



Universiteit
Leiden
The Netherlands

Advancing precision oncology: insights from C-terminal FGFR2/3 truncations and MYC-driven mTOR inhibitor resistance

Yemelyanenko, J.

Citation

Yemelyanenko, J. (2026, September 22). *Advancing precision oncology: insights from C-terminal FGFR2/3 truncations and MYC-driven mTOR inhibitor resistance*. Retrieved from <https://hdl.handle.net/1887/4309879>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4309879>

Note: To cite this publication please use the final published version (if applicable).

ADVANCING PRECISION ONCOLOGY

Insights from C-terminal FGFR2/3 truncations
and MYC-driven mTOR inhibitor resistance

Julia Yemelyanenko



ADVANCING PRECISION ONCOLOGY

Insights from C-terminal FGFR2/3
truncations and MYC-driven mTOR
inhibitor resistance

About the cover: The cover of this thesis symbolizes the wide spectrum of cancer patient identities, tumor types, and disease presentations encountered in oncology, represented by multiple abstract faces depicted in different colors. A subset of these faces is interconnected by lines, symbolizing shared genomic alterations that can range from single-nucleotide mutations to complex structural rearrangements. Although diverse in nature, such alterations may converge on common functional outcomes that drive tumor formation or therapy resistance. The ability to recognize these hidden molecular patterns relies on advanced next-generation sequencing technologies, robust computational and statistical frameworks, and the systematic functional validation of detected variants. Together, these approaches enable the identification of clinically meaningful alterations, including C-terminal truncations of *FGFR2* and *FGFR3*, as well as *MYC* activation associated with resistance to mTOR inhibitors, key discoveries explored throughout this thesis. By uncovering shared vulnerabilities among diverse tumors, this knowledge contributes to improved patient stratification and the development of more effective, individualized treatments, thereby advancing the goals of precision oncology.

Cover design: Sergio Vallés Martí | Instagram: @sergiarts

Thesis lay-out: Douwe Oppewal

Printing: Gildeprint Enschede | www.gildeprint.nl

ISBN: 978-94-6496-573-5

© Copyright, Julia Yemelyanenko, 2026

All rights reserved. No part of this book may be reproduced in any form or by any means without permission of the author.

ADVANCING PRECISION ONCOLOGY:

Insights from C-terminal FGFR2/3 truncations and MYC-driven mTOR inhibitor resistance

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr. S. de Rijcke,
volgens besluit van het college voor promoties
te verdedigen op dinsdag 22 september 2026
klokke 11:30 uur

door

Julia Yemelyanenko Lyalenko
geboren te Sevastopol, Oekraïne
in 1993

Promotor: Prof.dr. J.M.M. Jonkers

Co-promotor: Dr. D.K. Zingg

Promotiecommissie:

Prof.dr. M. van Eck

Prof.dr. E.C.M. de Lange

Prof.dr. M.M. Maurice (University Medical Center Utrecht)

Prof.dr. J.W.M. Martens (Erasmus University Medical Center)

Prof.dr. P. ten Dijke (Leiden University Medical Center)

The research described in this thesis was performed at the division of Molecular Pathology at the Netherlands Cancer Institute – Antoni van Leeuwenhoek Hospital (NKI-AVL) (Amsterdam, The Netherlands). The research was financially supported by the Dutch Cancer Society and Oncode Institute.

TABLE OF CONTENTS

	List of abbreviations	6
Chapter 1	General introduction and scope of this thesis	11
Chapter 2	FGFR rearrangements: oncogenic drivers and therapeutic targets <i>Trends Cancer</i> , 2026 May;12(5):461-476.	21
Chapter 3	Truncated <i>FGFR2</i> is a clinically actionable oncogene in multiple cancers <i>Nature</i> , 2022 Aug; 608(7923):609-617	49
Chapter 4	The C-terminal Kinase Domain–Binding and Suppression motif prevents constitutive activation of FGFR2 <i>Cancer Research</i> , 2025 Jun 24; 85(17):3234–57	121
Chapter 5	C-terminal Truncation and Fusion Partner Determine Oncogenicity of FGFR3 <i>Cancer Research</i> , 2026 Mar 16; 86(6):1372-1391	177
Chapter 6	MYC is a clinically significant driver of mTOR inhibitor resistance in breast cancer <i>Journal of Experimental Medicine</i> , 2023 Nov 6;220(11):e20211743	227
Chapter 7	General discussion	281
Appendices	English summary	308
	Nederlandse samenvatting	310
	Curriculum vitae	313
	List of publications	314

