50% of the cells expressed only GFP, while in the presence of doxycycline alone 22% of the cells expressed only mCherry. (Fig. 37D).

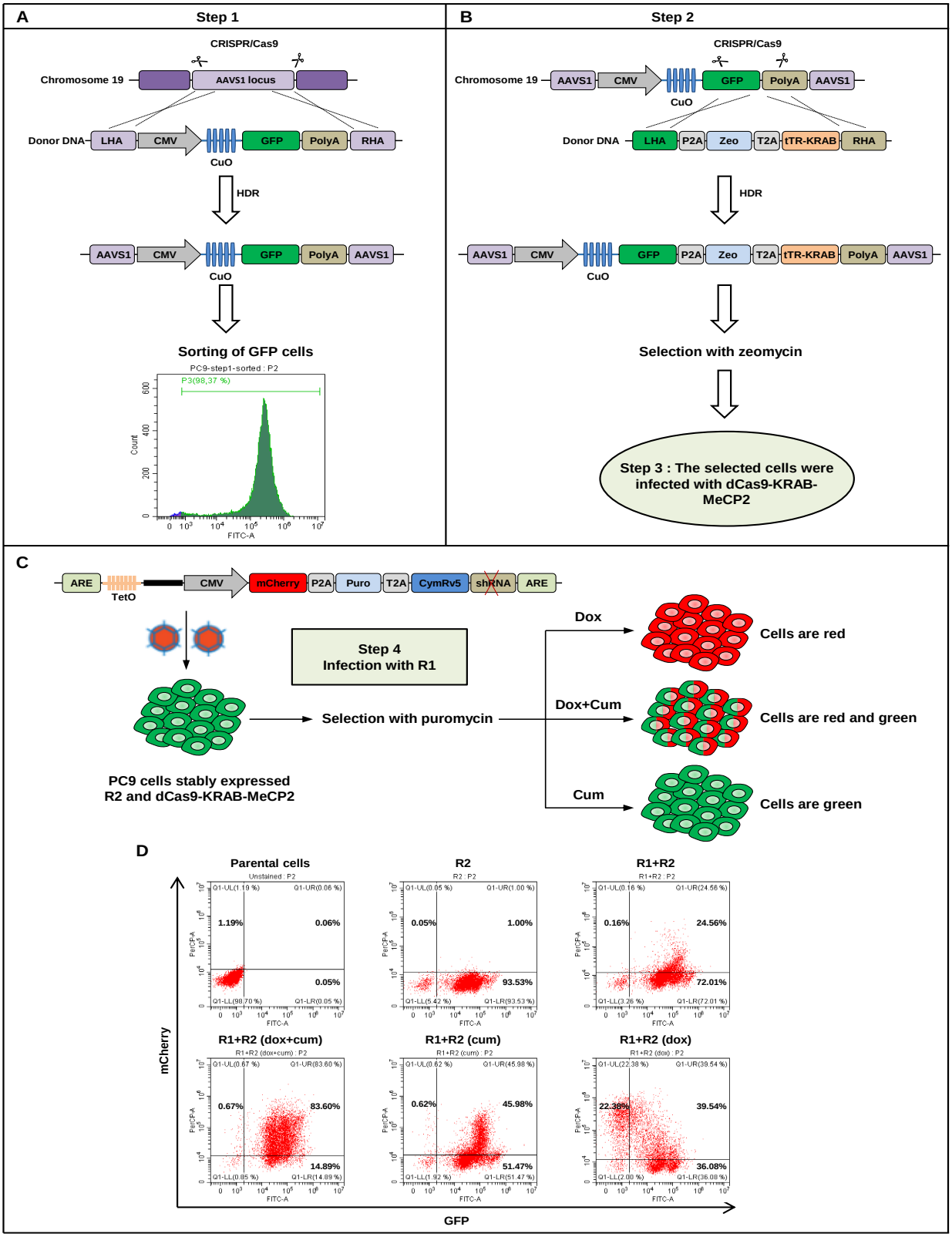

A
Step 1
CRISPR/Cas9
Chromosome 19
AAVS1 locus
Donor DNA
LHA
CMV
GFP
PolyA
RHA
CuO
HDR
AAVS1
Sorting of GFP cells
PC9-step1-sorted : P2
P3(98,37 %)
Count
FITC-A
B
Step 2
CRISPR/Cas9
Chromosome 19
AAVS1
CMV
GFP
PolyA
AAVS1
CuO
Donor DNA
LHA
P2A
Zeo
T2A
tTR-KRAB
RHA
HDR
Selection with zeomycin
Step 3 : The selected cells were infected with dCas9-KRAB-MeCP2
C
ARE
TetO
CMV
mCherry
P2A
Puro
T2A
CymRv5
shRNA
ARE
Step 4
Infection with R1
PC9 cells stably expressed
R2 and dCas9-KRAB-MeCP2
Selection with puromycin
Dox
Dox+Cum
Cum
Cells are red
Cells are red and green
Cells are green
D
Parental cells
Unstained : P2
R2
R2 : P2
R1+R2
R1+R2 : P2
1.19%
0.06%
0.05%
0.05%
1.00%
93.53%
0.16%
24.56%
72.01%
R1+R2 (dox+cum)
R1+R2 (cum)
R1+R2 (dox)
0.67%
83.60%
14.89%
0.62%
45.98%
51.47%
22.38%
39.54%
36.08%
mCherry
GFP
FITC-A
PerCP-A

Figure 37. Generation PC9 cells expressing the toggle switch components. (A-B) knock in of R2 into the AAVS1 locus using CRISPR/Cas9. (C-D) The cells were infected with R1 and the fluorescent levels of mCherry and GFP were assessed by FACS after 72h of doxycycline or cumate treatment.

These data indicate that our toggle switch circuit needs further optimization. One possibility is that three days of treatment are not sufficient to shift all the cells to the desired state. If longer treatments will yield similar results, we will consider the possibility of purifying the cells through sequential steps of induction. For example, after treatment with doxycycline and cumate, we can isolate double positive cells by FACS. From this population, $mCherry^{+}/GFP^{-}$ cells could then be sorted in the presence of doxycycline alone upon withdrawal of cumate. Finally, these cells could be grown in the presence of cumate alone for a last step of purification to select cells capable of shifting to a $mCherry^{-}/GFP^{+}$ state.

Discussion

1. Elucidate the global effects of anticancer treatment on the clonal architecture of a mass population of NSCLC cells

The emergence of resistance is a fundamental cancer property, which mostly derives from the fact that individual tumors are composed of an intricate pattern of heterogeneous subclonal populations, functioning as a complex reservoir that fuels the capacity of tumor cells to adapt to environmental conditions. Although most EGFR-mutant NSCLC patients initially respond to EGFR inhibitors, the vast majority of these tumors ultimately become resistant due to the emergence of small subpopulations of resistant or tolerant cells. In this study, we investigated the global effects of anticancer therapies on the clonal evolution of a mass population of NSCLC cells, as well as the mechanism(s) responsible for the heterogeneous drug sensitivity of the different cells. Using highly complex genetic barcodes to label individual clones, we showed that some subpopulations display a specific and predetermined response to treatment. These data indicate the certain cells that are primed to behave as tolerant or highly sensitive in response to a given therapy.

The origin of drug tolerance in cancer is debated. Some studies suggest that a subpopulation of cancer cells pre-exist in the treatment-naïve tumors, and are selectively enriched upon drug exposure. For example, Kurppa and colleagues reported that certain PC9 cell subpopulations are more likely to become tolerant to osimertinib treatment (Kurppa et al., 2020). Moreover, Hangaeur et al. found that pretreatment of HER2-positive breast cancer cell lines with a glutathione peroxidase 4 (GPX4) inhibitor reduces the number of persister cells that remain after treatment with lapatinib, suggesting that cells expressing high level of GPX4 pre-exist in the original mass population and are less sensitive to HER2 inhibition (Hangauer et al., 2017). In a recent preprint, Goyal et al. showed that distinct subpopulation of melanoma cells can adopt different transcriptional and functional profiles in the presence of BRAF inhibitors in a way that is probably predetermined before the onset of treatment (Goyal et al., 2021). These findings are consistent with previous reports indicating that persister cells may pre-exist before drug exposure (Raha et al., 2014; Shaffer et al., 2017). On the other hand, several studies suggest that DTP cells arise *de novo* during treatment. In support of this type of mechanism, Rambow and colleagues found that lineage markers characteristic of DTP cells in melanoma were not present in the

parental population, but their expression was induced by treatment with BRAF inhibitors (Rambow et al., 2018). Moreover, recent studies on colorectal and breast cancer argue that most tumor cells are equipotent in their capacity to enter and exit a drug tolerance state in response to chemotherapy. Both studies showed that these cells emerge stochastically by adopting diapause-like state (Dhimolea et al., 2021; Rehman et al., 2021). However, it is important to note that the two frameworks are not mutually exclusive (Shaffer et al., 2017). Our data support a model where both mechanisms can co-exist. In fact, our barcoding experiments showed that a subset of PC9 cells was consistently selected across multiple replicates, reflecting a predetermined response to the treatment, while other clones were enriched in a more stochastic manner. Also, the proportion of cells arising through one or the other of these two mechanisms probably depends on the compound used. Indeed, we noticed that the correlation between replicates was higher for targeted drugs, such as osimertinib, compared to chemotherapeutic agents (Fig. 38).

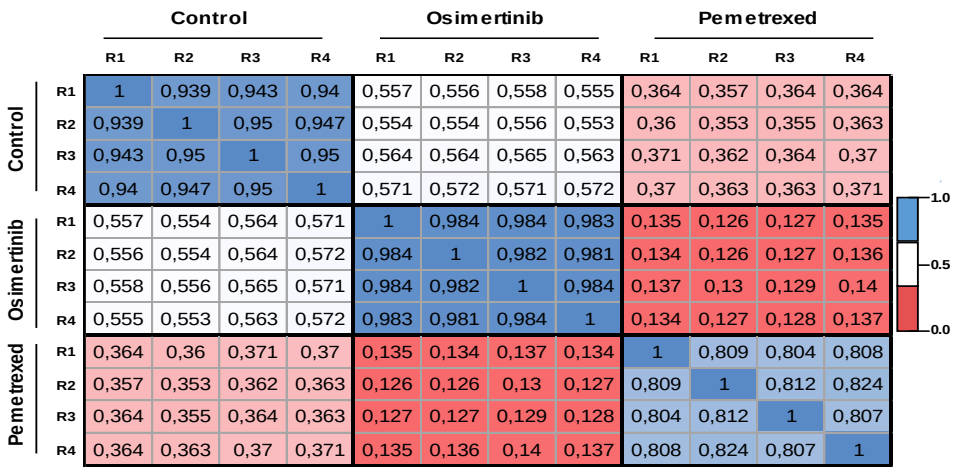


Figure 38. Distinct response to osimertinib and pemetrexed. PC9 barcoded cells were treated for 9 days with or without 100 nM osimertinib or100 nM pemetrexed. Spearman's pairwise correlation analysis of positively selected barcodes was generated.

The relationship between persistence and cancer stem cells (CSCs), and their contribution to disease recurrence and therapeutic resistance have long been debated. While studies have reported that DTP cells may display markers associated with CSCs, such as high level of aldehyde dehydrogenases (ALDH) (Raha et al., 2014), a major unanswered question is whether DTP cells bear properties that have been attributed to CSCs. It has been shown that CSCs possess an intrinsic resistance to chemotherapy due to high levels of drug efflux pumps, which can prevent accumulation of cytotoxic agents inside the cells, resulting in reduced sensitivity to a

variety of drugs (Segerman et al., 2016). In our study, we showed that tolerance to different treatment arises from multiple distinct subpopulations. These observations imply that NSCLC putative CSCs would not be generically tolerant to most treatments, but they should instead display a specific pattern of drug sensitivity. Of note, we performed a mouse experiment in which barcoded PC9 cells were subcutaneously grafted in immunodeficient mice to form tumors. We showed that the barcode complexity was decreased in the tumors, although the correlation between replicates was lower in tumors as compared to cells grown in vitro (Fig. 39 and data not shown), implying that the tumorigenic potential is not restricted to a small subset of the cell population. While preliminary, this experiment also showed higher correlation between tumors derived from mice treated with osimertinib (Fig. 39), implying that EGFR-TKIs can induce a specific shift in the clonal architecture of a NSCLC population not only *in vitro*, but also *in vivo*. To address the relationship between CSCs and drug sensitivity, future experiments could be focused on the isolation of barcoded PC9 cells that display CSC characteristics (*e.g.* ALDH^{high} vs ALDH^{low}) and comparison of their capacity to form tumors in serial xenotransplantation experiments, and/or testing their contribution in tumor relapse after therapy.

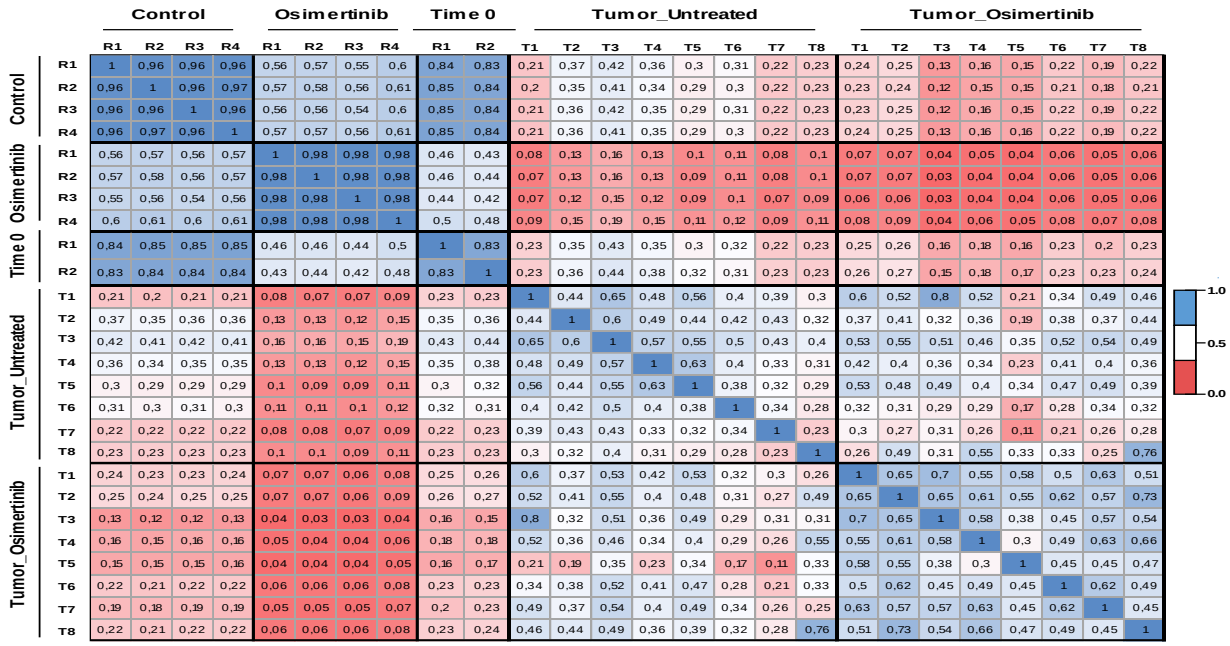


Figure 39. PC9 barcoded cells transplanted into immunodeficient mice. 5 x 10^5 PC9 cells were subcutaneously injected in the right and left flanks of SCID mice. Once the tumors reached a mean volume of about 100 mm^3, the mice were randomized and treated with either vehicle or osimertinib (1 mg/kg) for 4 weeks. The tumors were dissected and gDNA was extracted for barcode analysis by NGS. Spearman's pairwise correlation analysis of positively selected barcodes was generated.

Our findings indicate that distinct clones arise in response to different treatments. This observation could have important clinical implications in predicting effective drug combinations. In EGFR-mutant NSCLCs, reactivation of ERK1/2 has been identified as a resistance mechanism to EGFR-TKIs. Trickler et al. showed that resistance could be prevented by upfront combination with EGFR and MEK inhibitors, in both *in vitro* and *in vivo* models (Ercan et al., 2012; Tricker et al., 2015). Phase Ib/II clinical trials (NCT02143466/NCT03392246) are ongoing to evaluate the efficacy of this combination in patients with advanced EGFR-mutant NSCLC (Oxnard et al., 2020; Ramalingam et al., 2019). However, preclinical studies showed that a small fraction of NSCLC cells can adapt and survive initial treatment with this combination (Kurppa et al., 2020). In our experiments, several osimertinib tolerant clones were also selected in response to trametinib, implying that this combination could probably favor the selection of the same cell subpopulations (Fig. 40). These observations suggest that other combinatorial strategies could be designed in which associations of compounds targeting completely different pathways may be more effective in preventing the emergence of DTP cells.

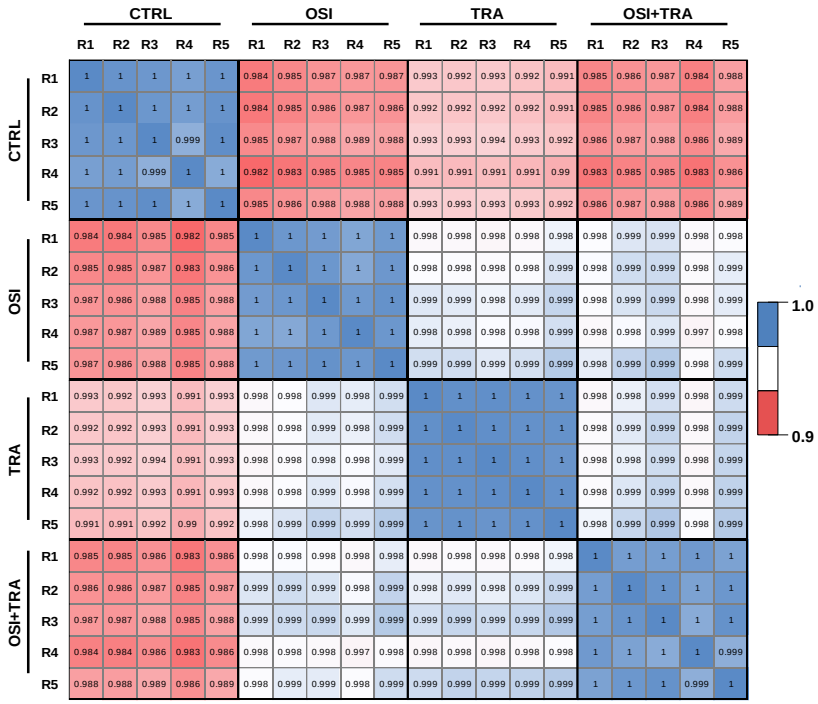


Figure 40. Predefined response to dual EGFR/MEK inhibition. PC9 barcoded cells were treated for two weeks with or without 100 nM osimertinib (OSI), 30 nM trametinib (TRA) or 100 nM osimertinib + 30 nM trametinib (OSI+TRA). The barcodes amplified from gDNA and analyzed by NGS. Spearman's pairwise correlation analysis of positively selected barcodes is represented. For this experiment, the cells were labeled using CRISPR-barcoding approach.

While successful with certain tumors, such as in the case of BRAF and MEK inhibitors in melanoma, associations of compounds that target the same pathway can also cause higher toxicity and lower therapeutic indexes (Karoulia et al., 2017). On the other hand, targeting parallel pathways with horizontal drug combinations holds promise to maximize drug efficacy. Our data obtained with CRISPR-barcoding and the lentiviral library indicate that the cell subpopulations intrinsically tolerant to osimertinib or pemetrexed were completely different, suggesting that a combination of these two drugs may prevent or delay the emergence of resistance in NSCLC. These effects may explain, at least in part, the effectiveness of the combination of gefitinib with chemotherapy in patients with EGFR-mutant NSCLCs (Noronha et al., 2020). A phase III clinical trial (FLAURA2) is ongoing to evaluate the effects of adding platinum and pemetrexed to osimertinib in the first-line setting (NCT04035486). The first results published from this trial demonstrated manageable safety and tolerability for this combination (Planchard et al., 2021; White et al., 2021). These observations suggest that a better understanding of the heterogeneity in the drug response of individual cells could be important to design more effective and durable therapies.

2. Identify the mode of action of new anti-cancer agent

In this study we used a lentiviral barcode library to investigate the variability in drug response of individual cancer cells exposed to anti-cancer therapies. We found that some subpopulations of PC9 NSCLC cells display a specific and predetermined response to treatment. We extended these findings and showed that each drug exerts a characteristic effect on the clonal architecture of the cell population, resulting in a specific barcode pattern. We reasoned that this functional and predetermined heterogeneity within the same population of cancer cells could be used as a specific signature to compare different compounds and investigate their mechanism of action. As a proof-of-concept of this new strategy, we have generated a database of barcode profiles from 87 known drugs targeting various pathways and used it to predict the mechanism of action of a new compound, X-13271.