LIST OF ABBREVIATIONS

100kGP:	100,000 Genomes Project
3'-UTR:	3'-untranslated region
AACR:	American Association for Cancer Research
ADCs:	antibody–drug conjugates
AI:	artificial intelligence
amp:	amplification
AUC:	area under the dose–response curve
AURKA:	Aurora kinase A
BPs:	breakpoints
CCA:	cholangiocarcinoma
CC:	coiled-coil
ChIP:	chromatin immunoprecipitation
Chr-2:	chromosome 2
CML:	chronic myeloid leukemia
co-IP:	co-immunoprecipitation
CR:	complete response
C-tail:	carboxy-terminal tail
ctDNA:	circulating tumor DNA
E18:	exon 18
EMA:	European Medicines Agency
ER+:	estrogen receptor–positive
ERK:	extracellular signal–regulated kinase
FDA:	U.S. Food and Drug Administration
FGF:	fibroblast growth factor
FGFR:	fibroblast growth factor receptor
<i>FGFR2</i>^{hotspot}:	<i>FGFR2</i> missense hotspot mutation
<i>FGFR2</i>^{ΔE18}:	exon 18–truncated <i>FGFR2</i>
<i>FGFR3</i>^{ΔE18}:	exon 18–truncated <i>FGFR3</i>
FGFRi:	FGFR inhibitors
FL:	full-length
FMI:	Foundation Medicine Inc.
GDC:	Genomic Data Commons
GEMMs:	genetically engineered mouse models
GENIE:	Genomics Evidence Neoplasia Information Exchange
GO:	Gene Ontology
GSVA:	gene set variation analysis
HER2:	human epidermal growth factor receptor 2
HMF:	Hartwig Medical Foundation
HR:	hazard ratio

HSPG:	heparan sulfate proteoglycan
I17:	intron 17
iCCA:	intrahepatic cholangiocarcinoma
IC50:	half-maximal inhibitory concentration
ICI:	immune checkpoint inhibitor
ICGC:	International Cancer Genome Consortium
IGR:	intergenic region
IHC:	immunohistochemistry
ILC:	invasive lobular carcinoma
IMD:	IRSp53/MIM homology domain
JM:	juxtamembrane domain
KD:	kinase domain
KDBS:	kinase domain-binding and suppression motif
KEP:	<i>K14-Cre;Cdh1^{F/F};Trp53^{F/F}</i> mouse model
LC-WGS:	low-coverage whole genome sequencing
LisH:	lissencephaly-1-homology domain
LUAD:	lung adenocarcinoma
LUSC:	lung squamous cell carcinoma
mAbs:	monoclonal antibodies
MAPK:	mitogen-activated protein kinase
MS:	mass spectrometry
MSI-H:	microsatellite instability-high
mTOR:	mechanistic target of rapamycin
mTORi:	mTOR inhibitors
MYCi:	MYC inhibitor
NGS:	next-generation sequencing
NSCLC:	non-small cell lung cancer
NXG:	<i>NOD-Prkdcscid-IL2rgTm1/Rj</i> immunocompromised mice
OS:	overall survival
PCAWG:	Pan-Cancer Analysis of Whole Genomes
PDAC:	pancreatic ductal adenocarcinoma
PDX:	patient-derived xenograft
PFS:	progression-free survival
PI3K:	phosphoinositide 3-kinase
PI3Ki:	PI3K inhibitors
PKC:	protein kinase C
PR:	partial response
RAF:	rapidly accelerated fibrosarcoma
RAS:	rat sarcoma
REs:	rearrangements
Ribo-seq:	ribosome profiling