One of the earliest approaches for high-throughput profiling of drug response involved the National Cancer Institute panel of 60 human cancer cell lines (NCI-60). Since its introduction in the early 1990s, the NCI-60 has been used to screen thousands of chemical compounds ranging

from natural products to FDA-approved drugs (Shoemaker, 2006). The screen is based on treating 60 different cell lines derived from various types of tumors with the compounds of interest for 48h, followed by measuring their viability using sulphorhodamine B. Data analysis tools, such as the COMPARE algorithm, can then be used to classify the drugs based on their patterns of growth inhibition (Holbeck et al., 2010). However, testing each cell line individually against thousands of drugs is time consuming and expensive. As an alternative approach to the NCI-60, the Golub's laboratory developed PRISM (profiling relative inhibition simultaneously in mixtures), a system in which different cell lines labeled with a genetic barcode are mixed in a single pool, which is then treated with the compounds. By measuring the relative proportion of the different barcodes, this approach is used to identify the cell lines that are more sensitive to the drugs tested (Yu et al., 2016). In both NCI-60 and PRISM, the effects of the compounds are compared on the viability of different cell lines, each derived from distinct individuals, tissues and tumor types. As a consequence, their response to treatment can be influenced by a wide range of factors that can be extremely hard to predict. On the contrary, our strategy takes advantage of the intrinsic heterogeneity within the same population of cancer cells sharing an identical (or almost identical) genotype and a very similar phenotype.

Our results suggest that the response of different cell subpopulations, as well as the barcode profiles is stable over time and it is not affected by freeze-thaw cycles. These data imply that the barcode profiles of the different compounds don't need to derive from the same experiment, but they can be successively added to the dataset. As a consequence, the dataset remains open and it can be progressively enriched over time, which should further improve the accuracy of the analysis. As a proof-of-concept of this approach, we applied it to investigate the mechanisms of action of a new molecule, X-13271, that was shown to exert a stronger growth inhibition in NSCLC cells, as compared to neuroblastoma cells. Beside the analysis of the effects of new anti-cancer drugs, this strategy method can also be used to assess the specificity of known compounds through comparison of their barcode profile with those of other molecules targeting the same pathway. For example, in our study we screened the cells against a panel of different EGFR inhibitors and chemotherapeutic agents, including alkylating, mitotic and antifolate drug, and we found that each class of these compounds displayed specific barcode profiles.

An obvious drawback of our strategy is that only drugs that inhibit the growth of PC9 cells can be profiled. Indeed, some drugs, such as in the case of targeted therapy, elicit responses only in cell lines harboring particular oncogenic mutations or expressing certain genes, thus making it difficult to investigate their mode of action using other cells. To overcome this limitation, it would be conceivable to apply a similar workflow to a small panel of cell lines, each representative of a tumor type of interest, which would allow choosing the appropriate model according to the class of drugs tested.

The method presented here is relatively fast and readily applicable, especially once a dataset of known compounds is established. To our knowledge, this is the first example of a strategy harnessing intratumor heterogeneity to profile drug response. By measuring the effects across different cell subpopulations, this approach could provide a powerful tool to investigate the mechanism of action of new compounds.

3. Isolate and characterize specific cell subpopulations displaying an intrinsic low or high sensitivity to treatment

In this thesis, we demonstrated that certain subpopulations of EGFR-mutant NSCLC cells show a predetermined low or high sensitivity to EGFR-TKIs and cytotoxic therapies, which could have important implications in the acquisition resistance to treatment. The approach that different groups have used to investigate drug tolerance relies on the selection of NSCLC cells with a particular drug, followed by characterization of the remaining population of DTP cells (Hata et al., 2016; Ramirez et al., 2016; Sharma et al., 2010). Though clinically relevant, these findings derive from selected cells, and they are not necessarily transposable to treatment-naïve cells. Moreover, although most of the studies are focused on drug tolerant/resistant cells, a better understanding on drug response can also derive from the characterization of highly sensitive cells, which, by definition, cannot be obtained through selection. To overcome these limitations, we devised Barcode-Tracker, a method that combines cellular barcoding with CRISPR/Cas9 technologies to isolate specific clones of interest from a heterogeneous cell population.

The laboratory of A. Raj has recently proposed a method, named Rewind, to capture the cells of interest directly from the original, unselected mass population. However, this approach,

which relies on RNA fluorescence *in situ* hybridization (RNA-FISH) to identify specific barcodes, does not enable to isolate living clones for their characterization and the fixed cells can only be analyzed by RNA sequencing or imaging-based assays (Emert et al., 2021). To identify specific subpopulations of living cells, we initially developed a system based on fluorescent reporters driven by barcode-containing promoters, which could be recognized by a dCas9 activator in complex with a sgRNA complementary to the barcode of interest. While this complex could induce the expression of GFP and mCherry, in our hands this approach also displayed a certain degree of leakiness that we couldn't prevent, so we decided to change our strategy and use a toggle switch circuit. While we were testing this new system, a paper from Anna Obenauf laboratory was published describing a method, named CaTCH, to isolate specific subclonal populations of cancer cells using a strategy very similar to our Barcode-Tracker activator system (Umkehrer et al., 2021). The authors used CaTCH to study the resistance mechanisms in response to dual BRAF and MEK inhibition in melanoma *in vivo*. The mouse melanoma cell line YUMM-1.7 (BRAF-V600E), previously transduced with the components of the CaTCH system, was subcutaneously injected in the flank of syngeneic C57BL/6 mice. After treatment with either vehicle or BRAF/MEK inhibitors, the tumors were dissociated and the cells grown in culture. From the two cells populations, derived from naïve or resistant tumors, the authors then isolated and characterized clones containing the same barcode. They found that the clone BC2646 displayed low sensitivity to BRAF/MEK inhibition even when isolated from naïve tumors. Mechanistically, the authors showed that, even before drug exposure, the clone BC2646 displayed higher activation of EGFR and MET as compared to the parental cells (Umkehrer et al., 2021). In another study, the same group showed that melanomas have reduced response to immunotherapy once they have developed resistance to targeted therapy. Using CaTCH, the authors isolated the same clonal lineages derived from naïve or resistant tumors and they showed that treatment with BRAF/MEK inhibitors can induce cross-resistance to immunotherapy (Haas et al., 2021). Despite the similarity between CaTCH and the Barcode-Tracker activator constructs, this strategy did not work well for us. Since the CaTCH is available through Addgene, future experiments could be focused on comparing the two systems using the same cell line and the same conditions.

Conclusions and Perspectives

We developed the Barcode-Tracker strategy to identify and isolate clones of interest based on their intrinsic and predetermined response to a panel of drugs. Once the system is optimized, a barcode library will be generated and introduced at low MOI into PC9 cells containing the other Barcode-Tracker components. After amplification, the cells will be frozen down or treated with different compounds, including EGFR-TKIs and chemotherapy agents, to identify a few barcodes of interest, *i.e.* labeling osimertinib tolerant or highly sensitive populations. The corresponding sgRNAs will be generated and inserted in freshly thawed batches of cells containing the library. After isolation of single clones expressing the fluorescent reporters, the cells will be amplified, and the barcodes will be confirmed by Sanger sequencing. The resistant/tolerant/sensitive phenotypes of these clones will be assessed in the presence of different concentrations of osimertinib. To gain insights into the molecular mechanisms responsible for the different levels of responsiveness to osimertinib, we will compare the gene expression profile of the different selected clones, as well as of the mass population of parental cells using gene set enrichment analysis.

In our study, we used a high complex lentiviral barcode library to investigate the effects of targeted and cytotoxic therapies on the clonal dynamics of different subpopulations of PC9 cells. We found that certain cells display a predefined response to each drug, resulting in a specific barcode pattern that can be used to compare different compounds and investigate their mechanism of action. As a proof-of-concept of this approach, we have generated barcode profiles from 87 known drugs targeting various pathways and used it to predict the mechanism of action of a new compound, X-13271. In the future, it will be important to increase the number of screened drugs and possibly apply a similar workflow to other cell lines, each representative of a tumor type of interest. Moreover, to identify mechanisms potentially responsible for the predetermined drug response of certain cell subpopulations, scRNA-Seq analysis of untreated cells was performed and we are in the process of analyzing the data. Future experiments could be designed to compare the transcriptomic profile of different subpopulations of cancer cells before and after treatment.

Methods

1. Cell culture

The EGFR-mutant NSCLC PC9 cell line harboring a deletion in exon 19 of EGFR (ΔE746-A750) were obtained from ECACC (distributed by Sigma-Aldrich). The neuroblastoma cell line SKNAS were a kind gift from Dr. Jaume Mora (Hospital Sant Joan de Déu, Barcelona, Spain). Both cell lines were grown in RPMI 1640 (Life Technologies #21875-034). The human embryonic kidney 293T cells were obtained from ATCC and maintained in DMEM (Life Technologies #41965-039). All media were supplemented with 10% fetal bovine serum (FBS) (Life Technologies #10270-106) and 0.5% penicillin/streptomycin (Sigma #P4333). All cell lines were cultured at 37 °C in a humidified incubator with 5% CO_2 and were periodically tested for mycoplasma infections.

2. Generation of PC9 single-cell subclones and drug-tolerant persisters

PC9 parental cells (treatment-naïve) were seeded into 96-well plates at a density of 0.1 cell/well. Wells containing only single colonies were expanded for 3 weeks and used for the experiments. Tolerant cells were generated by treating PC9 parental cells for 3 weeks with 1 μM osimertinib (PC9-OT), 2 μM gefitinib (PC9-GT), 100 nM pemetrexed (PC9-PT) or 500 nM cisplatin (PC9-CT). Fresh media containing the relevant drugs were added to the cells every 3 days. Cells that survived the treatment were considered DTP.

3. Cell viability assays

Short-term cell viability was measured using CellTiter-Glo Luminescent reagent (Promega #G7571) according to manufacturer protocol. Approximately, 2 x 10^3 cells per well were seeded into 96-well plates (Costar) and allowed to attach overnight. The following day, cells were treated with the appropriate drug and their viability were measured 72 h after treatment using Infinite F200 PRO (TECAN). Data were derived from at least three independent experiments.

For long-term colony forming assays, 25 x 10^3 cells per well were seeded in six-well plates. The following day, cells were treated for 7 days with the indicated drugs. Media were

then removed, and cells were washed with phosphate buffered saline (PBS) (Sigma #D8537), fixed with a 10% methanol, 10% acetic acid solution, and stained with a 1% crystal violet (Merck #115940). Plates were rinsed with water and left for drying at room temperature (RT), after which images were captured using a ChemiDoc Imaging System.

4. Western Blot

For Western blotting, 5 x 10^5 cells were seeded in T-25 flasks, treated the next day and lysed at specified time points in a buffer containing 50 mM HEPES pH 7.6, 150 mM NaCl, 5 mM EDTA, NP40 0.5%, 20 mM NaF, 2 mM Na_3VO_4, supplemented with protease inhibitor mini tablets (ThermoFisher Scientific #88665). Lysates were cleared by centrifugation at 14,000 g for 15 min at 4°C and protein concentration was determined using the Bradford assay (Bio-Rad). Sodium dodecyl sulfate (SDS) loading buffer was added to equal amounts of lysates, followed by SDS-polyacrylamide gel electrophoresis (PAGE) and transfer to polyvinylidene fluoride membranes (Bio-Rad) using a Trans-Blot Turbo Transfer System (Bio-Rad). Membranes were blocked for 1 h at room temperature in 5% nonfat dried milk (Sigma #70166) in PBS and incubated overnight at 4°C with the primary antibodies in 3% BSA PBS-tween. The following antibodies, purchased from Cell Signaling, were used: EGF Receptor (#2232), Phospho-EGF Receptor (Tyr1068) (D7A5) (#3777), p44/42 MAPK (Erk1/2) (L34F12) (#4696), Phospho-p44/42 MAPK (Erk1/2) (Thr202/Tyr204) (D13.14.4E) (#4370), Akt (pan) (40D4) (#2920), Phospho-Akt (Ser473) (D9E) (#4060), Phospho-S6 Ribosomal Protein (Ser240/244) (#2215s), Histone H3 (#9715) and Acetyl-Histone H3 (Lys27) (D5E4) (#8173). After washing, the membranes were incubated 1 h with horseradish peroxidase-conjugated secondary antibodies and the bands were visualized by chemiluminescence using the Clarity Western ECL substrate (Bio-Rad #1705060) and a ChemiDoc Imaging System (Bio-Rad). The images were analyzed using the Image Lab Software (Bio-Rad).

5. DNA damage assay

The level of DNA damage was detected by using the fluorescent anti-γH2AX antibody (Alexa Fluor® 647 Mouse anti-H2AX (pS139), BD Biosciences, #560447) according to the

manufacturer's instructions. Briefly, 5 x 10^5 cells were seeded in T-25 flasks and treated the next day with or without pemetrexed (100 nM). 48h after treatment, the cells were detached, fixed with 4% paraformaldehyde for 15 min at RT, and permeabilized with ice-cold 100% methanol, followed by antibody staining for 1h at RT. After incubation stained cells were washed twice with PBS and analyzed with a BD LSRFortessa™ Cell Analyzer.

6. RNA extraction and qPCR

The indicated cells were seeded at 5 x 10^5 cells into T-25 flasks (4 technical replicates per condition), treated for 48 h in the presence or the absence of osimertinib (100 nM). Cells were then washed with PBS and total RNA was extracted using TRI Reagent (Sigma-Aldrich) and purified with NucleoSpin RNA columns (Macherey-Nagel) following the manufacturer's instruction. RNA was reverse transcribed using SensiFAST™ cDNA Synthesis Kit (Bioline/Meridian Bioscience #BIO-65053). q-PCR were performed using Fast SYBR Green Master Mix (Applied Biosystems) on a QuantStudio Flex Real-Time PCR System (ThermoFisher Scientific). Relative expression of DUSP6 was performed according to the standard curve method using human TATA box binding protein (TBP) as a reference gene for normalization. The primers used were as follows: DUSP6-Fw (5'-CTGGAACGAGAATACGGGCG-3') and DUSP6-Rv (5'-CTTACTGAAGCCACCTTCCAGG-3'); TBP-Fw (5'-TTGTACCGCAGCTGCAAAAT-3') and TBP-Rv (5'-TATATTCGGCGTTTCGGGCA-3'). Statistical analysis was performed with the GraphPad Prism 4 software using Mann-Whitney's test.

7. CRISPR-barcoding

Partially degenerate barcode sequences were inserted into the safe harbor AAVS1 locus of PC9 cells using CRISPR/Cas9 technology as previously described (Guernet et al., 2016). Five replicates of the cells (5 x 10^5 cells per T-75 flask) were then treated with or without osimertinib (100 nM), gefitinib (2 μM), trametinib (30 nM) or pemetrexed (100 nM). Cells were collected following 2 weeks of treatment and gDNA was extracted using NucleoSpin® Tissue kit (Machery-Nagel). Targeted amplification of the integrated barcode in the AAVS1 locus was

performed using previously described (Guernet et al., 2016) primers containing Illumina adapter sequences and 6 bp unique indexes (Table 6). For each sample, we performed 3 PCR reactions, each from 500 ng of genomic DNA in a final volume of 50 µl, using Herculase II Fusion DNA Polymerase (Agilent technologies #600677) and the following program: 98°C for 5 minutes; followed by 27 cycles of 20 s at 98°C, 20 s at 60°C and 30 s at 72°C; final extension at 72°C for 3 min. The PCR products from the same sample were pooled and purified over 2% agarose gels (band size at 275 bp) using NucleoSpin® Gel and PCR clean-up kit (Macherey-Nagel). Purified amplicons were quantified by Qubit (Invitrogen #Q33231) and their quality was assessed by Bioanalyzer (Agilent #50671504). Sequencing was performed on an Illumina MiniSeq, using 150-cycles High Output Reagent Kit (#FC-420-1002). Counts of barcodes for each sample were extracted from FASTQ files using galaxy (https://usegalaxy.org). Spearman correlations were calculated using Excel (Microsoft).

8. Lentiviral barcode library construction

VIRHD containing puromycin and polyA sequence was digested overnight with EcoRV at 37°C followed by gel purification using NucleoSpin® Gel and PCR clean-up kit. The barcode library was ordered as an Ultramer DNA oligo from Integrated DNA Technologies (IDT) and the double-stranded barcodes were generated by preforming a single cycle extension reaction with a reverse primer (Table 7) using DreamTaq DNA Polymerase (ThermoFisher Scientific #K1081) according to the manufacturer's protocol. The resulting reaction was run on 2% agarose gel and purified using the NucleoSpin® Gel and PCR clean-up kit. The fragments were then ligated with 100 ng of EcoRV digested VIRHD backbone in 5:1 molar ratio using Gibson assembly (NEB #E2621S). Reactions were then purified by columns (NucleoSpin Gel and PCR Clean-up) and electroporated into One Shot™ TOP10 Electrocomp™ (Invitrogen #C404052). After 1 h recovery period, a serial dilution of the bacteria was plated onto LB-ampicillin plates to calculate the transformation efficiency. The remainder of the recovered bacteria was grown overnight in liquid culture (Terrific Broth, Sigma #T5574), containing 100 µg/ml ampicillin at 37 °C. Plasmids were extracted using NucleoBond Xtra Midi columns (Macherey-Nagel).

9. Lentivirus production and barcoding of PC9 cells

Lentiviral particles were produced by transfecting 293T cells with 5 µg of VIRHD containing the barcode sequences, 3 µg of pCMV-dR8.91 and 1.5 µg of pMD-VSV-G packaging plasmids using polyethylenimine (Polysciences #23966) according to the manufacturer's instructions. Lentivirus was collected at 48 and 72 h post-transfection, centrifuged and filtered. The virus was supplemented with 8 µg/ml polybrene, aliquoted and stored at −80 °C for later use.

PC9 parental cells were transduced with the library at low MOI as previously described (Bhang et al., 2015). After 24h incubation, virus was removed and cells were selected with medium containing 2 µg/ml puromycin (ChemCruz #C1518) for 3 days, after which 25 x 10^3 barcoded cells were isolated, quickly amplified and cryopreserved in aliquots of 2 x 10^6 cells per vial in liquid nitrogen.

10. Drug screen

Different vials of PC9 barcoded cells were thawed, expanded in culture for three passages, and screened against a custom compound library containing previously identified inhibitors with well-defined mechanisms of action. Our library is composed of 87 small molecules, which are either approved for clinical use or are tool compounds with known activities (Table 8). For the screen, the cells were plated in quadruplicate per treatment into T-75 flasks (5 x 10^5 cells per replicate), and treated the following day with different drugs. After 9 days of treatment, the cells were harvested and the gDNA was extracted using NucleoSpin® Tissue kit according to manufacturer's instructions. Barcodes were then PCR-amplified using primers containing Illumina adaptors and 6 bp unique index sequences (Table 9). For each sample, up to 1.5 µg of gDNA was used as a template, and was amplified using Herculase II Fusion DNA Polymerase in 3 individual PCR reactions each consisting of 10 µl of 5× Herculase II reaction buffer, 0.5 µl dNTP mix (25 mM each dNTP), 2 µl DMSO, 0.5 µl of Herculase II fusion DNA polymerase, 500 ng of gDNA, 2 µl primer (10 µM) Fw, 2 µl primer (10 µM) Rv and nuclease free water to a final volume of 50 µl. Thermal cycler conditions were: 98°C for 3 minutes; followed by 27 cycles of 20 s at 98°C, 20 s at 60°C and 30 s at 72°C; final extension at 72°C for 3 min. The PCR

products from the same sample were pooled, purified over 2% agarose gels (band size at 450 bp) using NucleoSpin® Gel and PCR clean-up kit, quantified using Qubit and their quality was assessed by Bioanalyzer. Sequencing was then performed on an Illumina MiniSeq with single-end 75 bp reads (#FC-420-1001).

11. Bioinformatics analysis of the drug screen data

After sequencing, barcode count data for each sample were extracted from FASTQ files using galaxy (https://usegalaxy.org). Our dataset contains in total 505 samples including 40 control samples and 18 time zeros samples representing 9 experiments spanning over 11 NGS runs. Barcode sequences associated with each sample were recorded in a barcode-count matrix, normalized by sample to one million (i.e. such that the sum over all barcodes per sample equaled one million), reaching a total of 3,120,235 unique barcodes. To filter out non-specific PCR amplification, barcodes detected in less than 5 control samples and 5 time zeros over all the samples were discarded. Filtered data composed of 12,305 barcodes was then analyzed using the statistical computing language R, specifically the package DeSEQ2 to evaluate log2 fold change (FC) between condition and control samples within each experiment to the drug-specific effects on the clonal architecture in terms of barcode profile. For determining the significance of the differences of logFC between two conditions, the Wald test was performed on log-transformed values, and p-values <0.05 were considered statistically significant.

12. Barcode-Tracker activator constructs and cell line generation

To generate Barcode-Tracker activator construct, two minimal promoters separated by the barcode cloning site, driven the expression of GFP and mCherry were cloned using Gibson assembly into VIRHD empty backbone containing puromycin resistance gene. A DNA barcode sequence was ordered from IDT and inserted into a barcode cloning site using BbsI restriction enzyme. The CMV and YB-TATA minimal promoters were ordered as Ultramer DNA oligo from IDT (Table 10). The GFP and mCherry were amplified using Herculase II Fusion DNA Polymerase from pEGFP-C1 (Clontech #6084-1) and mCherry2-N1 (Addgene #54517)

respectively. The sgRNAs were ordered from IDT as sense and antisense oligos (Table 10), annealed and cloned into Lenti-sgRNA-MS2-Zeo (Addgene #61427).

To establish the Barcode-Tracker cell line, 293T cells were first infected with the Leni-dCAS-VP64-Blast (Addgene #61425) and lenti-MPHv2 (Addgene #89308). The cells were then transfected with the Barcode-Tracker construct and a barcode-specific sgRNA (sgRNA-A/B) or a control sgRNA (sgRNA-mut) using polyethylenimine. After 72h of transfection, the cells were harvested and the fluorescent protein expression levels were measured on a Cytoflex flow cytometer (Beckman Coulter).

13. Construction of Barcode-Tracker toggle switch plasmids

The toggle switch is based on two repressor cassettes. In repressor 1 (R1), seven repeats of the Tet response element (ordered as Ultramer DNA oligo from IDT) and CMV promoter amplified by PCR from pBIND vector (Promega #E2440) were cloned into VIRHD. Downstream of the CMV promoter, we cloned mCherry, P2A, puromycin, T2A and cumate coding sequences (CymRV5), followed by shRNA targeting GFP. The cumate repressor was amplified from epB-UbC-CymRV5-nls-GFP-DEx2 (Addgene #119906). The self-cleaving peptides and the shRNA sequence were ordered as gBlocks from IDT. R1 is flanked by two anti-repressor elements, which were amplified from pLS-mP (Addgene #81225). Gibson assembly was used to clone the different components of R1.

Repressor 2 (R2) was cloned using Gibson assembly into pBluescript backbone, which serves as donor DNA vector for knock-in experiment. Two donor DNA vectors were constructed. The first construct consist of CMV promoter containing five repeats of cumate response elements, amplified by PCR from pCuo CA Rac1 (Addgene #84643), GFP and polyA site. The cassette is flanked by left and right homology arms complementary to AAVS1 site (amplified from cDNA). The second donor DNA vector is composed of P2A, zeomycin, T2A and TetR-KRAB, flanked by left and right homologous arms complementary to GFP and PolyA respectively. Tet fused to KRAB was amplified from pLVUT-tTR-KRAB (Addgene #11651).

14. Generation of stable cell line expressing the toggle switch circuits

Repressor 2 was inserted into the AAVS1 locus of PC9 cells using CRISPR/Cas9 technology. We first transfected the cells with a donor DNA vector containing CMV promoter, cumate operator, GFP and PolyA, and a sgRNA targeting the AAVS1 locus (sgRNA-1) (Table 11). After transfection, the GFP positive cells were sorted using FACS. The isolated cells were then transfected with a second donor DNA vector containing P2A, zeomycin, T2A and tTR-KRAB, and sgRNA-2 targeted the inserted locus (Table 11). After transfection, the cells were selected with 300 µg/ml zeomycin (Invitrogen #R25001) to enrich for cells expressing R2. To introduce the dCas9 repressor, the cells expressing R2 was infected with the dCas9-KRAB-MeCP2 (Addgene #110824) and then selected with 10 µg/ml blasticidin (ThermoFisher Scientific #A1113903). All transfection were performed using a Nucleofector II and Amaxa Nucleofector kits (Lonza #VCA-1005) according to the manufacturer's instructions.

To test the toggle switch system, the cells containing R2 and dCas9-KRAB-MeCP2 were transduced with R1 and selected with 1 µg/ml puromycin (ChemCruz #C1518). The resulting cells were then treated with doxycycline (Santa Cruz Biotechnology #sc-204734) or cumic acid (Sigma #268402) and the expression levels of GFP and mCherry were assessed using Cytoflex flow cytometer.

Table 6. List of oligonucleotides used for CRISPR-Barcoding.

Name	Sequence 5'-3'
AAVS1-Barcode-Sequence	GTGTCCCCGAGCTGGGACCACCTTATATTCCCAGGGCCGGTTAATGTG GCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACANNNNNatN NNNgtcGACAGGATTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCC TTCCTAGTCTCCTGATATTGGGTCTAACCCCCACCTCCTGTTAGGCA
NGS_AAVS1_Fw	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCT CTTCCGATCTGTTAATGTGGCTCTGGTTCTGG
NGS_AAVS1_Ind 1_Rv	CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 2_Rv	CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 3_Rv	CAAGCAGAAGACGGCATACGAGATGCCTAAGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 4_Rv	CAAGCAGAAGACGGCATACGAGATTGGTCAGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 5_Rv	CAAGCAGAAGACGGCATACGAGATCACTGTGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 6_Rv	CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 7_Rv	CAAGCAGAAGACGGCATACGAGATGATCTGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 8_Rv	CAAGCAGAAGACGGCATACGAGATTCAAGTGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 9_Rv	CAAGCAGAAGACGGCATACGAGATCTGATCGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 10_Rv	CAAGCAGAAGACGGCATACGAGATAAGCTAGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 11_Rv	CAAGCAGAAGACGGCATACGAGATGTAGCCGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 12_Rv	CAAGCAGAAGACGGCATACGAGATTACAAGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 13_Rv	CAAGCAGAAGACGGCATACGAGATTTGACTGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 14_Rv	CAAGCAGAAGACGGCATACGAGATGGAACTGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 15_Rv	CAAGCAGAAGACGGCATACGAGATTGACATGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 16_Rv	CAAGCAGAAGACGGCATACGAGATGGACGGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 17_Rv	CAAGCAGAAGACGGCATACGAGATCTCTACGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC

Name	Sequence 5'-3'
NGS_AAVS1_Ind 18_Rv	CAAGCAGAAGACGGCATACGAGATGCGGACGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 19_Rv	CAAGCAGAAGACGGCATACGAGATTTTCACGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 20_Rv	CAAGCAGAAGACGGCATACGAGATGGCCACGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 21_Rv	CAAGCAGAAGACGGCATACGAGATCGAAACGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 22_Rv	CAAGCAGAAGACGGCATACGAGATCGTACGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 23_Rv	CAAGCAGAAGACGGCATACGAGATCCACTCGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 24_Rv	CAAGCAGAAGACGGCATACGAGATGCTACCGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 25_Rv	CAAGCAGAAGACGGCATACGAGATATCAGTGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 26_Rv	CAAGCAGAAGACGGCATACGAGATGCTCATGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 27_Rv	CAAGCAGAAGACGGCATACGAGATAGGAATGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 28_Rv	CAAGCAGAAGACGGCATACGAGATCTTTTGGTGACTGGAGTTCAGACG TGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 29_Rv	CAAGCAGAAGACGGCATACGAGATTAGTTGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 30_Rv	CAAGCAGAAGACGGCATACGAGATCCGGTGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 31_Rv	CAAGCAGAAGACGGCATACGAGATATCGTGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 32_Rv	CAAGCAGAAGACGGCATACGAGATTGAGTGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 33_Rv	CAAGCAGAAGACGGCATACGAGATCGCCTGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 34_Rv	CAAGCAGAAGACGGCATACGAGATGCCATGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 35_Rv	CAAGCAGAAGACGGCATACGAGATAAAATGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 36_Rv	CAAGCAGAAGACGGCATACGAGATTGTTGGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 37_Rv	CAAGCAGAAGACGGCATACGAGATATTCCGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC

Name	Sequence 5'-3'
NGS_AAVS1_Ind 38_Rv	CAAGCAGAAGACGGCATACGAGATAGCTAGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 39_Rv	CAAGCAGAAGACGGCATACGAGATGTATAGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 40_Rv	CAAGCAGAAGACGGCATACGAGATTCTGAGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 41_Rv	CAAGCAGAAGACGGCATACGAGATGTCGTCGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 42_Rv	CAAGCAGAAGACGGCATACGAGATCGATTAGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 43_Rv	CAAGCAGAAGACGGCATACGAGATGCTGTAGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 44_Rv	CAAGCAGAAGACGGCATACGAGATATTATAGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 45_Rv	CAAGCAGAAGACGGCATACGAGATGAATGAGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 46_Rv	CAAGCAGAAGACGGCATACGAGATTCGGGAGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 47_Rv	CAAGCAGAAGACGGCATACGAGATCTTCGAGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC
NGS_AAVS1_Ind 48_Rv	CAAGCAGAAGACGGCATACGAGATTGCCGAGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC

Table 7. List of oligonucleotides used to generate the lentiviral barcode library.

Name	Sequence 5'-3'
Lenti-Barcode-Fw	GCAATAAACAAGTTCAAATCCCTCGGAAGCGAT*CWNSNNGANSATNWGCNNN NACTSNNACANSNNWNASG*ATCCGATGATCACTGTTCAGCTCAGGCACCGGG
Lenti-Barcode-Rv	CCCGGTGCCTGAGCTGAACAG

Table 8. Compounds used for this study.

Drug	Concentration	Source	Catalog number
Osimertinib	100 nM	LC laboratories	C-7200
WZ4002	1 µM	Santa Cruz Biotechnology	sc-364655
Rociletinib	1 µM	MedChemExpress	HY-15729
Lazertinib	1 µM	MedChemExpress	HY-109061
Mavelertinib	1 µM	MedChemExpress	HY-12972

Drug	Concentration	Source	Catalog number
Gefitinib	2 µM	Santa Cruz Biotechnology	sc-202166A
Gefitinib based PROTAC-1	10 µM	MedChemExpress	HY-123921
Sorafenib	5 µM	Santa Cruz Biotechnology	sc-220125A
Regorafenib	10 µM	TargetMOI	T1792
Sunitinib Malate	15 µM	SelleckChem	S1042
Lenvatinib	40 µM	LC laboratories	L-5400
Cabozantinib	10 µM	LC laboratories	C-8901
Pemetrexed	100 nM	SelleckChem	S1135
Methotrexate	40 nM	MedChemExpress	HY-14519
Pralatrexate	3 nM	MedChemExpress	HY-10446
5-Fluorouracil	6 µM	MedChemExpress	HY-90006
Carboplatin	7 µM	MedChemExpress	HY-17393
Oxaliplatin	5 µM	MedChemExpress	HY-17371
Paclitaxel	5 nM	MedChemExpress	HY-B0015
Cisplatin	500 nM	SelleckChem	S1166
Doxorubicin hydrochloride	80 nM	Abcam	ab120629
Mitomycin C	0.02 µg/ml	Sigma	M0503
U0126-EtOH	30 µM	MedChemExpress	HY-12031
Trametinib	30 nM	Santa Cruz Biotechnology	sc-364639
SAHA (Vorinostat)	2 µM	Sigma	SML0061
Trichostatin A	110 nM	Sigma	T1952
Sodium lbutyrate	2 mM	Santa Cruz Biotechnology	sc-202341B
Sodium phenylbutyrate (4PBA)	2 mM	Santa Cruz Biotechnology	sc-200652A
Tazemetostat (EPZ-6438)	40 µM	SelleckChem	S7128
5-Azacytidine	1,5 µM	Sigma	A2385
Capivasertib (AZD5363)	50 µM	MedChemExpress	HY-15431
Chloroquine	40 µM	Sigma	C6628
Spautin-1	7 µM	SelleckChem	S7888
SBI-0206965	5 µM	Sigma	SML1540
Bafilomycin A1	1,2 nM	SelleckChem	S1413
Trimethoprim	600 µM	Acros Organics	738-70-5
Nigericin	100 nM	Invitrogen	N1495
Erastin	15 µM	SelleckChem	S7242
Ferrostatin-1	60 µM	Sigma	SML0583
Deferoxamine mesylate salt	7 µM	Sigma	D9533
S63845	10 µM	MedChemExpress	HY-100741
Navitoclax	6 µM	MedChemExpress	HY-10087
eFT-508	25 µM	SelleckChem	S8275
CGP57380	10 µM	Sigma	C0993
XAV-939	50 µM	MedChemExpress	HY-15147
LGK-974	5 µM	APEXBIO	B2307
MG-132	510 nM	MedChemExpress	HY-13259

Bortezomib	20 nM	MedChemExpress	HY-10227
THZ1	90 nM	MedChemExpress	HY-80013
Palbociclib isethionate	10 µM	MedChemExpress	HY-A0065
STO-609	30 µM	MedChemExpress	HY-19805
Napabucasin	1 µM	SelleckChem	S7977
Stat3 inhibitor V, stattic	7 µM	Santa Cruz Biotechnology	sc-202818
AZD1480	5 µM	Santa Cruz Biotechnology	sc-364735
SHP099 HCL	50 µM	SelleckChem	S8278
LIM Kinase Inhibitor I, LIMKi 3	30 µM	Calbiochem	435930
Y-27632 dihydrochloride	300 µM	MedChemExpress	HY-10583
Ibudilast	400 µM	MedChemExpress	HY-B0763
Phenformin HCL	50 µM	SelleckChem	S2542
Dorsomorphin (Compound C) 2HCl	8 µM	SelleckChem	S7306
LY3009120	3 µM	SelleckChem	S7842
Vemurafenib	7 µM	MedChemExpress	HY-12057
RAF265 (CHIR-265)	5 µM	SelleckChem	S2161
CCT196969	1 µM	SelleckChem	S7743
ML-210	30 µM	MedChemExpress	HY-100003
IMR-1	80 µM	MedChemExpress	HY-100431
CCT020312	3 µM	MedChemExpress	HY-119240
Tunicamycin	0.5 µg/ml	Sigma	SML1287
Thapsigargin	8 nM	Calbiochem	586006-2MG
L-755,507	20 µM	Santa Cruz Biotechnology	sc-204045
Infigratinib (BGJ398)	7 µM	SelleckChem	S2183
Pluripotin (SC-1)	10 µM	Sigma	SML0858
Olaparib	20 µM	MedChemExpress	HY-10162
Crizotinib	2 µM	Sigma	PZ0191
TAE684 (NVP-TAE684)	1 µM	SelleckChem	S1108
Sotorasib (AMG-510)	35 µM	MedChemExpress	HY-114277
Nutlin-3	25 µM	SelleckChem	S1061
N-acetylcysteine (NAC)	15 mM	Sigma	A9165
Mefloquine hydrochloride	20 µM	MedChemExpress	HY-17437A
Probenecid	500 µM	MedChemExpress	HY-B0545
Saracatinib	500 nM	MedChemExpress	HY-10234
Alisertib (MLN 8237)	50 nM	MedChemExpress	HY-10971
BMS-536924	2 µM	MedChemExpress	HY-10262
JSH-23	60 µM	MedChemExpress	HY-13982
Temsirolimus	6 µM	MedChemExpress	HY-50910
AZD8186	50 µM	MedChemExpress	HY-12330
Levofloxacin	250 µM	MedChemExpress	HY-B0330

Table 9. Primers used for deep sequencing for lentiviral barcode library.

Name	Sequence 5'-3'
NGS_Lenti Puro_Fw	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCT CTTCCGATCTGTTCAAATCCCTCGGAAGC
NGS_Lenti Puro_Ind 1_Rv	CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 2_Rv	CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 3_Rv	CAAGCAGAAGACGGCATACGAGATGCCTAAGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 4_Rv	CAAGCAGAAGACGGCATACGAGATTGGTCAGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 5_Rv	CAAGCAGAAGACGGCATACGAGATCACTGTGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 6_Rv	CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 7_Rv	CAAGCAGAAGACGGCATACGAGATGATCTGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 8_Rv	CAAGCAGAAGACGGCATACGAGATTCAAGTGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 9_Rv	CAAGCAGAAGACGGCATACGAGATCTGATCGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 10_Rv	CAAGCAGAAGACGGCATACGAGATAAGCTAGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 11_Rv	CAAGCAGAAGACGGCATACGAGATGTAGCCGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 12_Rv	CAAGCAGAAGACGGCATACGAGATTACAAGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 13_Rv	CAAGCAGAAGACGGCATACGAGATTTGACTGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 14_Rv	CAAGCAGAAGACGGCATACGAGATGGAACTGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 15_Rv	CAAGCAGAAGACGGCATACGAGATTGACATGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 16_Rv	CAAGCAGAAGACGGCATACGAGATGGACGGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 17_Rv	CAAGCAGAAGACGGCATACGAGATCTCTACGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 18_Rv	CAAGCAGAAGACGGCATACGAGATGCGGACGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC

Name	**Sequence 5'-3'**
NGS_Lenti Puro_Ind 19_Rv	CAAGCAGAAGACGGCATACGAGATTTTCACGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 20_Rv	CAAGCAGAAGACGGCATACGAGATGGCCACGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 21_Rv	CAAGCAGAAGACGGCATACGAGATCGAAACGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 22_Rv	CAAGCAGAAGACGGCATACGAGATCGTACGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 23_Rv	CAAGCAGAAGACGGCATACGAGATCCACTCGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 24_Rv	CAAGCAGAAGACGGCATACGAGATGCTACCGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 25_Rv	CAAGCAGAAGACGGCATACGAGATATCAGTGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 26_Rv	CAAGCAGAAGACGGCATACGAGATGCTCATGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 27_Rv	CAAGCAGAAGACGGCATACGAGATAGGAATGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 28_Rv	CAAGCAGAAGACGGCATACGAGATCTTTTGGTGACTGGAGTTCAGACG TGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 29_Rv	CAAGCAGAAGACGGCATACGAGATTAGTTGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 30_Rv	CAAGCAGAAGACGGCATACGAGATCCGGTGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 31_Rv	CAAGCAGAAGACGGCATACGAGATATCGTGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 32_Rv	CAAGCAGAAGACGGCATACGAGATTGAGTGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 33_Rv	CAAGCAGAAGACGGCATACGAGATCGCCTGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 34_Rv	CAAGCAGAAGACGGCATACGAGATGCCATGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 35_Rv	CAAGCAGAAGACGGCATACGAGATAAAATGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 36_Rv	CAAGCAGAAGACGGCATACGAGATTGTTGGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 37_Rv	CAAGCAGAAGACGGCATACGAGATATTCCGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 38_Rv	CAAGCAGAAGACGGCATACGAGATAGCTAGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC

Name	Sequence 5'-3'
NGS_Lenti Puro_Ind 39_Rv	CAAGCAGAAGACGGCATACGAGATGTATAGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 40_Rv	CAAGCAGAAGACGGCATACGAGATTCTGAGGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 41_Rv	CAAGCAGAAGACGGCATACGAGATGTCGTCGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 42_Rv	CAAGCAGAAGACGGCATACGAGATCGATTAGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 43_Rv	CAAGCAGAAGACGGCATACGAGATGCTGTAGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 44_Rv	CAAGCAGAAGACGGCATACGAGATATTATAGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 45_Rv	CAAGCAGAAGACGGCATACGAGATGAATGAGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 46_Rv	CAAGCAGAAGACGGCATACGAGATTCGGGAGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 47_Rv	CAAGCAGAAGACGGCATACGAGATCTTCGAGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC
NGS_Lenti Puro_Ind 48_Rv	CAAGCAGAAGACGGCATACGAGATTGCCGAGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGTTCCTGGCCACCGTC

Table 10. List of oligonucleotides used to generate the Barcode-Tracker constructs.