RNA-seq:	RNA sequencing
RPPA:	reverse phase protein array profiling
RR:	response rate
RTK:	receptor tyrosine kinase
RT-qPCR:	reverse transcription quantitative PCR
SB:	Sleeping Beauty transposon system
SD:	stable disease
siRNAs:	small interfering RNAs
STAT:	signal transducer and activator of transcription
TACC3:	transforming acidic coiled-coil containing protein 3
TAM:	tamoxifen
TCGA:	The Cancer Genome Atlas
TEs:	translation efficiencies
TF:	transcription factor
TIDE:	Tracking of Indels by Decomposition
TKI:	tyrosine kinase inhibitor
TMB-H:	tumor mutation burden–high
TMT:	tandem mass tag
tRNA:	transfer RNA
VUS:	variants of unknown significance
WES:	whole exome sequencing
WGS:	whole genome sequencing
WGTS:	whole genome and transcriptome sequencing
WT:	wild-type



CHAPTER 1

GENERAL INTRODUCTION AND SCOPE OF THIS THESIS

Cancer is the second-leading cause of death worldwide, surpassed only by cardiovascular disease [1]. Globally, about 1 in 5 people will develop cancer during their lifetime, with 1 in 9 men and 1 in 12 women dying from the disease [2]. In 2022, the World Health Organization's International Agency for Research on Cancer reported approximately 20 million new cancer cases and 9.7 million deaths worldwide [2]. In 2023, an estimated 18.5 million new cases and 10.4 million deaths occurred globally, excluding non-melanoma skin cancers [1]. These numbers will continue to grow, as the global cancer burden is expected to almost double by 2050, reaching over 35 million new cases and 18.6 million deaths annually, largely driven by population growth, aging, and lifestyle-associated risk factors [1]. Metastatic disease, which accounts for two-thirds – and in some reports up to 90% – of cancer-related deaths in solid tumors [3], remains particularly devastating, with a median survival of only about 10 months for affected patients [4]. Therefore, even though over the past two decades major advances have deepened our understanding of the molecular mechanisms underlying cancer initiation, progression, metastasis, and therapy response, much remains to be done to transform these discoveries into lasting improvements in cancer care.

Part of the enduring challenge in curing cancer stems from its complexity and diversity. No two tumors are truly alike. While tumors of the same type may share hallmark molecular features, they differ markedly between patients, particularly at advanced disease stages. This variability extends beyond genomic alterations to include differences in the tumor microenvironment, immune landscape, and epigenetic regulation, all of which shape tumor behavior and therapeutic response. At the genomic level, it is extremely rare for two metastatic tumors to share an identical fingerprint [5–8]. Additionally, treatment exposure further shapes the tumor genome and contributes to this diversity by imposing an evolutionary bottleneck that selects for therapy-resistant clones. One of the largest pan-cancer analyses of metastatic solid tumors to date showed that 96% of driver mutations within individual metastatic lesions are clonal [7]. Consistently, roughly half of patients with annotated treatment histories (53%) harbor treatment-enriched known or suspected tumor drivers [8], underscoring the strong selective pressure exerted by anticancer therapies during tumor evolution.

The growing understanding of the vast tumor heterogeneity and therapeutic selection pressures has ushered in an era of precision oncology. One-size-fits-all therapeutic approaches are suboptimal and are increasingly being replaced by targeted therapies tailored to the molecular features of each patient's individual tumor [9,10].

Precision oncology: reality or illusion?

The central tenet of precision oncology is to deliver the right drug(s) to the right patient at the right time [11]. Chronic myeloid leukemia (CML) is the poster child for precision oncology. Since its introduction in 2002 as first-line treatment, imatinib, a tyrosine

kinase inhibitor (TKI) targeting the BCR-ABL1 fusion oncoprotein that drives CML, has revolutionized therapy and transformed a once fatal malignancy into a manageable chronic condition with a near-normal life expectancy when treated early [12]. Other successes in precision oncology include trastuzumab for HER2-positive breast cancer [13]; EGFR [14], ALK [15] and RET inhibitors [16] for lung cancers; and NTRK inhibitors for TRK fusion-positive cancers [17]. These targeted drugs have improved clinical outcomes in solid cancers, but with much more modest impact compared to imatinib in CML. Nevertheless, biomarker-guided therapy is the strongest determinant of improved clinical outcomes such as higher response rate (RR) and longer progression-free survival (PFS) and overall survival (OS) [18–21]. Notably, targeted therapies are not intrinsically more effective than cytotoxic agents; rather, their success depends on appropriate biomarker selection to guide their use. When administered without a matching biomarker, their efficacy is minimal and can be even inferior to that achieved with conventional chemotherapy [19,21].