Name	Sequence 5'-3'
mCMV-Fw	CCTCGCCCTTGCTCACCATCTCGAGGCGATCTGACGGTTCACTAAACGAGCT CTGCTTATATAGGCCTCCCACCGTACACGCCTACCTCGACATACGTAAACCA TGGTCTTCGAGAAGACATAGGTTTAGACTCTAGAGGGTATATAATG
mCMV-Rv	CGCCCTTGCTCACCATGGTGGGATCCGGTGGCGACCGGTGGATGCCGGAAT GCCAAGCTTTTTACCAACAGTACCGGAATGCCAAGCTGGAAGTCGAGCTTC CATTATATACCCTCTAGAGTCTAAACCTATGTCTTC
mYB-TATA-Fw	ATGCTCGAGCTACTACTACCAGAACTGTACAGCTATCAGCAGCTGGCCCCC ATTATATACCCTCTAGCAAACCATGGTCTTCGAGAAGACATAGGTTG
mYB-TATA-Rv	CATGGATCCGGTGGCTGATCGAGCGGTCAAGCGTTCTGGTAGTAGAGTAGT GGCCCCCATTATATACCCTCTAGCAACCTATGTCTTCTCGAAGACCA
Barcode sequence-Fw	GTTCCGCGTTACATAACTTA
Barcode sequence-Rv	TAAGTTATGTAACGCGGAAC
sgRNA (A)-Fw	CACCGTTCCGCGTTACATAACTTA

Name	Sequence 5'-3'
sgRNA (A)-Rv	AAACTAAGTTATGTAACGCGGAA
sgRNA (B)-Fw	CACCGTAAGTTATGTAACGCGGAAC
sgRNA (B)-Rv	AAACGTTCCGCGTTACATAACTTAC
sgRNA (mut)-Fw	CACCGTTGCTCATTACATAACTTA
sgRNA (mut)-Rv	AAACTAAGTTATGTAATGAGCAAC

Table 11. List of sgRNA used to knock in repressor 2 into the AAVS1 locus.

Name	Sequence 5'-3'
sgRNA (1)-Fw	GTCCCCTCCACCCCACAGTG
sgRNA (1)-Rv	CACTGTGGGGTGGAGGGGAC
sgRNA (2)-Fw	GTCACCAATCCTGTCCCTAG
sgRNA (2)-Rv	CTAGGGACAGGATTGGTGAC

References

Aceto, N., Bardia, A., Miyamoto, D.T., Donaldson, M.C., Wittner, B.S., Spencer, J.A., Yu, M., Pely, A., Engstrom, A., Zhu, H., Brannigan, B.W., Kapur, R., Stott, S.L., Shioda, T., Ramaswamy, S., Ting, D.T., Lin, C.P., Toner, M., Haber, D.A., and Maheswaran, S. (2014). Circulating Tumor Cell Clusters are Oligoclonal Precursors of Breast Cancer Metastasis. Cell *158*, 1110–1122.

Ahn, M.-J., Han, J.-Y., Lee, K.H., Kim, S.-W., Kim, D.-W., Lee, Y.-G., Cho, E.K., Kim, J.-H., Lee, G.-W., Lee, J.-S., Min, Y.J., Kim, J.-S., Lee, S.S., Kim, H.R., Hong, M.H., Ahn, J.S., Sun, J.-M., Kim, H.T., Lee, D.H., Kim, S., and Cho, B.C. (2019). Lazertinib in patients with EGFR mutation-positive advanced non-small-cell lung cancer: results from the dose escalation and dose expansion parts of a first-in-human, open-label, multicentre, phase 1-2 study. Lancet Oncol *20*, 1681–1690.

Alemany, A., Florescu, M., Baron, C.S., Peterson-Maduro, J., and van Oudenaarden, A. (2018). Whole-organism clone tracing using single-cell sequencing. Nature *556*, 108–112.

Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J., and Clarke, M.F. (2003). Prospective identification of tumorigenic breast cancer cells. PNAS *100*, 3983–3988.

Almendro, V., Marusyk, A., and Polyak, K. (2013). Cellular heterogeneity and molecular evolution in cancer. Annu Rev Pathol *8*, 277–302.

Anand, U., Dey, A., Chandel, A.K.S., Sanyal, R., Mishra, A., Pandey, D.K., De Falco, V., Upadhyay, A., Kandimalla, R., Chaudhary, A., Dhanjal, J.K., Dewanjee, S., Vallamkondu, J., and Pérez de la Lastra, J.M. (2022). Cancer chemotherapy and beyond: Current status, drug candidates, associated risks and progress in targeted therapeutics. Genes & Diseases.

Anderson, K., Lutz, C., van Delft, F.W., Bateman, C.M., Guo, Y., Colman, S.M., Kempski, H., Moorman, A.V., Titley, I., Swansbury, J., Kearney, L., Enver, T., and Greaves, M. (2011). Genetic variegation of clonal architecture and propagating cells in leukaemia. Nature *469*, 356–361.

Arasada, R.R., Shilo, K., Yamada, T., Zhang, J., Yano, S., Ghanem, R., Wang, W., Takeuchi, S., Fukuda, K., Katakami, N., Tomii, K., Ogushi, F., Nishioka, Y., Talabere, T., Misra, S., Duan, W., Fadda, P., Rahman, M.A., Nana-Sinkam, P., Evans, J., Amann, J., Tchekneva, E.E., Dikov, M.M., and Carbone, D.P. (2018). Notch3-dependent β-catenin signaling mediates EGFR TKI drug persistence in EGFR mutant NSCLC. Nat Commun *9*, 3198.

Awad, M.M. (2016). Impaired c-Met Receptor Degradation Mediated by MET Exon 14 Mutations in Non–Small-Cell Lung Cancer. JCO *34*, 879–881.

Baron, C.S., and van Oudenaarden, A. (2019). Unravelling cellular relationships during development and regeneration using genetic lineage tracing. Nat Rev Mol Cell Biol *20*, 753–765.

Bergethon, K., Shaw, A.T., Ignatius Ou, S.-H., Katayama, R., Lovly, C.M., McDonald, N.T., Massion, P.P., Siwak-Tapp, C., Gonzalez, A., Fang, R., Mark, E.J., Batten, J.M., Chen, H., Wilner, K.D., Kwak, E.L., Clark, J.W., Carbone, D.P., Ji, H., Engelman, J.A., Mino-Kenudson, M., Pao, W., and Iafrate, A.J. (2012). ROS1 Rearrangements Define a Unique Molecular Class of Lung Cancers. JCO *30*, 863–870.

Berns, K., Hijmans, E.M., Mullenders, J., Brummelkamp, T.R., Velds, A., Heimerikx, M., Kerkhoven, R.M., Madiredjo, M., Nijkamp, W., Weigelt, B., Agami, R., Ge, W., Cavet, G., Linsley, P.S., Beijersbergen, R.L., and Bernards, R. (2004). A large-scale RNAi screen in human cells identifies new components of the p53 pathway. Nature *428*, 431–437.

Bhang, H.C., Ruddy, D.A., Krishnamurthy Radhakrishna, V., Caushi, J.X., Zhao, R., Hims, M.M., Singh, A.P., Kao, I., Rakiec, D., Shaw, P., Balak, M., Raza, A., Ackley, E., Keen, N., Schlabach, M.R., Palmer, M., Leary, R.J., Chiang, D.Y., Sellers, W.R., Michor, F., Cooke, V.G., Korn, J.M., and Stegmeier, F. (2015). Studying clonal dynamics in response to cancer therapy using high-complexity barcoding. Nat. Med. *21*, 440–448.

Biddy, B.A., Kong, W., Kamimoto, K., Guo, C., Waye, S.E., Sun, T., and Morris, S.A. (2018). Single-cell mapping of lineage and identity in direct reprogramming. Nature *564*, 219–224.

Birchmeier, C., Sharma, S., and Wigler, M. (1987). Expression and rearrangement of the ROS1 gene in human glioblastoma cells. Proc Natl Acad Sci U S A *84*, 9270–9274.

Bock, C., Datlinger, P., Chardon, F., Coelho, M.A., Dong, M.B., Lawson, K.A., Lu, T., Maroc, L., Norman, T.M., Song, B., Stanley, G., Chen, S., Garnett, M., Li, W., Moffat, J., Qi, L.S., Shapiro, R.S., Shendure, J., Weissman, J.S., and Zhuang, X. (2022). High-content CRISPR screening. Nat Rev Methods Primers *2*, 1–23.

Bonnet, D., and Dick, J.E. (1997). Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nat Med *3*, 730–737.

Bowling, S., Sritharan, D., Osorio, F.G., Nguyen, M., Cheung, P., Rodriguez-Fraticelli, A., Patel, S., Yuan, W.-C., Fujiwara, Y., Li, B.E., Orkin, S.H., Hormoz, S., and Camargo, F.D. (2020). An Engineered CRISPR-Cas9 Mouse Line for Simultaneous Readout of Lineage Histories and Gene Expression Profiles in Single Cells. Cell *181*, 1410-1422.e27.

Brahmer, J., Reckamp, K.L., Baas, P., Crinò, L., Eberhardt, W.E.E., Poddubskaya, E., Antonia, S., Pluzanski, A., Vokes, E.E., Holgado, E., Waterhouse, D., Ready, N., Gainor, J., Arén Frontera, O., Havel, L., Steins, M., Garassino, M.C., Aerts, J.G., Domine, M., Paz-Ares, L., Reck, M., Baudelet, C., Harbison, C.T., Lestini, B., and Spigel, D.R. (2015). Nivolumab versus Docetaxel in Advanced Squamous-Cell Non-Small-Cell Lung Cancer. N Engl J Med *373*, 123–135.

Bramlett, C., Jiang, D., Nogalska, A., Eerdeng, J., Contreras, J., and Lu, R. (2020). Clonal tracking using embedded viral barcoding and high-throughput sequencing. Nat Protoc *15*, 1436–1458.

Buckingham, M.E., and Meilhac, S.M. (2011). Tracing cells for tracking cell lineage and clonal behavior. Dev Cell *21*, 394–409.

Bukowski, K., Kciuk, M., and Kontek, R. (2020). Mechanisms of Multidrug Resistance in Cancer Chemotherapy. Int J Mol Sci *21*, E3233.

Burslem, G.M., and Crews, C.M. (2020). Proteolysis-Targeting Chimeras as Therapeutics and Tools for Biological Discovery. Cell *181*, 102–114.

Byers, L.A., Diao, L., Wang, J., Saintigny, P., Girard, L., Peyton, M., Shen, L., Fan, Y., Giri, U., Tumula, P.K., Nilsson, M.B., Gudikote, J., Tran, H., Cardnell, R.J.G., Bearss, D.J., Warner, S.L., Foulks, J.M., Kanner, S.B., Gandhi, V., Krett, N., Rosen, S.T., Kim, E.S., Herbst, R.S., Blumenschein, G.R., Lee, J.J., Lippman, S.M., Ang, K.K., Mills, G.B., Hong, W.K., Weinstein, J.N., Wistuba, I.I., Coombes, K.R., Minna, J.D., and Heymach, J.V. (2013). An epithelial-mesenchymal transition gene signature predicts resistance to EGFR and PI3K inhibitors and identifies Axl as a therapeutic target for overcoming EGFR inhibitor resistance. Clin Cancer Res *19*, 279–290.

Canon, J., Rex, K., Saiki, A.Y., Mohr, C., Cooke, K., Bagal, D., Gaida, K., Holt, T., Knutson, C.G., Koppada, N., Lanman, B.A., Werner, J., Rapaport, A.S., San Miguel, T., Ortiz, R., Osgood, T., Sun, J.-R., Zhu, X., McCarter, J.D., Volak, L.P., Houk, B.E., Fakih, M.G., O'Neil, B.H., Price, T.J., Falchook, G.S., Desai, J., Kuo, J., Govindan, R., Hong, D.S., Ouyang, W., Henary, H., Arvedson, T., Cee, V.J., and Lipford, J.R. (2019). The clinical KRAS(G12C) inhibitor AMG 510 drives anti-tumour immunity. Nature *575*, 217–223.

Caunt, C.J., and Keyse, S.M. (2013). Dual-specificity MAP kinase phosphatases (MKPs): shaping the outcome of MAP kinase signalling. FEBS J *280*, 489–504.

Caunt, C.J., Sale, M.J., Smith, P.D., and Cook, S.J. (2015). MEK1 and MEK2 inhibitors and cancer therapy: the long and winding road. Nat Rev Cancer *15*, 577–592.

Chan, M.M., Smith, Z.D., Grosswendt, S., Kretzmer, H., Norman, T.M., Adamson, B., Jost, M., Quinn, J.J., Yang, D., Jones, M.G., Khodaverdian, A., Yosef, N., Meissner, A., and Weissman, J.S. (2019). Molecular recording of mammalian embryogenesis. Nature *570*, 77–82.

Chang, M.T., Shanahan, F., Nguyen, T.T.T., Staben, S.T., Gazzard, L., Yamazoe, S., Wertz, I.E., Piskol, R., Yang, Y.A., Modrusan, Z., Haley, B., Evangelista, M., Malek, S., Foster, S.A., and Ye, X. (2022). Identifying transcriptional programs underlying cancer drug response with TraCe-seq. Nat Biotechnol *40*, 86–93.

Chen, J., Sprouffske, K., Huang, Q., and Maley, C.C. (2011). Solving the Puzzle of Metastasis: The Evolution of Cell Migration in Neoplasms. PLOS ONE *6*, e17933.

Chen, Z., Fillmore, C.M., Hammerman, P.S., Kim, C.F., and Wong, K.-K. (2014). Non-small-cell lung cancers: a heterogeneous set of diseases. Nat Rev Cancer *14*, 535–546.

Comoglio, P.M., Trusolino, L., and Boccaccio, C. (2018). Known and novel roles of the MET oncogene in cancer: a coherent approach to targeted therapy. Nat Rev Cancer *18*, 341–358.

Cortot, A.B., Kherrouche, Z., Descarpentries, C., Wislez, M., Baldacci, S., Furlan, A., and Tulasne, D. (2017). Exon 14 Deleted MET Receptor as a New Biomarker and Target in Cancers. JNCI: Journal of the National Cancer Institute *109*, djw262.

Cuaron, J.J., Yorke, E.D., Foster, A., Hsu, M., Zhang, Z., Liu, F., Jackson, A., Mychalczak, B., Rosenzweig, K.E., Wu, A.J., and Rimner, A. (2013). Stereotactic Body Radiation Therapy for Primary Lung Cancers >3 Centimeters. Journal of Thoracic Oncology *8*, 1396–1401.

Dale, B., Cheng, M., Park, K.-S., Kaniskan, H.Ü., Xiong, Y., and Jin, J. (2021). Advancing targeted protein degradation for cancer therapy. Nat Rev Cancer *21*, 638–654.

Darwin, C. (1859). On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life. (London: John Murray).

Davidson, M.R., Gazdar, A.F., and Clarke, B.E. (2013). The pivotal role of pathology in the management of lung cancer. Journal of Thoracic Disease *5*.

Davis, A., Gao, R., and Navin, N. (2017). Tumor evolution: Linear, branching, neutral or punctuated? Biochimica et Biophysica Acta (BBA) - Reviews on Cancer *1867*, 151–161.

Deans, T.L., Cantor, C.R., and Collins, J.J. (2007). A Tunable Genetic Switch Based on RNAi and Repressor Proteins for Regulating Gene Expression in Mammalian Cells. Cell *130*, 363–372.

Deuschle, U., Meyer, W.K., and Thiesen, H.J. (1995). Tetracycline-reversible silencing of eukaryotic promoters. Mol Cell Biol *15*, 1907–1914.

Devarakonda, S., Morgensztern, D., and Govindan, R. (2015). Genomic alterations in lung adenocarcinoma. The Lancet Oncology *16*, e342–e351.

Dhimolea, E., de Matos Simoes, R., Kansara, D., Al'Khafaji, A., Bouyssou, J., Weng, X., Sharma, S., Raja, J., Awate, P., Shirasaki, R., Tang, H., Glassner, B.J., Liu, Z., Gao, D., Bryan, J., Bender, S., Roth, J., Scheffer, M., Jeselsohn, R., Gray, N.S., Georgakoudi, I., Vazquez, F., Tsherniak, A., Chen, Y., Welm, A., Duy, C., Melnick, A., Bartholdy, B., Brown, M., Culhane, A.C., and Mitsiades, C.S. (2021). An Embryonic Diapause-like Adaptation with Suppressed Myc Activity Enables Tumor Treatment Persistence. Cancer Cell *39*, 240-256.e11.

Dogan, S., Shen, R., Ang, D.C., Johnson, M.L., D'Angelo, S.P., Paik, P.K., Brzostowski, E.B., Riely, G.J., Kris, M.G., Zakowski, M.F., and Ladanyi, M. (2012). Molecular Epidemiology of EGFR and KRAS Mutations in 3,026 Lung Adenocarcinomas: Higher Susceptibility of Women to Smoking-Related KRAS-Mutant Cancers. Clin Cancer Res *18*, 6169–6177.

Drilon, A., Siena, S., Ou, S.-H.I., Patel, M., Ahn, M.J., Lee, J., Bauer, T.M., Farago, A.F., Wheler, J.J., Liu, S.V., Doebele, R., Giannetta, L., Cerea, G., Marrapese, G., Schirru, M., Amatu, A., Bencardino, K., Palmeri, L., Sartore-Bianchi, A., Vanzulli, A., Cresta, S., Damian, S., Duca, M., Ardini, E., Li, G., Christiansen, J., Kowalski, K., Johnson, A.D., Patel, R., Luo, D., Chow-Maneval, E., Hornby, Z., Multani, P.S., Shaw, A.T., and De Braud, F.G. (2017). Safety and Antitumor Activity of the Multitargeted Pan-TRK, ROS1, and ALK Inhibitor Entrectinib: Combined Results from Two Phase I Trials (ALKA-372-001 and STARTRK-1). Cancer Discov *7*, 400–409.

Drilon, A., Siena, S., Dziadziuszko, R., Barlesi, F., Krebs, M.G., Shaw, A.T., de Braud, F., Rolfo, C., Ahn, M.-J., Wolf, J., Seto, T., Cho, B.C., Patel, M.R., Chiu, C.-H., John, T., Goto, K., Karapetis, C.S., Arkenau, H.-T., Kim, S.-W., Ohe, Y., Li, Y.-C., Chae, Y.K., Chung, C.H., Otterson, G.A., Murakami, H., Lin, C.-C., Tan, D.S.W., Prenen, H., Riehl, T., Chow-Maneval, E., Simmons, B., Cui, N., Johnson, A., Eng, S., Wilson, T.R., and Doebele, R.C. (2020). Entrectinib in ROS1 fusion-positive non-small-cell lung cancer: integrated analysis of three phase 1–2 trials. Lancet Oncol *21*, 261–270.

Du, Y., Chen, Y., Wang, Y., Chen, J., Lu, X., Zhang, L., Li, Y., Wang, Z., Ye, G., and Zhang, G. (2022). HJM-561, a Potent, Selective, and Orally Bioavailable EGFR PROTAC that Overcomes Osimertinib-Resistant EGFR Triple Mutations. Molecular Cancer Therapeutics *21*, 1060–1066.

Ede, C., Chen, X., Lin, M.-Y., and Chen, Y.Y. (2016). Quantitative Analyses of Core Promoters Enable Precise Engineering of Regulated Gene Expression in Mammalian Cells. ACS Synth Biol *5*, 395–404.

Emert, B.L., Cote, C.J., Torre, E.A., Dardani, I.P., Jiang, C.L., Jain, N., Shaffer, S.M., and Raj, A. (2021). Variability within rare cell states enables multiple paths toward drug resistance. Nat Biotechnol *39*, 865–876.

Engelman, J.A., Zejnullahu, K., Mitsudomi, T., Song, Y., Hyland, C., Park, J.O., Lindeman, N., Gale, C.-M., Zhao, X., Christensen, J., Kosaka, T., Holmes, A.J., Rogers, A.M., Cappuzzo, F., Mok, T., Lee, C., Johnson, B.E., Cantley, L.C., and Jänne, P.A. (2007). MET Amplification Leads to Gefitinib Resistance in Lung Cancer by Activating ERBB3 Signaling. Science.

Ercan, D., Xu, C., Yanagita, M., Monast, C.S., Pratilas, C.A., Montero, J., Butaney, M., Shimamura, T., Sholl, L., Ivanova, E.V., Tadi, M., Rogers, A., Repellin, C., Capelletti, M., Maertens, O., Goetz, E.M., Letai, A., Garraway, L.A., Lazzara, M.J., Rosen, N., Gray, N.S., Wong, K.-K., and Jänne, P.A. (2012). Reactivation of ERK Signaling Causes Resistance to EGFR Kinase Inhibitors. Cancer Discovery *2*, 934–947.

Ferlenghi, F., Scalvini, L., Vacondio, F., Castelli, R., Bozza, N., Marseglia, G., Rivara, S., Lodola, A., La Monica, S., Minari, R., Petronini, P.G., Alfieri, R., Tiseo, M., and Mor, M. (2021). A sulfonyl fluoride derivative inhibits EGFRL858R/T790M/C797S by covalent modification of the catalytic lysine. European Journal of Medicinal Chemistry *225*, 113786.

Fisher, R.A., Gollan, B., and Helaine, S. (2017). Persistent bacterial infections and persister cells. Nat Rev Microbiol *15*, 453–464.

Friedlaender, A., Subbiah, V., Russo, A., Banna, G.L., Malapelle, U., Rolfo, C., and Addeo, A. (2022). EGFR and HER2 exon 20 insertions in solid tumours: from biology to treatment. Nat Rev Clin Oncol *19*, 51–69.

Gainor, J.F., and Shaw, A.T. (2013). Novel Targets in Non-Small Cell Lung Cancer: ROS1 and RET Fusions. Oncologist *18*, 865–875.

Gardner, T.S., Cantor, C.R., and Collins, J.J. (2000). Construction of a genetic toggle switch in Escherichia coli. Nature *403*, 339–342.

Gazdar, A.F., Bunn, P.A., and Minna, J.D. (2017). Small-cell lung cancer: what we know, what we need to know and the path forward. Nat Rev Cancer *17*, 725–737.

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., Lim, W.A., Weissman, J.S., and Qi, L.S. (2013). CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. Cell *154*, 442–451.

Gisselsson, D., Pettersson, L., Höglund, M., Heidenblad, M., Gorunova, L., Wiegant, J., Mertens, F., Cin, P.D., Mitelman, F., and Mandahl, N. (2000). Chromosomal breakage-fusion-bridge events cause genetic intratumor heterogeneity. PNAS *97*, 5357–5362.

Goldstraw, P., Chansky, K., Crowley, J., Rami-Porta, R., Asamura, H., Eberhardt, W.E.E., Nicholson, A.G., Groome, P., Mitchell, A., Bolejack, V., Goldstraw, P., Rami-Porta, R., Asamura, H., Ball, D., Beer, D.G., Beyruti, R., Bolejack, V., Chansky, K., Crowley, J., Detterbeck, F., Eberhardt, W.E.E., Edwards, J., Galateau-Sallé, F., Giroux, D., Gleeson, F., Groome, P., Huang, J., Kennedy, C., Kim, J., Kim, Y.T., Kingsbury, L., Kondo, H., Krasnik, M., Kubota, K., Lerut, A., Lyons, G., Marino, M., Marom, E.M., Meerbeeck, J. van, et al. (2016). The IASLC Lung Cancer Staging Project: Proposals for Revision of the TNM Stage Groupings in the Forthcoming (Eighth) Edition of the TNM Classification for Lung Cancer. Journal of Thoracic Oncology *11*, 39–51.

Goyal, Y., Dardani, I.P., Busch, G.T., Emert, B., Fingerman, D., Kaur, A., Jain, N., Mellis, I.A., Li, J., Kiani, K., Fane, M.E., Weeraratna, A.T., Herlyn, M., and Raj, A. (2021). Pre-determined diversity in resistant fates emerges from homogenous cells after anti-cancer drug treatment. 2021.12.08.471833.

Grumolato, L., Liu, G., Mong, P., Mudbhary, R., Biswas, R., Arroyave, R., Vijayakumar, S., Economides, A.N., and Aaronson, S.A. (2010). Canonical and noncanonical Wnts use a common mechanism to activate completely unrelated coreceptors. Genes Dev *24*, 2517–2530.

Guan, J., Umapathy, G., Yamazaki, Y., Wolfstetter, G., Mendoza, P., Pfeifer, K., Mohammed, A., Hugosson, F., Zhang, H., Hsu, A.W., Halenbeck, R., Hallberg, B., and Palmer, R.H. (2015). FAM150A and FAM150B are activating ligands for anaplastic lymphoma kinase. Elife *4*, e09811.

Guernet, A., Mungamuri, S.K., Cartier, D., Sachidanandam, R., Jayaprakash, A., Adriouch, S., Vezain, M., Charbonnier, F., Rohkin, G., Coutant, S., Yao, S., Ainani, H., Alexandre, D., Tournier, I., Boyer, O., Aaronson, S.A., Anouar, Y., and Grumolato, L. (2016). CRISPR-Barcoding for Intratumor Genetic Heterogeneity Modeling and Functional Analysis of Oncogenic Driver Mutations. Mol Cell *63*, 526–538.

Guler, G.D., Tindell, C.A., Pitti, R., Wilson, C., Nichols, K., KaiWai Cheung, T., Kim, H.-J., Wongchenko, M., Yan, Y., Haley, B., Cuellar, T., Webster, J., Alag, N., Hegde, G., Jackson, E., Nance, T.L., Giresi, P.G., Chen, K.-B., Liu, J., Jhunjhunwala, S., Settleman, J., Stephan, J.-P., Arnott, D., and Classon, M. (2017). Repression of Stress-Induced LINE-1 Expression Protects Cancer Cell Subpopulations from Lethal Drug Exposure. Cancer Cell *32*, 221-237.e13.

Gundem, G., Van Loo, P., Kremeyer, B., Alexandrov, L.B., Tubio, J.M.C., Papaemmanuil, E., Brewer, D.S., Kallio, H.M.L., Högnäs, G., Annala, M., Kivinummi, K., Goody, V., Latimer, C., O'Meara, S., Dawson, K.J., Isaacs, W., Emmert-Buck, M.R., Nykter, M., Foster, C., Kote-Jarai, Z., Easton, D., Whitaker, H.C., ICGC Prostate Group, Neal, D.E., Cooper, C.S., Eeles, R.A., Visakorpi, T., Campbell, P.J., McDermott, U., Wedge, D.C., and Bova, G.S. (2015). The evolutionary history of lethal metastatic prostate cancer. Nature *520*, 353–357.

Haas, L., Elewaut, A., Gerard, C.L., Umkehrer, C., Leiendecker, L., Pedersen, M., Krecioch, I., Hoffmann, D., Novatchkova, M., Kuttke, M., Neumann, T., da Silva, I.P., Witthock, H., Cuendet, M.A., Carotta, S., Harrington, K.J., Zuber, J., Scolyer, R.A., Long, G.V., Wilmott, J.S., Michielin, O., Vanharanta, S., Wiesner, T., and Obenauf, A.C. (2021). Acquired resistance to anti-MAPK targeted therapy confers an immune-evasive tumor microenvironment and cross-resistance to immunotherapy in melanoma. Nat Cancer *2*, 693–708.

Haderk, F., Fernández-Méndez, C., Čech, L., Yu, J., Meraz, I.M., Olivas, V., Rabago, D.B., Kerr, D.L., Gomez, C., Allegakoen, D.V., Guan, J., Shah, K.N., Herrington, K.A., Gbenedio, O.M., Nanjo, S., Majidi, M., Tamaki, W., Rotow, J.K., McCoach, C.E., Riess, J.W., Gutkind, J.S., Tang, T.T., Post, L., Huang, B., Santisteban, P., Goodarzi, H., Bandyopadhyay, S., Kuo, C.J., Roose, J.P., Wu, W., Blakely, C.M., Roth, J.A., and Bivona, T.G. (2021). A focal adhesion kinase-YAP signaling axis drives drug tolerant persister cells and residual disease in lung cancer. 2021.10.23.465573.

Hall, A. (1984). Oncogenes--implications for human cancer: a review. J R Soc Med *77*, 410–416.

Hanahan, D. (2022). Hallmarks of Cancer: New Dimensions. Cancer Discovery *12*, 31–46.

Hangauer, M.J., Viswanathan, V.S., Ryan, M.J., Bole, D., Eaton, J.K., Matov, A., Galeas, J., Dhruv, H.D., Berens, M.E., Schreiber, S.L., McCormick, F., and McManus, M.T. (2017). Drug-tolerant persister cancer cells are vulnerable to GPX4 inhibition. Nature *551*, 247–250.

Hanna, N.H., Schneider, B.J., Temin, S., Baker, S., Brahmer, J., Ellis, P.M., Gaspar, L.E., Haddad, R.Y., Hesketh, P.J., Jain, D., Jaiyesimi, I., Johnson, D.H., Leighl, N.B., Phillips, T., Riely, G.J., Robinson, A.G., Rosell, R., Schiller, J.H., Singh, N., Spigel, D.R., Stabler, J.O., Tashbar, J., and Masters, G. (2020). Therapy for Stage IV Non–Small-Cell Lung Cancer Without Driver Alterations: ASCO and OH (CCO) Joint Guideline Update. JCO *38*, 1608–1632.

Hansford, S., and Huntsman, D.G. (2014). Boveri at 100: Theodor Boveri and genetic predisposition to cancer. The Journal of Pathology *234*, 142–145.

Hata, A.N., Niederst, M.J., Archibald, H.L., Gomez-Caraballo, M., Siddiqui, F.M., Mulvey, H.E., Maruvka, Y.E., Ji, F., Bhang, H.C., Krishnamurthy Radhakrishna, V., Siravegna, G., Hu, H., Raoof, S., Lockerman, E., Kalsy, A., Lee, D., Keating, C.L., Ruddy, D.A., Damon, L.J., Crystal, A.S., Costa, C., Piotrowska, Z., Bardelli, A., Iafrate, A.J., Sadreyev, R.I., Stegmeier, F., Getz, G., Sequist, L.V., Faber, A.C., and Engelman, J.A. (2016). Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition. Nat. Med. *22*, 262–269.

Herbst, R.S., Baas, P., Kim, D.-W., Felip, E., Pérez-Gracia, J.L., Han, J.-Y., Molina, J., Kim, J.-H., Arvis, C.D., Ahn, M.-J., Majem, M., Fidler, M.J., de Castro, G., Garrido, M., Lubiniecki, G.M., Shentu, Y., Im, E., Dolled-Filhart, M., and Garon, E.B. (2016). Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. Lancet *387*, 1540–1550.

Herbst, R.S., Morgensztern, D., and Boshoff, C. (2018). The biology and management of non-small cell lung cancer. Nature *553*, 446–454.

Holbeck, S.L., Collins, J.M., and Doroshow, J.H. (2010). Analysis of Food and Drug Administration-approved anticancer agents in the NCI60 panel of human tumor cell lines. Mol Cancer Ther *9*, 1451–1460.

Hong, D.S., Fakih, M.G., Strickler, J.H., Desai, J., Durm, G.A., Shapiro, G.I., Falchook, G.S., Price, T.J., Sacher, A., Denlinger, C.S., Bang, Y.-J., Dy, G.K., Krauss, J.C., Kuboki, Y., Kuo, J.C., Coveler, A.L., Park, K., Kim, T.W., Barlesi, F., Munster, P.N., Ramalingam, S.S., Burns, T.F., Meric-Bernstam, F., Henary, H., Ngang, J., Ngarmchamnanrith, G., Kim, J., Houk, B.E., Canon, J., Lipford, J.R., Friberg, G., Lito, P., Govindan, R., and Li, B.T. (2020). KRASG12C Inhibition with Sotorasib in Advanced Solid Tumors. New England Journal of Medicine *383*, 1207–1217.

Howington, J.A., Blum, M.G., Chang, A.C., Balekian, A.A., and Murthy, S.C. (2013). Treatment of Stage I and II Non-small Cell Lung Cancer: Diagnosis and Management of Lung Cancer, 3rd ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. CHEST *143*, e278S-e313S.

Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and Applications of CRISPR-Cas9 for Genome Engineering. Cell *157*, 1262–1278.

Huang, L., and Fu, L. (2015). Mechanisms of resistance to EGFR tyrosine kinase inhibitors. Acta Pharmaceutica Sinica B *5*, 390–401.

Jänne, P.A., Yang, J.C.-H., Kim, D.-W., Planchard, D., Ohe, Y., Ramalingam, S.S., Ahn, M.-J., Kim, S.-W., Su, W.-C., Horn, L., Haggstrom, D., Felip, E., Kim, J.-H., Frewer, P., Cantarini, M., Brown, K.H., Dickinson, P.A., Ghiorghiu, S., and Ranson, M. (2015a). AZD9291 in EGFR Inhibitor–Resistant Non–Small-Cell Lung Cancer. New England Journal of Medicine *372*, 1689–1699.

Jänne, P.A., Ahn, M.-J., Kim, D.-W., Kim, S.-W., Planchard, D., Ramalingam, S.S., Frewer, P., Cantarini, M., Ghiorghiu, S., and Yang, J.C.-H. (2015b). A Phase I Study of AZD9291 in Patients with Egfr-Tki-Resistant Advanced Nsclc – Updated Progression Free Survival and Duration of Response Data. Annals of Oncology *26*, i57.

Jia, Y., Yun, C.-H., Park, E., Ercan, D., Manuia, M., Juarez, J., Xu, C., Rhee, K., Chen, T., Zhang, H., Palakurthi, S., Jang, J., Lelais, G., DiDonato, M., Bursulaya, B., Michellys, P.-Y., Epple, R., Marsilje, T.H., McNeill, M., Lu, W., Harris, J., Bender, S., Wong, K.-K., Jänne, P.A., and Eck, M.J. (2016). Overcoming EGFR(T790M) and EGFR(C797S) resistance with mutant-selective allosteric inhibitors. Nature *534*, 129–132.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science *337*, 816–821.

Joung, J., Konermann, S., Gootenberg, J.S., Abudayyeh, O.O., Platt, R.J., Brigham, M.D., Sanjana, N.E., and Zhang, F. (2017). Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. Nat Protoc *12*, 828–863.

Junker, J.P., Spanjaard, B., Peterson-Maduro, J., Alemany, A., Hu, B., Florescu, M., and van Oudenaarden, A. (2016). Massively parallel clonal analysis using CRISPR/Cas9 induced genetic scars (Systems Biology).

Kalhor, R., Mali, P., and Church, G.M. (2017). Rapidly evolving homing CRISPR barcodes. Nat Methods *14*, 195–200.

Kalhor, R., Kalhor, K., Mejia, L., Leeper, K., Graveline, A., Mali, P., and Church, G.M. (2018). Developmental barcoding of whole mouse via homing CRISPR. Science *361*, eaat9804.

Karachaliou, N., Mayo, C., Costa, C., Magrí, I., Gimenez-Capitan, A., Molina-Vila, M.A., and Rosell, R. (2013). KRAS Mutations in Lung Cancer. Clinical Lung Cancer *14*, 205–214.

Karoulia, Z., Gavathiotis, E., and Poulikakos, P.I. (2017). New perspectives for targeting RAF kinase in human cancer. Nat Rev Cancer *17*, 676–691.

Katakami, N., Atagi, S., Goto, K., Hida, T., Horai, T., Inoue, A., Ichinose, Y., Koboyashi, K., Takeda, K., Kiura, K., Nishio, K., Seki, Y., Ebisawa, R., Shahidi, M., and Yamamoto, N. (2013). LUX-Lung 4: a phase II trial of afatinib in patients with advanced non-small-cell lung cancer who progressed during prior treatment with erlotinib, gefitinib, or both. J Clin Oncol *31*, 3335–3341.

Kate, S., Chougule, A., Joshi, A., Noronha, V., Patil, V., Dusane, R., Solanki, L., Tiwrekar, P., Trivedi, V., and Prabhash, K. (2019). Outcome of uncommon EGFR mutation positive newly diagnosed advanced non-small cell lung cancer patients: a single center retrospective analysis. Lung Cancer (Auckl) *10*, 1–10.

Kebschull, J.M., and Zador, A.M. (2018). Cellular barcoding: lineage tracing, screening and beyond. Nat Methods *15*, 871–879.

Kenfield, S.A., Wei, E.K., Stampfer, M.J., Rosner, B.A., and Colditz, G.A. (2008). Comparison of Aspects of Smoking Among Four Histologic Types of Lung Cancer. Tob Control *17*, 198–204.

Kis, Z., Pereira, H.S., Homma, T., Pedrigi, R.M., and Krams, R. (2015). Mammalian synthetic biology: emerging medical applications. J R Soc Interface *12*, 20141000.

Kobayashi, S., Canepa, H.M., Bailey, A.S., Nakayama, S., Yamaguchi, N., Goldstein, M.A., Huberman, M.S., and Costa, D.B. (2013). Compound EGFR mutations and response to EGFR tyrosine kinase inhibitors. J Thorac Oncol *8*, 45–51.

Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., Nureki, O., and Zhang, F. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature *517*, 583–588.

Kreso, A., and Dick, J.E. (2014). Evolution of the Cancer Stem Cell Model. Cell Stem Cell *14*, 275–291.

Kretzschmar, K., and Watt, F.M. (2012). Lineage tracing. Cell *148*, 33–45.

Kurppa, K.J., Liu, Y., To, C., Zhang, T., Fan, M., Vajdi, A., Knelson, E.H., Xie, Y., Lim, K., Cejas, P., Portell, A., Lizotte, P.H., Ficarro, S.B., Li, S., Chen, T., Haikala, H.M., Wang, H., Bahcall, M., Gao, Y., Shalhout, S., Boettcher, S., Shin, B.H., Thai, T., Wilkens, M.K., Tillgren, M.L., Mushajiang, M., Xu, M., Choi, J., Bertram, A.A., Ebert, B.L., Beroukhim, R., Bandopadhayay, P., Awad, M.M., Gokhale, P.C., Kirschmeier, P.T., Marto, J.A., Camargo, F.D., Haq, R., Paweletz, C.P., et al. (2020). Treatment-Induced Tumor Dormancy through YAP-Mediated Transcriptional Reprogramming of the Apoptotic Pathway. Cancer Cell *37*, 104-122.e12.

Kwaks, T.H.J., Barnett, P., Hemrika, W., Siersma, T., Sewalt, R.G.A.B., Satijn, D.P.E., Brons, J.F., van Blokland, R., Kwakman, P., Kruckeberg, A.L., Kelder, A., and Otte, A.P. (2003). Identification of anti-repressor elements that confer high and stable protein production in mammalian cells. Nat Biotechnol *21*, 553–558.

Kyrchanova, O., Chetverina, D., Maksimenko, O., Kullyev, A., and Georgiev, P. (2008). Orientation-dependent interaction between Drosophila insulators is a property of this class of regulatory elements. Nucleic Acids Res *36*, 7019–7028.

Lan, X., Jörg, D.J., Cavalli, F.M.G., Richards, L.M., Nguyen, L.V., Vanner, R.J., Guilhamon, P., Lee, L., Kushida, M.M., Pellacani, D., Park, N.I., Coutinho, F.J., Whetstone, H., Selvadurai, H.J., Che, C., Luu, B., Carles, A., Moksa, M., Rastegar, N., Head, R., Dolma, S., Prinos, P., Cusimano, M.D., Das, S., Bernstein, M., Arrowsmith, C.H., Mungall, A.J., Moore, R.A., Ma, Y., Gallo, M., Lupien, M., Pugh, T.J., Taylor, M.D., Hirst, M., Eaves, C.J., Simons, B.D., and Dirks, P.B. (2017). Fate mapping of human glioblastoma reveals an invariant stem cell hierarchy. Nature *549*, 227–232.

Lee, C.K., Wu, Y.-L., Ding, P.N., Lord, S.J., Inoue, A., Zhou, C., Mitsudomi, T., Rosell, R., Pavlakis, N., Links, M., Gebski, V., Gralla, R.J., and Yang, J.C.-H. (2015). Impact of Specific Epidermal Growth Factor Receptor (EGFR) Mutations and Clinical Characteristics on Outcomes After Treatment With EGFR Tyrosine Kinase Inhibitors Versus Chemotherapy in EGFR-Mutant Lung Cancer: A Meta-Analysis. J Clin Oncol *33*, 1958–1965.

Li, C., Heidt, D.G., Dalerba, P., Burant, C.F., Zhang, L., Adsay, V., Wicha, M., Clarke, M.F., and Simeone, D.M. (2007). Identification of Pancreatic Cancer Stem Cells. Cancer Res *67*, 1030–1037.

Liau, B.B., Sievers, C., Donohue, L.K., Gillespie, S.M., Flavahan, W.A., Miller, T.E., Venteicher, A.S., Hebert, C.H., Carey, C.D., Rodig, S.J., Shareef, S.J., Najm, F.J., van Galen, P., Wakimoto, H., Cahill, D.P., Rich, J.N., Aster, J.C., Suvà, M.L., Patel, A.P., and Bernstein, B.E. (2017). Adaptive Chromatin Remodeling Drives Glioblastoma Stem Cell Plasticity and Drug Tolerance. Cell Stem Cell *20*, 233-246.e7.