Despite significant promise and a few isolated success stories, the overall impact of targeted therapies on patient outcomes has unfortunately been limited. A meta-analysis of 346 phase I trials involving over 13,000 patients found that biomarker-based selection improved response rates; however, only 30% of patients with refractory cancers responded at all, and median PFS was just 5.7 months compared with 2.95 months in unselected cohorts [21]. Similarly, reviews of cancer drugs approved by the European Medicines Agency (EMA) and the U.S. Food and Drug Administration (FDA) have shown that most offered minimal or no gains in OS or quality of life [22,23]. On average, new anticancer agents approved by the EMA and FDA between 2003 and 2013 extended OS by only about 3.4 months relative to previously available treatments [24].

The limited benefits of targeted therapies likely stem from multiple factors. Drug approvals frequently rely on surrogate endpoints that show weak correlation with OS [25]. In addition, the coexistence of distinct genomic alterations across tumor subclones can undermine the effectiveness of single-agent monotherapies, emphasizing the need for rational combination strategies that balance efficacy with tolerable toxicity. Timing also plays a critical role: most targeted agents are tested in heavily pretreated metastatic patients who have exhausted standard-of-care options. Administering these therapies earlier, before extensive clonal evolution and emergence of resistance, may be key to improving outcomes [10,26]. This underscores the importance of incorporating routine molecular profiling at diagnosis to match patients with the most effective therapy from the outset.

The evolution of genome sequencing: from the first human genome to routine molecular profiling

Sequencing of the first full human genome was a landmark scientific achievement that required enormous resources, global collaboration, and more than a decade to complete. The Human Genome Project ultimately cost about \$2.7-3 billion USD, took 13 years

(1990-2003), and involved thousands of researchers at more than 20 research centers and universities across the world [27]. This is in stark contrast to the speed and cost efficiency of modern whole genome sequencing (WGS). Today, the cost for WGS is projected to be around \$200, with a goal of reaching \$100 per genome in the future. Even though WGS for clinical use requires extensive data analysis, interpretation, and use of consumables, which can raise total costs to several thousand dollars per genome [28], the dramatic reduction in sequencing costs has enabled genomic profiling of hundreds of cancer-associated genes as part of routine cancer care in many developed countries. This, in turn, has enabled the identification of hundreds of actionable genomic alterations and spurred the development of new targeted and immunotherapies. However, despite these significant advances, the functional impact of alterations in many oncogenic pathways remain incompletely understood.

From genomic profiling to biological and clinical insight

Translating genomic profiling results into clinical decision-making critically depends on understanding the functional consequences of individual genetic alterations. However, many variants detected in known cancer driver genes remain classified as variants of unknown significance (VUS), limiting their utility for therapy selection. This is particularly consequential when such variants occur in genes that are amenable to targeting by clinically approved therapies. There is therefore a pressing need for systematic and rapid functional characterization of VUS to distinguish biologically relevant driver events from neutral (passenger) variants and thereby support accurate and personalized treatment allocation.

Recent advances in high- and moderate-throughput functional genomics platforms have enabled the systematic interrogation of the oncogenic potential and drug sensitivity of large numbers of cancer-associated variants across clinically actionable genes. These approaches have been particularly instrumental in uncovering the oncogenic activity and underlying mechanisms of rare and low-frequency variants that would be difficult to prioritize based on recurrence alone [29–35]. For example, functional annotation of rare mutations in pancreatic ductal adenocarcinoma (PDAC) identified activating mutations in NADK, uncovering a previously underappreciated role for redox homeostasis in tumorigenesis and highlighting a potential therapeutic vulnerability [32]. Similarly, saturation mutagenesis and large-scale functional screens of kinases such as EGFR [29,33], HER2 [34] and FGFRs [35] have uncovered numerous low-prevalence activating and resistance-conferring mutations, many of which localize to outside the kinase domain yet remain sensitive to approved targeted therapies. Collectively, these studies demonstrate that systematic functional interrogation of rare variants can reveal novel mechanisms of oncogenic activation and expand the population of patients who may benefit from existing drugs.