Lim, S.M., Kim, D.-W., Jung, J.E., Lee, G., Ryou, J.-H., Kang, S.-U., Lee, Y.-H., Shin, H.-J., Yum, S.Y., Yim, E., Lee, S.-Y., and Ahn, J.S. (2021). 1365TiP A phase I/II, open-label study of BBT-176, a triple mutation targeting EGFR TKI, in patients with NSCLC who progressed after prior EGFR TKI therapy. Annals of Oncology *32*, S1035.

Lin, J.J., and Shaw, A.T. (2017). Recent Advances in Targeting ROS1 in Lung Cancer. Journal of Thoracic Oncology *12*, 1611–1625.

Lin, C.-C., Shih, J.-Y., Yu, C.-J., Ho, C.-C., Liao, W.-Y., Lee, J.-H., Tsai, T.-H., Su, K.-Y., Hsieh, M.-S., Chang, Y.-L., Bai, Y.-Y., Huang, D.D.-R., Thress, K.S., and Yang, J.C.-H. (2018). Outcomes in patients with non-small-cell lung cancer and acquired Thr790Met mutation treated with osimertinib: a genomic study. The Lancet Respiratory Medicine *6*, 107–116.

Lin, J.J., Riely, G.J., and Shaw, A.T. (2017). Targeting ALK: Precision Medicine Takes on Drug Resistance. Cancer Discov *7*, 137–155.

Liu, X., Zhang, X., Yang, L., Tian, X., Dong, T., Ding, C.Z., Hu, L., Wu, L., Zhao, L., Mao, J., Ji, Q., Yan, S., Zhu, Z., Xia, Y., Chan, C., and Chen, S. (2019). Abstract 1320: Preclinical evaluation of TQB3804, a potent EGFR C797S inhibitor. Cancer Res *79*, 1320–1320.

Livet, J., Weissman, T.A., Kang, H., Draft, R.W., Lu, J., Bennis, R.A., Sanes, J.R., and Lichtman, J.W. (2007). Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. Nature *450*, 56–62.

Malhotra, J., Malvezzi, M., Negri, E., Vecchia, C.L., and Boffetta, P. (2016). Risk factors for lung cancer worldwide. European Respiratory Journal *48*, 889–902.

Marine, J.-C., Dawson, S.-J., and Dawson, M.A. (2020). Non-genetic mechanisms of therapeutic resistance in cancer. Nat Rev Cancer *20*, 743–756.

Martincorena, I., and Campbell, P.J. (2015). Somatic mutation in cancer and normal cells. Science *349*, 1483–1489.

Martins, F., Sofiya, L., Sykiotis, G.P., Lamine, F., Maillard, M., Fraga, M., Shabafrouz, K., Ribi, C., Cairoli, A., Guex-Crosier, Y., Kuntzer, T., Michielin, O., Peters, S., Coukos, G., Spertini, F., Thompson, J.A., and Obeid, M. (2019). Adverse effects of immune-checkpoint inhibitors: epidemiology, management and surveillance. Nat Rev Clin Oncol *16*, 563–580.

Marusyk, A., Almendro, V., and Polyak, K. (2012). Intra-tumour heterogeneity: a looking glass for cancer? Nat Rev Cancer *12*, 323–334.

Mathieu, L.N., Larkins, E., Akinboro, O., Roy, P., Amatya, A.K., Fiero, M.H., Mishra-Kalyani, P.S., Helms, W.S., Myers, C.E., Skinner, A.M., Aungst, S., Jin, R., Zhao, H., Xia, H., Zirkelbach, J.F., Bi, Y., Li, Y., Liu, J., Grimstein, M., Zhang, X., Woods, S., Reece, K., Abukhdeir, A.M., Ghosh, S., Philip, R., Tang, S., Goldberg, K.B., Pazdur, R., Beaver, J.A., and Singh, H. (2022). FDA Approval Summary: Capmatinib and Tepotinib for the Treatment of Metastatic NSCLC Harboring MET Exon 14 Skipping Mutations or Alterations. Clin Cancer Res *28*, 249–254.

Maynard, A., McCoach, C.E., Rotow, J.K., Harris, L., Haderk, F., Kerr, D.L., Yu, E.A., Schenk, E.L., Tan, W., Zee, A., Tan, M., Gui, P., Lea, T., Wu, W., Urisman, A., Jones, K., Sit, R., Kolli, P.K., Seeley, E., Gesthalter, Y., Le, D.D., Yamauchi, K.A., Naeger, D.M., Bandyopadhyay, S., Shah, K., Cech, L., Thomas, N.J., Gupta, A., Gonzalez, M., Do, H., Tan, L., Bacaltos, B., Gomez-Sjoberg, R., Gubens, M., Jahan, T., Kratz, J.R., Jablons, D., Neff, N., Doebele, R.C., et al. (2020). Therapy-Induced Evolution of Human Lung Cancer Revealed by Single-Cell RNA Sequencing. Cell *182*, 1232-1251.e22.

McGranahan, N., and Swanton, C. (2015). Biological and Therapeutic Impact of Intratumor Heterogeneity in Cancer Evolution. Cancer Cell *27*, 15–26.

McKenna, A., Findlay, G.M., Gagnon, J.A., Horwitz, M.S., Schier, A.F., and Shendure, J. (2016). Whole-organism lineage tracing by combinatorial and cumulative genome editing. Science *353*, aaf7907.

McPherson, A., Roth, A., Laks, E., Masud, T., Bashashati, A., Zhang, A.W., Ha, G., Biele, J., Yap, D., Wan, A., Prentice, L.M., Khattra, J., Smith, M.A., Nielsen, C.B., Mullaly, S.C., Kalloger, S., Karnezis, A., Shumansky, K., Siu, C., Rosner, J., Chan, H.L., Ho, J., Melnyk, N., Senz, J., Yang, W., Moore, R., Mungall, A.J., Marra, M.A., Bouchard-Côté, A., Gilks, C.B., Huntsman, D.G., McAlpine, J.N., Aparicio, S., and Shah, S.P. (2016). Divergent modes of clonal spread and intraperitoneal mixing in high-grade serous ovarian cancer. Nature Genetics *48*, 758–767.

Merlo, L.M.F., Pepper, J.W., Reid, B.J., and Maley, C.C. (2006). Cancer as an evolutionary and ecological process. Nat Rev Cancer *6*, 924–935.

Mikubo, M., Inoue, Y., Liu, G., and Tsao, M.-S. (2021). Mechanism of Drug Tolerant Persister Cancer Cells: The Landscape and Clinical Implication for Therapy. J Thorac Oncol *16*, 1798–1809.

Miller, V.A., Hirsh, V., Cadranel, J., Chen, Y.-M., Park, K., Kim, S.-W., Zhou, C., Su, W.-C., Wang, M., Sun, Y., Heo, D.S., Crino, L., Tan, E.-H., Chao, T.-Y., Shahidi, M., Cong, X.J., Lorence, R.M., and Yang, J.C.-H. (2012). Afatinib versus placebo for patients with advanced, metastatic non-small-cell lung cancer after failure of erlotinib, gefitinib, or both, and one or two lines of chemotherapy (LUX-Lung 1): a phase 2b/3 randomised trial. The Lancet Oncology *13*, 528–538.

Milone, M.C., and O'Doherty, U. (2018). Clinical use of lentiviral vectors. Leukemia *32*, 1529–1541.

Mitsudomi, T., Morita, S., Yatabe, Y., Negoro, S., Okamoto, I., Tsurutani, J., Seto, T., Satouchi, M., Tada, H., Hirashima, T., Asami, K., Katakami, N., Takada, M., Yoshioka, H., Shibata, K., Kudoh, S., Shimizu, E., Saito, H., Toyooka, S., Nakagawa, K., Fukuoka, M., and West Japan Oncology Group (2010). Gefitinib versus cisplatin plus docetaxel in patients with non-small-cell lung cancer harbouring mutations of the epidermal growth factor receptor (WJTOG3405): an open label, randomised phase 3 trial. Lancet Oncol *11*, 121–128.

Mok, T.S., Wu, Y.-L., Ahn, M.-J., Garassino, M.C., Kim, H.R., Ramalingam, S.S., Shepherd, F.A., He, Y., Akamatsu, H., Theelen, W.S.M.E., Lee, C.K., Sebastian, M., Templeton, A., Mann, H., Marotti, M., Ghiorghiu, S., Papadimitrakopoulou, V.A., and AURA3 Investigators (2017). Osimertinib or Platinum-Pemetrexed in EGFR T790M-Positive Lung Cancer. N Engl J Med *376*, 629–640.

Moore, A.R., Rosenberg, S.C., McCormick, F., and Malek, S. (2020). RAS-targeted therapies: is the undruggable drugged? Nat Rev Drug Discov *19*, 533–552.

Morgillo, F., Della Corte, C.M., Fasano, M., and Ciardiello, F. (2016). Mechanisms of resistance to EGFR-targeted drugs: lung cancer. ESMO Open *1*, e000060.

Moris, N., Pina, C., and Arias, A.M. (2016). Transition states and cell fate decisions in epigenetic landscapes. Nat Rev Genet *17*, 693–703.

Mullick, A., Xu, Y., Warren, R., Koutroumanis, M., Guilbault, C., Broussau, S., Malenfant, F., Bourget, L., Lamoureux, L., Lo, R., Caron, A.W., Pilotte, A., and Massie, B. (2006). The cumate gene-switch: a system for regulated expression in mammalian cells. BMC Biotechnology *6*, 43.

Murray, P.B., Lax, I., Reshetnyak, A., Ligon, G.F., Lillquist, J.S., Natoli, E.J., Shi, X., Folta-Stogniew, E., Gunel, M., Alvarado, D., and Schlessinger, J. (2015). Heparin is an activating ligand of the orphan receptor tyrosine kinase ALK. Sci Signal *8*, ra6.

Naik, S.H., Schumacher, T.N., and Perié, L. (2014). Cellular barcoding: a technical appraisal. Exp Hematol *42*, 598–608.

Navarria, P., Ascolese, A.M., Mancosu, P., Alongi, F., Clerici, E., Tozzi, A., Iftode, C., Reggiori, G., Tomatis, S., Infante, M., Alloisio, M., Testori, A., Fogliata, A., Cozzi, L., Morenghi, E., and Scorsetti, M. (2013). Volumetric modulated arc therapy with flattening filter free (FFF) beams for stereotactic body radiation therapy (SBRT) in patients with medically inoperable early stage non small cell lung cancer (NSCLC). Radiotherapy and Oncology *107*, 414–418.

Nguyen, L.V., Cox, C.L., Eirew, P., Knapp, D.J.H.F., Pellacani, D., Kannan, N., Carles, A., Moksa, M., Balani, S., Shah, S., Hirst, M., Aparicio, S., and Eaves, C.J. (2014). DNA barcoding reveals diverse growth kinetics of human breast tumour subclones in serially passaged xenografts. Nat Commun *5*, 5871.

Nguyen, L.V., Pellacani, D., Lefort, S., Kannan, N., Osako, T., Makarem, M., Cox, C.L., Kennedy, W., Beer, P., Carles, A., Moksa, M., Bilenky, M., Balani, S., Babovic, S., Sun, I., Rosin, M., Aparicio, S., Hirst, M., and Eaves, C.J. (2015). Barcoding reveals complex clonal dynamics of de novo transformed human mammary cells. Nature *528*, 267–271.

Niederst, M.J., Hu, H., Mulvey, H.E., Lockerman, E.L., Garcia, A.R., Piotrowska, Z., Sequist, L.V., and Engelman, J.A. (2015). The Allelic Context of the C797S Mutation Acquired upon Treatment with Third-Generation EGFR Inhibitors Impacts Sensitivity to Subsequent Treatment Strategies. Clin Cancer Res *21*, 3924–3933.

Noronha, V., Patil, V.M., Joshi, A., Menon, N., Chougule, A., Mahajan, A., Janu, A., Purandare, N., Kumar, R., More, S., Goud, S., Kadam, N., Daware, N., Bhattacharjee, A., Shah, S., Yadav, A., Trivedi, V., Behel, V., Dutt, A., Banavali, S.D., and Prabhash, K. (2020). Gefitinib Versus Gefitinib Plus Pemetrexed and Carboplatin Chemotherapy in EGFR-Mutated Lung Cancer. J Clin Oncol *38*, 124–136.

Novello, S., Barlesi, F., Califano, R., Cufer, T., Ekman, S., Levra, M.G., Kerr, K., Popat, S., Reck, M., Senan, S., Simo, G.V., Vansteenkiste, J., Peters, S., and ESMO Guidelines Committee (2016). Metastatic non-small-cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol *27*, v1–v27.

Nowell, P.C. (1976). The clonal evolution of tumor cell populations. Science *194*, 23–28.

O'Brien, C.A., Pollett, A., Gallinger, S., and Dick, J.E. (2007). A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. Nature *445*, 106–110.

Oren, Y., Tsabar, M., Cuoco, M.S., Amir-Zilberstein, L., Cabanos, H.F., Hütter, J.-C., Hu, B., Thakore, P.I., Tabaka, M., Fulco, C.P., Colgan, W., Cuevas, B.M., Hurvitz, S.A., Slamon, D.J., Deik, A., Pierce, K.A., Clish, C., Hata, A.N., Zaganjor, E., Lahav, G., Politi, K., Brugge, J.S., and Regev, A. (2021). Cycling cancer persister cells arise from lineages with distinct programs. Nature *596*, 576–582.

Ortiz, M.A., Mikhailova, T., Li, X., Porter, B.A., Bah, A., and Kotula, L. (2021). Src family kinases, adaptor proteins and the actin cytoskeleton in epithelial-to-mesenchymal transition. Cell Communication and Signaling *19*, 67.

Oxnard, G.R., Hu, Y., Mileham, K.F., Husain, H., Costa, D.B., Tracy, P., Feeney, N., Sholl, L.M., Dahlberg, S.E., Redig, A.J., Kwiatkowski, D.J., Rabin, M.S., Paweletz, C.P., Thress, K.S., and Jänne, P.A. (2018). Assessment of Resistance Mechanisms and Clinical Implications in Patients With EGFR T790M-Positive Lung Cancer and Acquired Resistance to Osimertinib. JAMA Oncol *4*, 1527–1534.

Oxnard, G.R., Yang, J.C.-H., Yu, H., Kim, S.-W., Saka, H., Horn, L., Goto, K., Ohe, Y., Mann, H., Thress, K.S., Frigault, M.M., Vishwanathan, K., Ghiorghiu, D., Ramalingam, S.S., and Ahn, M.-J. (2020). TATTON: a multi-arm, phase Ib trial of osimertinib combined with selumetinib, savolitinib, or durvalumab in EGFR-mutant lung cancer. Annals of Oncology *31*, 507–516.

Paddison, P.J., Silva, J.M., Conklin, D.S., Schlabach, M., Li, M., Aruleba, S., Balija, V., O'Shaughnessy, A., Gnoj, L., Scobie, K., Chang, K., Westbrook, T., Cleary, M., Sachidanandam, R., McCombie, W.R., Elledge, S.J., and Hannon, G.J. (2004). A resource for large-scale RNA-interference-based screens in mammals. Nature *428*, 427–431.

Paik, P.K., Felip, E., Veillon, R., Sakai, H., Cortot, A.B., Garassino, M.C., Mazieres, J., Viteri, S., Senellart, H., Van Meerbeeck, J., Raskin, J., Reinmuth, N., Conte, P., Kowalski, D., Cho, B.C., Patel, J.D., Horn, L., Griesinger, F., Han, J.-Y., Kim, Y.-C., Chang, G.-C., Tsai, C.-L., Yang, J.C.-H., Chen, Y.-M., Smit, E.F., van der Wekken, A.J., Kato, T., Juraeva, D., Stroh, C., Bruns, R., Straub, J., Johne, A., Scheele, J., Heymach, J.V., and Le, X. (2020). Tepotinib in Non–Small-Cell Lung Cancer with MET Exon 14 Skipping Mutations. New England Journal of Medicine *383*, 931–943.

Pao, W., Miller, V.A., Politi, K.A., Riely, G.J., Somwar, R., Zakowski, M.F., Kris, M.G., and Varmus, H. (2005). Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. PLoS Med *2*, e73.

Papadimitrakopoulou, V.A., Wu, Y.-L., Han, J.-Y., Ahn, M.-J., Ramalingam, S.S., John, T., Okamoto, I., Yang, J.C.-H., Bulusu, K.C., Laus, G., Collins, B., Barrett, J.C., Chmielecki, J., and Mok, T.S.K. (2018). Analysis of resistance mechanisms to osimertinib in patients with EGFR T790M advanced NSCLC from the AURA3 study. Annals of Oncology *29*, viii741.

Passaro, A., Jänne, P.A., Mok, T., and Peters, S. (2021). Overcoming therapy resistance in EGFR-mutant lung cancer. Nat Cancer *2*, 377–391.

Pauklin, S., and Vallier, L. (2013). The cell-cycle state of stem cells determines cell fate propensity. Cell *155*, 135–147.

Pauls, S., Geldmacher-Voss, B., and Campos-Ortega, J.A. (2001). A zebrafish histone variant H2A.F/Z and a transgenic H2A.F/Z:GFP fusion protein for in vivo studies of embryonic development. Dev Genes Evol *211*, 603–610.

Pei, W., Feyerabend, T.B., Rössler, J., Wang, X., Postrach, D., Busch, K., Rode, I., Klapproth, K., Dietlein, N., Quedenau, C., Chen, W., Sauer, S., Wolf, S., Höfer, T., and Rodewald, H.-R. (2017). Polylox barcoding reveals haematopoietic stem cell fates realized in vivo. Nature *548*, 456–460.

Perli, S.D., Cui, C.H., and Lu, T.K. (2016). Continuous genetic recording with self-targeting CRISPR-Cas in human cells. Science *353*, aag0511.

Peschard, P., Fournier, T.M., Lamorte, L., Naujokas, M.A., Band, H., Langdon, W.Y., and Park, M. (2001). Mutation of the c-Cbl TKB domain binding site on the Met receptor tyrosine kinase converts it into a transforming protein. Mol Cell *8*, 995–1004.

Piotrowska, Z., Isozaki, H., Lennerz, J.K., Gainor, J.F., Lennes, I.T., Zhu, V.W., Marcoux, N., Banwait, M.K., Digumarthy, S.R., Su, W., Yoda, S., Riley, A.K., Nangia, V., Lin, J.J., Nagy, R.J., Lanman, R.B., Dias-Santagata, D., Mino-Kenudson, M., Iafrate, A.J., Heist, R.S., Shaw, A.T., Evans, E.K., Clifford, C., Ou, S.-H.I., Wolf, B., Hata, A.N., and Sequist, L.V. (2018). Landscape of Acquired Resistance to Osimertinib in EGFR-Mutant NSCLC and Clinical Validation of Combined EGFR and RET Inhibition with Osimertinib and BLU-667 for Acquired RET Fusion. Cancer Discov *8*, 1529–1539.

Pisco, A.O., Brock, A., Zhou, J., Moor, A., Mojtahedi, M., Jackson, D., and Huang, S. (2013). Non-Darwinian dynamics in therapy-induced cancer drug resistance. Nat Commun *4*, 2467.

Planchard, D., Feng, P.-H., Karaseva, N., Kim, S.-W., Kim, T.M., Lee, C.K., Poltoratskiy, A., Yanagitani, N., Marshall, R., Huang, X., Howarth, P., Jänne, P.A., and Kobayashi, K. (2021). Osimertinib plus platinum-pemetrexed in newly diagnosed epidermal growth factor receptor mutation-positive advanced/metastatic non-small-cell lung cancer: safety run-in results from the FLAURA2 study. ESMO Open *6*, 100271.

Prasetyanti, P.R., and Medema, J.P. (2017). Intra-tumor heterogeneity from a cancer stem cell perspective. Molecular Cancer *16*, 41.

Pratilas, C.A., Taylor, B.S., Ye, Q., Viale, A., Sander, C., Solit, D.B., and Rosen, N. (2009). V600EBRAF is associated with disabled feedback inhibition of RAF–MEK signaling and elevated transcriptional output of the pathway. Proceedings of the National Academy of Sciences *106*, 4519–4524.

Qu, X., Liu, H., Song, X., Sun, N., Zhong, H., Qiu, X., Yang, X., and Jiang, B. (2021). Effective degradation of EGFRL858R+T790M mutant proteins by CRBN-based PROTACs through both proteosome and autophagy/lysosome degradation systems. European Journal of Medicinal Chemistry *218*, 113328.

Raha, D., Wilson, T.R., Peng, J., Peterson, D., Yue, P., Evangelista, M., Wilson, C., Merchant, M., and Settleman, J. (2014a). The cancer stem cell marker aldehyde dehydrogenase is required to maintain a drug-tolerant tumor cell subpopulation. Cancer Res *74*, 3579–3590.

Raj, A., and van Oudenaarden, A. (2008). Nature, nurture, or chance: stochastic gene expression and its consequences. Cell *135*, 216–226.

Raj, B., Wagner, D.E., McKenna, A., Pandey, S., Klein, A.M., Shendure, J., Gagnon, J.A., and Schier, A.F. (2018). Simultaneous single-cell profiling of lineages and cell types in the vertebrate brain. Nat Biotechnol *36*, 442–450.

Ramalingam, S.S., Cheng, Y., Zhou, C., Ohe, Y., Imamura, F., Cho, B.C., Lin, M.-C., Majem, M., Shah, R., Rukazenkov, Y., Todd, A., Markovets, A., Barrett, J.C., Chmielecki, J., and Gray, J. (2018). Mechanisms of acquired resistance to first-line osimertinib: Preliminary data from the phase III FLAURA study. Annals of Oncology *29*, viii740.

Ramalingam, S.S., Saka, H., Ahn, M.-J., Yu, H., Yu, H., Horn, L., Hida, T., Cantarini, M., Verheijen, R., Wessen, J., Oxnard, G., and Ohe, Y. (2019). Abstract CT034: Osimertinib plus selumetinib for patients (pts) with EGFR-mutant (EGFRm) NSCLC following disease progression on an EGFR-TKI: Results from the Phase Ib TATTON study. Cancer Research *79*, CT034.

Ramalingam, S.S., Vansteenkiste, J., Planchard, D., Cho, B.C., Gray, J.E., Ohe, Y., Zhou, C., Reungwetwattana, T., Cheng, Y., Chewaskulyong, B., Shah, R., Cobo, M., Lee, K.H., Cheema, P., Tiseo, M., John, T., Lin, M.-C., Imamura, F., Kurata, T., Todd, A., Hodge, R., Saggese, M., Rukazenkov, Y., Soria, J.-C., and FLAURA Investigators (2020). Overall Survival with Osimertinib in Untreated, EGFR-Mutated Advanced NSCLC. N Engl J Med *382*, 41–50.

Rambow, F., Rogiers, A., Marin-Bejar, O., Aibar, S., Femel, J., Dewaele, M., Karras, P., Brown, D., Chang, Y.H., Debiec-Rychter, M., Adriaens, C., Radaelli, E., Wolter, P., Bechter, O., Dummer, R., Levesque, M., Piris, A., Frederick, D.T., Boland, G., Flaherty, K.T., van den Oord, J., Voet, T., Aerts, S., Lund, A.W., and Marine, J.-C. (2018). Toward Minimal Residual Disease-Directed Therapy in Melanoma. Cell *174*, 843-855.e19.

Ramirez, M., Rajaram, S., Steininger, R.J., Osipchuk, D., Roth, M.A., Morinishi, L.S., Evans, L., Ji, W., Hsu, C.-H., Thurley, K., Wei, S., Zhou, A., Koduru, P.R., Posner, B.A., Wu, L.F., and Altschuler, S.J. (2016). Diverse drug-resistance mechanisms can emerge from drug-tolerant cancer persister cells. Nat Commun *7*, 10690.

Ramón Y Cajal, S., Sesé, M., Capdevila, C., Aasen, T., De Mattos-Arruda, L., Diaz-Cano, S.J., Hernández-Losa, J., and Castellví, J. (2020). Clinical implications of intratumor heterogeneity: challenges and opportunities. J Mol Med (Berl) *98*, 161–177.

Reck, M., Spira, A., Besse, B., Wolf, J., Skoulidis, F., Borghaei, H., Goto, K., Park, K., Griesinger, F., Font, E.F., Boyer, M., Barrios, C.H., Goss, G., Yang, H., Obiozor, C., and Ramalingam, S. (2021). MO01.32 CodeBreaK 200: A Phase 3 Multicenter Study of Sotorasib, a KRAS(G12C) Inhibitor, versus Docetaxel in Patients with Previously Treated Advanced Non-Small Cell Lung Cancer (NSCLC) Harboring KRAS p.G12C Mutation. Journal of Thoracic Oncology *16*, S29.

Recondo, G., Che, J., Jänne, P.A., and Awad, M.M. (2020). Targeting MET Dysregulation in Cancer. Cancer Discov *10*, 922–934.

Rehman, S.K., Haynes, J., Collignon, E., Brown, K.R., Wang, Y., Nixon, A.M.L., Bruce, J.P., Wintersinger, J.A., Mer, A.S., Lo, E.B.L., Leung, C., Lima-Fernandes, E., Pedley, N.M., Soares, F., McGibbon, S., He, H.H., Pollet, A., Pugh, T.J., Haibe-Kains, B., Morris, Q., Ramalho-Santos, M., Goyal, S., Moffat, J., and O'Brien, C.A. (2021). Colorectal Cancer Cells Enter a Diapause-like DTP State to Survive Chemotherapy. Cell *184*, 226-242.e21.

Reshetnyak, A.V., Murray, P.B., Shi, X., Mo, E.S., Mohanty, J., Tome, F., Bai, H., Gunel, M., Lax, I., and Schlessinger, J. (2015). Augmentor α and β (FAM150) are ligands of the receptor tyrosine kinases ALK and LTK: Hierarchy and specificity of ligand–receptor interactions. PNAS *112*, 15862–15867.

Riely, G.J., Neal, J.W., Camidge, D.R., Spira, A.I., Piotrowska, Z., Costa, D.B., Tsao, A.S., Patel, J.D., Gadgeel, S.M., Bazhenova, L., Zhu, V.W., West, H.L., Mekhail, T., Gentzler, R.D., Nguyen, D., Vincent, S., Zhang, S., Lin, J., Bunn, V., Jin, S., Li, S., and Jänne, P.A. (2021). Activity and Safety of Mobocertinib (TAK-788) in Previously Treated Non–Small Cell Lung Cancer with EGFR Exon 20 Insertion Mutations from a Phase I/II Trial. Cancer Discov *11*, 1688–1699.

Rikova, K., Guo, A., Zeng, Q., Possemato, A., Yu, J., Haack, H., Nardone, J., Lee, K., Reeves, C., Li, Y., Hu, Y., Tan, Z., Stokes, M., Sullivan, L., Mitchell, J., Wetzel, R., Macneill, J., Ren, J.M., Yuan, J., Bakalarski, C.E., Villen, J., Kornhauser, J.M., Smith, B., Li, D., Zhou, X., Gygi, S.P., Gu, T.-L., Polakiewicz, R.D., Rush, J., and Comb, M.J. (2007). Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. Cell *131*, 1190–1203.

Rizvi, N.A., Hellmann, M.D., Snyder, A., Kvistborg, P., Makarov, V., Havel, J.J., Lee, W., Yuan, J., Wong, P., Ho, T.S., Miller, M.L., Rekhtman, N., Moreira, A.L., Ibrahim, F., Bruggeman, C., Gasmi, B., Zappasodi, R., Maeda, Y., Sander, C., Garon, E.B., Merghoub, T., Wolchok, J.D., Schumacher, T.N., and Chan, T.A. (2015). Mutational landscape determines sensitivity to PD-1 blockade in non–small cell lung cancer. Science *348*, 124–128.

Roskoski, R. (2014). The ErbB/HER family of protein-tyrosine kinases and cancer. Pharmacological Research *79*, 34–74.

Rotow, J., and Bivona, T.G. (2017). Understanding and targeting resistance mechanisms in NSCLC. Nat Rev Cancer *17*, 637–658.

Russo, M., Sogari, A., and Bardelli, A. (2021). Adaptive Evolution: How Bacteria and Cancer Cells Survive Stressful Conditions and Drug Treatment. Cancer Discov *11*, 1886–1895.

Schepers, K., Swart, E., van Heijst, J.W.J., Gerlach, C., Castrucci, M., Sie, D., Heimerikx, M., Velds, A., Kerkhoven, R.M., Arens, R., and Schumacher, T.N.M. (2008). Dissecting T cell lineage relationships by cellular barcoding. J Exp Med *205*, 2309–2318.

Schrock, A.B., Frampton, G.M., Suh, J., Chalmers, Z.R., Rosenzweig, M., Erlich, R.L., Halmos, B., Goldman, J., Forde, P., Leuenberger, K., Peled, N., Kalemkerian, G.P., Ross, J.S., Stephens, P.J., Miller, V.A., Ali, S.M., and Ou, S.-H.I. (2016). Characterization of 298 Patients with Lung Cancer Harboring MET Exon 14 Skipping Alterations. Journal of Thoracic Oncology *11*, 1493–1502.

Segerman, A., Niklasson, M., Haglund, C., Bergström, T., Jarvius, M., Xie, Y., Westermark, A., Sönmez, D., Hermansson, A., Kastemar, M., Naimaie-Ali, Z., Nyberg, F., Berglund, M., Sundström, M., Hesselager, G., Uhrbom, L., Gustafsson, M., Larsson, R., Fryknäs, M., Segerman, B., and Westermark, B. (2016). Clonal Variation in Drug and Radiation Response among Glioma-Initiating Cells Is Linked to Proneural-Mesenchymal Transition. Cell Rep *17*, 2994–3009.

Sehgal, K., Patell, R., Rangachari, D., and Costa, D.B. (2018). Targeting ROS1 rearrangements in non-small cell lung cancer with crizotinib and other kinase inhibitors. Transl Cancer Res *7*, S779–S786.

Sequist, L.V., Waltman, B.A., Dias-Santagata, D., Digumarthy, S., Turke, A.B., Fidias, P., Bergethon, K., Shaw, A.T., Gettinger, S., Cosper, A.K., Akhavanfard, S., Heist, R.S., Temel, J., Christensen, J.G., Wain, J.C., Lynch, T.J., Vernovsky, K., Mark, E.J., Lanuti, M., Iafrate, A.J., Mino-Kenudson, M., and Engelman, J.A. (2011). Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors. Sci Transl Med *3*, 75ra26.

Sequist, L.V., Soria, J.-C., Goldman, J.W., Wakelee, H.A., Gadgeel, S.M., Varga, A., Papadimitrakopoulou, V., Solomon, B.J., Oxnard, G.R., Dziadziuszko, R., Aisner, D.L., Doebele, R.C., Galasso, C., Garon, E.B., Heist, R.S., Logan, J., Neal, J.W., Mendenhall, M.A., Nichols, S., Piotrowska, Z., Wozniak, A.J., Raponi, M., Karlovich, C.A., Jaw-Tsai, S., Isaacson, J., Despain, D., Matheny, S.L., Rolfe, L., Allen, A.R., and Camidge, D.R. (2015). Rociletinib in EGFR-Mutated Non–Small-Cell Lung Cancer. New England Journal of Medicine *372*, 1700–1709.

Shaffer, S.M., Dunagin, M.C., Torborg, S.R., Torre, E.A., Emert, B., Krepler, C., Beqiri, M., Sproesser, K., Brafford, P.A., Xiao, M., Eggan, E., Anastopoulos, I.N., Vargas-Garcia, C.A., Singh, A., Nathanson, K.L., Herlyn, M., and Raj, A. (2017). Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. Nature *546*, 431–435.

Shah, K.N., Bhatt, R., Rotow, J., Rohrberg, J., Olivas, V., Wang, V.E., Hemmati, G., Martins, M.M., Maynard, A., Kuhn, J., Galeas, J., Donnella, H.J., Kaushik, S., Ku, A., Dumont, S., Krings, G., Haringsma, H.J., Robillard, L., Simmons, A.D., Harding, T.C., McCormick, F., Goga, A., Blakely, C.M., Bivona, T.G., and Bandyopadhyay, S. (2019). Aurora kinase A drives the evolution of resistance to third-generation EGFR inhibitors in lung cancer. Nat Med *25*, 111–118.

Shan, Y., Eastwood, M.P., Zhang, X., Kim, E.T., Arkhipov, A., Dror, R.O., Jumper, J., Kuriyan, J., and Shaw, D.E. (2012). Oncogenic Mutations Counteract Intrinsic Disorder in the EGFR Kinase and Promote Receptor Dimerization. Cell *149*, 860–870.

Sharma, P., and Allison, J.P. (2015). Immune Checkpoint Targeting in Cancer Therapy: Towards Combination Strategies with Curative Potential. Cell *161*, 205–214.

Sharma, S.V., Bell, D.W., Settleman, J., and Haber, D.A. (2007). Epidermal growth factor receptor mutations in lung cancer. Nat Rev Cancer *7*, 169–181.

Sharma, S.V., Lee, D.Y., Li, B., Quinlan, M.P., Takahashi, F., Maheswaran, S., McDermott, U., Azizian, N., Zou, L., Fischbach, M.A., Wong, K.-K., Brandstetter, K., Wittner, B., Ramaswamy, S., Classon, M., and Settleman, J. (2010). A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. Cell *141*, 69–80.

Shaw, A.T., and Solomon, B. (2011). Targeting Anaplastic Lymphoma Kinase in Lung Cancer. Clin Cancer Res *17*, 2081–2086.

Shaw, A.T., Ou, S.-H.I., Bang, Y.-J., Camidge, D.R., Solomon, B.J., Salgia, R., Riely, G.J., Varella-Garcia, M., Shapiro, G.I., Costa, D.B., Doebele, R.C., Le, L.P., Zheng, Z., Tan, W., Stephenson, P., Shreeve, S.M., Tye, L.M., Christensen, J.G., Wilner, K.D., Clark, J.W., and Iafrate, A.J. (2014). Crizotinib in ROS1-Rearranged Non–Small-Cell Lung Cancer. New England Journal of Medicine *371*, 1963–1971.

Shaw, A.T., Felip, E., Bauer, T.M., Besse, B., Navarro, A., Postel-Vinay, S., Gainor, J.F., Johnson, M., Dietrich, J., James, L.P., Clancy, J.S., Chen, J., Martini, J.-F., Abbattista, A., and Solomon, B.J. (2017). Lorlatinib in non-small-cell lung cancer with ALK or ROS1 rearrangement: an international, multicentre, open-label, single-arm first-in-man phase 1 trial. Lancet Oncol *18*, 1590–1599.

Shaw, A.T., Riely, G.J., Bang, Y.-J., Kim, D.-W., Camidge, D.R., Solomon, B.J., Varella-Garcia, M., Iafrate, A.J., Shapiro, G.I., Usari, T., Wang, S.C., Wilner, K.D., Clark, J.W., and Ou, S.-H.I. (2019). Crizotinib in ROS1-rearranged advanced non-small-cell lung cancer (NSCLC): updated results, including overall survival, from PROFILE 1001. Ann Oncol *30*, 1121–1126.

Shen, S., Vagner, S., and Robert, C. (2020). Persistent Cancer Cells: The Deadly Survivors. Cell *183*, 860–874.

Shoemaker, R.H. (2006). The NCI60 human tumour cell line anticancer drug screen. Nat Rev Cancer *6*, 813–823.

Simanshu, D.K., Nissley, D.V., and McCormick, F. (2017). RAS Proteins and Their Regulators in Human Disease. Cell *170*, 17–33.

Singh, A., and Settleman, J. (2010). EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer. Oncogene *29*, 4741–4751.

Skoulidis, F., and Heymach, J.V. (2019). Co-occurring genomic alterations in non-small-cell lung cancer biology and therapy. Nat Rev Cancer *19*, 495–509.

Slusarczyk, A.L., Lin, A., and Weiss, R. (2012). Foundations for the design and implementation of synthetic genetic circuits. Nat Rev Genet *13*, 406–420.

Snippert, H.J., van der Flier, L.G., Sato, T., van Es, J.H., van den Born, M., Kroon-Veenboer, C., Barker, N., Klein, A.M., van Rheenen, J., Simons, B.D., and Clevers, H. (2010). Intestinal Crypt Homeostasis Results from Neutral Competition between Symmetrically Dividing Lgr5 Stem Cells. Cell *143*, 134–144.

Soda, M., Choi, Y.L., Enomoto, M., Takada, S., Yamashita, Y., Ishikawa, S., Fujiwara, S., Watanabe, H., Kurashina, K., Hatanaka, H., Bando, M., Ohno, S., Ishikawa, Y., Aburatani, H., Niki, T., Sohara, Y., Sugiyama, Y., and Mano, H. (2007). Identification of the transforming EML4–ALK fusion gene in non-small-cell lung cancer. Nature *448*, 561–566.

Soria, J.-C., Ohe, Y., Vansteenkiste, J., Reungwetwattana, T., Chewaskulyong, B., Lee, K.H., Dechaphunkul, A., Imamura, F., Nogami, N., Kurata, T., Okamoto, I., Zhou, C., Cho, B.C., Cheng, Y., Cho, E.K., Voon, P.J., Planchard, D., Su, W.-C., Gray, J.E., Lee, S.-M., Hodge, R., Marotti, M., Rukazenkov, Y., Ramalingam, S.S., and FLAURA Investigators (2018). Osimertinib in Untreated EGFR-Mutated Advanced Non-Small-Cell Lung Cancer. N Engl J Med *378*, 113–125.

Sottoriva, A., Kang, H., Ma, Z., Graham, T.A., Salomon, M.P., Zhao, J., Marjoram, P., Siegmund, K., Press, M.F., Shibata, D., and Curtis, C. (2015). A Big Bang model of human colorectal tumor growth. Nature Genetics *47*, 209–216.

Spanjaard, B., Hu, B., Mitic, N., Olivares-Chauvet, P., Janjuha, S., Ninov, N., and Junker, J.P. (2018). Simultaneous lineage tracing and cell-type identification using CRISPR–Cas9-induced genetic scars. Nat Biotechnol *36*, 469–473.

Spigel, D., Goto, K., Camidge, D.R., Elamin, Y., Langen, A.J. de, Leighl, N.B., Minchom, A., Piotrowska, Z., Planchard, D., Reckamp, K., Albayya, F., Green, J., Kim, S., Louie-Gao, M., Sawtell, R., Zalutskaya, A., and Co, B.C. (2021). Abstract P230: A phase 1/2 study of BLU-945, a highly potent and selective inhibitor of epidermal growth factor receptor (EGFR) resistance mutations, in patients with EGFR-mutant non-small cell lung cancer (NSCLC). Mol Cancer Ther *20*, P230–P230.

Stratton, M.R., Campbell, P.J., and Futreal, P.A. (2009). The cancer genome. Nature *458*, 719–724.

Su, Y., Wei, W., Robert, L., Xue, M., Tsoi, J., Garcia-Diaz, A., Moreno, B.H., Kim, J., Ng, R.H., Lee, J.W., Koya, R.C., Comin-Anduix, B., Graeber, T.G., Ribas, A., and Heath, J.R. (2017). Single-cell analysis resolves the cell state transition and signaling dynamics associated with melanoma drug-induced resistance. PNAS *114*, 13679–13684.

Sun, C., Wang, L., Huang, S., Heynen, G.J.J.E., Prahallad, A., Robert, C., Haanen, J., Blank, C., Wesseling, J., Willems, S.M., Zecchin, D., Hobor, S., Bajpe, P.K., Lieftink, C., Mateus, C., Vagner, S., Grernrum, W., Hofland, I., Schlicker, A., Wessels, L.F.A., Beijersbergen, R.L., Bardelli, A., Di Nicolantonio, F., Eggermont, A.M.M., and Bernards, R. (2014a). Reversible and adaptive resistance to BRAF(V600E) inhibition in melanoma. Nature *508*, 118–122.

Sun, J., Ramos, A., Chapman, B., Johnnidis, J.B., Le, L., Ho, Y.-J., Klein, A., Hofmann, O., and Camargo, F.D. (2014b). Clonal dynamics of native haematopoiesis. Nature *514*, 322–327.

Swayden, M., Chhouri, H., Anouar, Y., and Grumolato, L. (2020). Tolerant/Persister Cancer Cells and the Path to Resistance to Targeted Therapy. Cells *9*, 2601.

Tabassum, D.P., and Polyak, K. (2015). Tumorigenesis: it takes a village. Nat Rev Cancer *15*, 473–483.

Takezawa, K., Pirazzoli, V., Arcila, M.E., Nebhan, C.A., Song, X., de Stanchina, E., Ohashi, K., Janjigian, Y.Y., Spitzler, P.J., Melnick, M.A., Riely, G.J., Kris, M.G., Miller, V.A., Ladanyi, M., Politi, K., and Pao, W. (2012). HER2 amplification: a potential mechanism of acquired resistance to EGFR inhibition in EGFR-mutant lung cancers that lack the second-site EGFRT790M mutation. Cancer Discov *2*, 922–933.

Tanaka, K., Yu, H.A., Yang, S., Han, S., Selcuklu, S.D., Kim, K., Ramani, S., Ganesan, Y.T., Moyer, A., Sinha, S., Xie, Y., Ishizawa, K., Osmanbeyoglu, H.U., Lyu, Y., Roper, N., Guha, U., Rudin, C.M., Kris, M.G., Hsieh, J.J., and Cheng, E.H. (2021). Targeting Aurora B kinase prevents and overcomes resistance to EGFR inhibitors in lung cancer by enhancing BIM- and PUMA-mediated apoptosis. Cancer Cell *39*, 1245-1261.e6.

Taniguchi, H., Yamada, T., Wang, R., Tanimura, K., Adachi, Y., Nishiyama, A., Tanimoto, A., Takeuchi, S., Araujo, L.H., Boroni, M., Yoshimura, A., Shiotsu, S., Matsumoto, I., Watanabe, S., Kikuchi, T., Miura, S., Tanaka, H., Kitazaki, T., Yamaguchi, H., Mukae, H., Uchino, J., Uehara, H., Takayama, K., and Yano, S. (2019). AXL confers intrinsic resistance to osimertinib and advances the emergence of tolerant cells. Nat Commun *10*, 259.

Thress, K.S., Paweletz, C.P., Felip, E., Cho, B.C., Stetson, D., Dougherty, B., Lai, Z., Markovets, A., Vivancos, A., Kuang, Y., Ercan, D., Matthews, S., Cantarini, M., Barrett, J.C., Jänne, P.A., and Oxnard, G.R. (2015). Acquired EGFR C797S mediates resistance to AZD9291 in advanced non-small cell lung cancer harboring EGFR T790M. Nat Med *21*, 560–562.

Timmerman, R., Paulus, R., Galvin, J., Michalski, J., Straube, W., Bradley, J., Fakiris, A., Bezjak, A., Videtic, G., Johnstone, D., Fowler, J., Gore, E., and Choy, H. (2010). Stereotactic Body Radiation Therapy for Inoperable Early Stage Lung Cancer. JAMA *303*, 1070–1076.

To, C., Beyett, T.S., Jang, J., Feng, W.W., Bahcall, M., Haikala, H.M., Shin, B.H., Heppner, D.E., Rana, J.K., Leeper, B.A., Soroko, K.M., Poitras, M.J., Gokhale, P.C., Kobayashi, Y., Wahid, K., Kurppa, K.J., Gero, T.W., Cameron, M.D., Ogino, A., Mushajiang, M., Xu, C., Zhang, Y., Scott, D.A., Eck, M.J., Gray, N.S., and Jänne, P.A. (2022). An allosteric inhibitor against the therapy-resistant mutant forms of EGFR in non-small cell lung cancer. Nat Cancer 1–16.

Torre, E.A., Arai, E., Bayatpour, S., Jiang, C.L., Beck, L.E., Emert, B.L., Shaffer, S.M., Mellis, I.A., Fane, M.E., Alicea, G.M., Budinich, K.A., Weeraratna, A.T., Shi, J., and Raj, A. (2021). Genetic screening for single-cell variability modulators driving therapy resistance. Nat Genet *53*, 76–85.

Touil, Y., Igoudjil, W., Corvaisier, M., Dessein, A.-F., Vandomme, J., Monté, D., Stechly, L., Skrypek, N., Langlois, C., Grard, G., Millet, G., Leteurtre, E., Dumont, P., Truant, S., Pruvot, F.-R., Hebbar, M., Fan, F., Ellis, L.M., Formstecher, P., Van Seuningen, I., Gespach, C., Polakowska, R., and Huet, G. (2014). Colon cancer cells escape 5FU chemotherapy-induced cell death by entering stemness and quiescence associated with the c-Yes/YAP axis. Clin Cancer Res *20*, 837–846.

Tricker, E.M., Xu, C., Uddin, S., Capelletti, M., Ercan, D., Ogino, A., Pratilas, C.A., Rosen, N., Gray, N.S., Wong, K.-K., and Jänne, P.A. (2015). Combined EGFR/MEK Inhibition Prevents the Emergence of Resistance in EGFR-Mutant Lung Cancer. Cancer Discovery *5*, 960–971.

Turke, A.B., Zejnullahu, K., Wu, Y.-L., Song, Y., Dias-Santagata, D., Lifshits, E., Toschi, L., Rogers, A., Mok, T., Sequist, L., Lindeman, N.I., Murphy, C., Akhavanfard, S., Yeap, B.Y., Xiao, Y., Capelletti, M., Iafrate, A.J., Lee, C., Christensen, J.G., Engelman, J.A., and Jänne, P.A. (2010). Pre-existence and clonal selection of MET amplification in EGFR mutant NSCLC. Cancer Cell *17*, 77–88.

Umkehrer, C., Holstein, F., Formenti, L., Jude, J., Froussios, K., Neumann, T., Cronin, S.M., Haas, L., Lipp, J.J., Burkard, T.R., Fellner, M., Wiesner, T., Zuber, J., and Obenauf, A.C. (2021). Isolating live cell clones from barcoded populations using CRISPRa-inducible reporters. Nat Biotechnol *39*, 174–178.

Van Der Steen, N., Caparello, C., Rolfo, C., Pauwels, P., Peters, G.J., and Giovannetti, E. (2016). New developments in the management of non-small-cell lung cancer, focus on rociletinib: what went wrong? Onco Targets Ther *9*, 6065–6074.

Vander Heiden, M.G., Cantley, L.C., and Thompson, C.B. (2009). Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. Science *324*, 1029–1033.

VanHorn, S., and Morris, S.A. (2021). Next-Generation Lineage Tracing and Fate Mapping to Interrogate Development. Developmental Cell *56*, 7–21.

Vasconcelos, P.E.N.S., Gergis, C., Viray, H., Varkaris, A., Fujii, M., Rangachari, D., VanderLaan, P.A., Kobayashi, I.S., Kobayashi, S.S., and Costa, D.B. (2020). EGFR-A763_Y764insFQEA Is a Unique Exon 20 Insertion Mutation That Displays Sensitivity to Approved and In-Development Lung Cancer EGFR Tyrosine Kinase Inhibitors. JTO Clin Res Rep *1*, 100051.

Vellanki, P.J., Mulkey, F., Jaigirdar, A.A., Rodriguez, L., Wang, Y., Xu, Y., Zhao, H., Liu, J., Howe, G., Wang, J., Choo, Q., Golding, S.J., Mansell, V., Korsah, K., Spillman, D., de Claro, R.A., Pazdur, R., Beaver, J.A., and Singh, H. (2021). FDA Approval Summary: Nivolumab with Ipilimumab and Chemotherapy for Metastatic Non-Small Cell Lung Cancer, a Collaborative Project Orbis Review. Clin Cancer Res *27*, 3522–3527.

Vlashi, E., and Pajonk, F. (2015). Cancer Stem Cells, Cancer Cell Plasticity and Radiation Therapy. Semin Cancer Biol *0*, 28–35.

Vogelstein, B., and Kinzler, K.W. (2015). The Path to Cancer --Three Strikes and You're Out. N. Engl. J. Med. *373*, 1895–1898.

Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz, L.A., and Kinzler, K.W. (2013). Cancer Genome Landscapes. Science *339*, 1546–1558.

Walsh, C., and Cepko, C.L. (1992). Widespread dispersion of neuronal clones across functional regions of the cerebral cortex. Science *255*, 434–440.

Wang, M., Herbst, R.S., and Boshoff, C. (2021). Toward personalized treatment approaches for non-small-cell lung cancer. Nat Med *27*, 1345–1356.

Wang, Q., Yang, S., Wang, K., and Sun, S.-Y. (2019). MET inhibitors for targeted therapy of EGFR TKI-resistant lung cancer. J Hematol Oncol *12*, 63.

Wang, T., Wei, J.J., Sabatini, D.M., and Lander, E.S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. Science *343*, 80–84.

Weber, W., and Fussenegger, M. (2012). Emerging biomedical applications of synthetic biology. Nat Rev Genet *13*, 21–35.

Weinreb, C., Rodriguez-Fraticelli, A., Camargo, F., and Klein, A.M. (2020). Lineage tracing on transcriptional landscapes links state to fate during differentiation. Science *367*, eaaw3381.

Weinstein, I.B. (2002). Addiction to Oncogenes--the Achilles Heal of Cancer. Science *297*, 63–64.

Welch, J.S., Ley, T.J., Link, D.C., Miller, C.A., Larson, D.E., Koboldt, D.C., Wartman, L.D., Lamprecht, T.L., Liu, F., Xia, J., Kandoth, C., Fulton, R.S., McLellan, M.D., Dooling, D.J., Wallis, J.W., Chen, K., Harris, C.C., Schmidt, H.K., Kalicki-Veizer, J.M., Lu, C., Zhang, Q., Lin, L., O'Laughlin, M.D., McMichael, J.F., Delehaunty, K.D., Fulton, L.A., Magrini, V.J., McGrath, S.D., Demeter, R.T., Vickery, T.L., Hundal, J., Cook, L.L., Swift, G.W., Reed, J.P., Alldredge, P.A., Wylie, T.N., Walker, J.R., Watson, M.A., Heath, S.E., et al. (2012). The Origin and Evolution of Mutations in Acute Myeloid Leukemia. Cell *150*, 264–278.

Weng, C.-H., Chen, L.-Y., Lin, Y.-C., Shih, J.-Y., Lin, Y.-C., Tseng, R.-Y., Chiu, A.-C., Yeh, Y.-H., Liu, C., Lin, Y.-T., Fang, J.-M., and Chen, C.-C. (2019). Epithelial-mesenchymal transition (EMT) beyond EGFR mutations per se is a common mechanism for acquired resistance to EGFR TKI. Oncogene *38*, 455–468.

Westover, D., Zugazagoitia, J., Cho, B.C., Lovly, C.M., and Paz-Ares, L. (2018). Mechanisms of acquired resistance to first- and second-generation EGFR tyrosine kinase inhibitors. Annals of Oncology *29*, i10–i19.

White, M.N., Piotrowska, Z., Stirling, K., Liu, S.V., Banwait, M.K., Cunanan, K., Sequist, L.V., Wakelee, H.A., Hausrath, D., and Neal, J.W. (2021). Combining Osimertinib With Chemotherapy in EGFR-Mutant NSCLC at Progression. Clin Lung Cancer *22*, 201–209.

Wilson, W.R., and Hay, M.P. (2011). Targeting hypoxia in cancer therapy. Nat Rev Cancer *11*, 393–410.

Wu, X., Gu, Z., Chen, Y., Chen, B., Chen, W., Weng, L., and Liu, X. (2019). Application of PD-1 Blockade in Cancer Immunotherapy. Comput Struct Biotechnol J *17*, 661–674.

Wu, Y.-L., Zhang, L., Kim, D.-W., Liu, X., Lee, D.H., Yang, J.C.-H., Ahn, M.-J., Vansteenkiste, J.F., Su, W.-C., Felip, E., Chia, V., Glaser, S., Pultar, P., Zhao, S., Peng, B., Akimov, M., and Tan, D.S.W. (2018). Phase Ib/II Study of Capmatinib (INC280) Plus Gefitinib After Failure of Epidermal Growth Factor Receptor (EGFR) Inhibitor Therapy in Patients With EGFR-Mutated, MET Factor-Dysregulated Non-Small-Cell Lung Cancer. J Clin Oncol *36*, 3101–3109.

Yaeger, R., and Corcoran, R.B. (2019). Targeting Alterations in the RAF–MEK Pathway. Cancer Discov *9*, 329–341.

Yang, J.C.-H., Camidge, D.R., Yang, C.-T., Zhou, J., Guo, R., Chiu, C.-H., Chang, G.-C., Shiah, H.-S., Chen, Y., Wang, C.-C., Berz, D., Su, W.-C., Yang, N., Wang, Z., Fang, J., Chen, J., Nikolinakos, P., Lu, Y., Pan, H., Maniam, A., Bazhenova, L., Shirai, K., Jahanzeb, M., Willis, M., Masood, N., Chowhan, N., Hsia, T.-C., Jian, H., and Lu, S. (2020). Safety, Efficacy, and Pharmacokinetics of Almonertinib (HS-10296) in Pretreated Patients With EGFR-Mutated Advanced NSCLC: A Multicenter, Open-label, Phase 1 Trial. Journal of Thoracic Oncology *15*, 1907–1918.

Yang, Z., Yang, N., Ou, Q., Xiang, Y., Jiang, T., Wu, X., Bao, H., Tong, X., Wang, X., Shao, Y.W., Liu, Y., Wang, Y., and Zhou, C. (2018). Investigating Novel Resistance Mechanisms to Third-Generation EGFR Tyrosine Kinase Inhibitor Osimertinib in Non–Small Cell Lung Cancer Patients. Clinical Cancer Research *24*, 3097–3107.

Yao, Z., Torres, N.M., Tao, A., Gao, Y., Luo, L., Li, Q., de Stanchina, E., Abdel-Wahab, O., Solit, D.B., Poulikakos, P.I., and Rosen, N. (2015). BRAF Mutants Evade ERK-Dependent Feedback by Different Mechanisms that Determine Their Sensitivity to Pharmacologic Inhibition. Cancer Cell *28*, 370–383.

Yao, Z., Yaeger, R., Rodrik-Outmezguine, V.S., Tao, A., Torres, N.M., Chang, M.T., Drosten, M., Zhao, H., Cecchi, F., Hembrough, T., Michels, J., Baumert, H., Miles, L., Campbell, N.M., de Stanchina, E., Solit, D.B., Barbacid, M., Taylor, B.S., and Rosen, N. (2017). Tumours with class 3 BRAF mutants are sensitive to the inhibition of activated RAS. Nature *548*, 234–238.

Yates, L.R., and Campbell, P.J. (2012). Evolution of the cancer genome. Nat Rev Genet *13*, 795–806.

Yates, L.R., Gerstung, M., Knappskog, S., Desmedt, C., Gundem, G., Van Loo, P., Aas, T., Alexandrov, L.B., Larsimont, D., Davies, H., Li, Y., Ju, Y.S., Ramakrishna, M., Haugland, H.K., Lilleng, P.K., Nik-Zainal, S., McLaren, S., Butler, A., Martin, S., Glodzik, D., Menzies, A., Raine, K., Hinton, J., Jones, D., Mudie, L.J., Jiang, B., Vincent, D., Greene-Colozzi, A., Adnet, P.-Y., Fatima, A., Maetens, M., Ignatiadis, M., Stratton, M.R., Sotiriou, C., Richardson, A.L., Lønning, P.E., Wedge, D.C., and Campbell, P.J. (2015). Subclonal diversification of primary breast cancer revealed by multiregion sequencing. Nat Med *21*, 751–759.

Yeo, C.J. (1999). Tumor suppressor genes: A short review. Surgery *125*, 363–366.

Yeo, N.C., Chavez, A., Lance-Byrne, A., Chan, Y., Menn, D., Milanova, D., Kuo, C.-C., Guo, X., Sharma, S., Tung, A., Cecchi, R.J., Tuttle, M., Pradhan, S., Lim, E.T., Davidsohn, N., Ebrahimkhani, M.R., Collins, J.J., Lewis, N.E., Kiani, S., and Church, G.M. (2018). An enhanced CRISPR repressor for targeted mammalian gene regulation. Nat Methods *15*, 611–616.

Yochum, Z.A., Cades, J., Wang, H., Chatterjee, S., Simons, B.W., O'Brien, J.P., Khetarpal, S.K., Lemtiri-Chlieh, G., Myers, K.V., Huang, E.H.-B., Rudin, C.M., Tran, P.T., and Burns, T.F. (2019). Targeting the EMT transcription factor TWIST1 overcomes resistance to EGFR inhibitors in EGFR-mutant non-small cell lung cancer. Oncogene *38*, 656–670.

Yu, C., Mannan, A.M., Yvone, G.M., Ross, K.N., Zhang, Y.-L., Marton, M.A., Taylor, B.R., Crenshaw, A., Gould, J.Z., Tamayo, P., Weir, B.A., Tsherniak, A., Wong, B., Garraway, L.A., Shamji, A.F., Palmer, M.A., Foley, M.A., Winckler, W., Schreiber, S.L., Kung, A.L., and Golub, T.R. (2016). High-throughput identification of genotype-specific cancer vulnerabilities in mixtures of barcoded tumor cell lines. Nat Biotechnol *34*, 419–423.

Yu, H.A., Arcila, M.E., Rekhtman, N., Sima, C.S., Zakowski, M.F., Pao, W., Kris, M.G., Miller, V.A., Ladanyi, M., and Riely, G.J. (2013). Analysis of Tumor Specimens at the Time of Acquired Resistance to EGFR TKI therapy in 155 patients with EGFR mutant Lung Cancers. Clin Cancer Res *19*, 2240–2247.

Yun, C.-H., Mengwasser, K.E., Toms, A.V., Woo, M.S., Greulich, H., Wong, K.-K., Meyerson, M., and Eck, M.J. (2008). The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. Proc Natl Acad Sci U S A *105*, 2070–2075.

Zappa, C., and Mousa, S.A. (2016). Non-small cell lung cancer: current treatment and future advances. Transl Lung Cancer Res *5*, 288–300.

Zhang, H., Zhao, H.-Y., Xi, X.-X., Liu, Y.-J., Xin, M., Mao, S., Zhang, J.-J., Lu, A.-X., and Zhang, S.-Q. (2020). Discovery of potent epidermal growth factor receptor (EGFR) degraders by proteolysis targeting chimera (PROTAC). Eur J Med Chem *189*, 112061.

Zhang, Y., He, B., Zhou, D., Li, M., and Hu, C. (2018). Newly emergent acquired EGFR exon 18 G724S mutation after resistance of a T790M specific EGFR inhibitor osimertinib in non-small-cell lung cancer: a case report. Onco Targets Ther *12*, 51–56.

Zhang, Y.-L., Yuan, J.-Q., Wang, K.-F., Fu, X.-H., Han, X.-R., Threapleton, D., Yang, Z.-Y., Mao, C., and Tang, J.-L. (2016). The prevalence of EGFR mutation in patients with non-small cell lung cancer: a systematic review and meta-analysis. Oncotarget *7*, 78985–78993.

Zhang, Z., Lee, J.C., Lin, L., Olivas, V., Au, V., LaFramboise, T., Abdel-Rahman, M., Wang, X., Levine, A.D., Rho, J.K., Choi, Y.J., Choi, C.-M., Kim, S.-W., Jang, S.J., Park, Y.S., Kim, W.S., Lee, D.H., Lee, J.-S., Miller, V.A., Arcila, M., Ladanyi, M., Moonsamy, P., Sawyers, C., Boggon, T.J., Ma, P.C., Costa, C., Taron, M., Rosell, R., Halmos, B., and Bivona, T.G. (2012). Activation of the AXL kinase causes resistance to EGFR-targeted therapy in lung cancer. Nat Genet *44*, 852–860.

Zhao, H.-Y., Yang, X.-Y., Lei, H., Xi, X.-X., Lu, S.-M., Zhang, J.-J., Xin, M., and Zhang, S.-Q. (2020). Discovery of potent small molecule PROTACs targeting mutant EGFR. European Journal of Medicinal Chemistry *208*, 112781.

Zhou, C., Wu, Y.L., Chen, G., Feng, J., Liu, X.-Q., Wang, C., Zhang, S., Wang, J., Zhou, S., Ren, S., Lu, S., Zhang, L., Hu, C., Hu, C., Luo, Y., Chen, L., Ye, M., Huang, J., Zhi, X., Zhang, Y., Xiu, Q., Ma, J., Zhang, L., and You, C. (2015). Final overall survival results from a randomised, phase III study of erlotinib versus chemotherapy as first-line treatment of EGFR mutation-positive advanced non-small-cell lung cancer (OPTIMAL, CTONG-0802). Annals of Oncology *26*, 1877–1883.

Zhou, W., Ercan, D., Chen, L., Yun, C.-H., Li, D., Capelletti, M., Cortot, A.B., Chirieac, L., Iacob, R.E., Padera, R., Engen, J.R., Wong, K.-K., Eck, M.J., Gray, N.S., and Jänne, P.A. (2009). Novel mutant-selective EGFR kinase inhibitors against EGFR T790M. Nature *462*, 1070–1074.