Despite these successes, current functional annotation efforts remain heavily biased toward single-nucleotide variants and small indels, leaving other classes of VUS, including structural variants and gene fusions, far less systematically explored. Indeed, both experimental screening platforms and *in silico* saturation mutagenesis approaches have predominantly focused on missense alterations [29–31,33–36], whereas the functional and clinical consequences of complex rearrangements, domain truncations and non-canonical splice variants remain poorly characterized. Addressing this imbalance will be essential for achieving a comprehensive functional interpretation of the cancer genome and for fully realizing the promise of precision oncology.

SCOPE OF THIS THESIS

Building on these knowledge gaps, the present thesis aims to advance our understanding of the functional and clinical implications of genomic alterations in the fibroblast growth factor receptor (FGFR) family, as well as to elucidate mechanisms of therapeutic resistance within the PI3K-AKT-mTOR pathway.

In **Chapter 2** [37], we review the biological and clinical relevance of canonical in-frame FGFR fusions and complex non-canonical structural FGFR variants in cancer. We discuss their oncogenic mechanisms, their value as predictive biomarkers of therapeutic response, and the emerging challenges and opportunities for FGFR-targeted therapy. Central to this review is the discovery that exon 18 (E18) truncations in *FGFR2* and *FGFR3* are critical for oncogenic activation, a finding that, as later detailed in **Chapters 3-5** [38–40], has redefined our understanding of FGFR-driven tumorigenesis and informed the design of clinical trials evaluating next-generation FGFR inhibitors. Together, these insights highlight the importance of more precisely defining which *FGFR* genomic alterations are truly oncogenic and therapeutically actionable.

In **Chapter 3** [38], we use a combination of transposon-based screening, mouse tumor modeling and functional assays to identify C-terminal truncation of E18 in *FGFR2* as a potent single-driver oncogenic event. Analysis of human oncogenomic datasets reveals that this truncation can arise through multiple mechanisms (e.g. rearrangements, partial amplifications, and truncating mutations), all converging on the production of E18-truncated *FGFR2* (*FGFR2^{ΔE18}*) oncogenic transcripts. In contrast, the oncogenic competence of full-length *FGFR2* amplification depends on the presence of additional co-driver alterations. Importantly, we show that expression of *FGFR2^{ΔE18}* is a predictive biomarker of FGFR-targeted therapy response, with direct implications for patient selection and the design of therapeutic strategies.

Building on these findings, in **Chapter 4** [39], we investigate the molecular mechanisms through which loss of the C-terminal tail leads to FGFR2 oncogenic activation. Through a systematic functional analysis of *Fgfr2*-E18 variants, we identify a previously unrecognized phenylalanine-serine (F703-S704) motif within the C-terminal region that mediates autoinhibitory binding of the C-tail to the kinase domain. This ‘kinase domain-binding and suppression’ (KDBS) motif is essential for restraining FGFR2 kinase activity under normal conditions. Disruption or truncation of this motif, either through E18 loss or targeted mutation, abolishes this autoinhibitory control, resulting in constitutive activation of FGFR2 and aberrant stimulation of the PI3K-AKT-mTOR and STAT3 signaling pathways. We show that permutation of the KDBS motif, together with additional regulatory C-terminal sites, fully phenocopies the oncogenic competence of FGFR2^{ΔE18}, establishing the KDBS motif as a central molecular switch controlling FGFR2 activation. These insights reveal a structural and mechanistic basis for oncogenic FGFR2 activation through C-terminal truncation and uncover potential therapeutic vulnerabilities, as combined inhibition of FGFR2 with PI3K-AKT-mTOR or STAT3 pathway inhibitors may overcome on-target resistance to FGFR-targeted therapies in FGFR2^{ΔE18}-positive cancers.

Prompted by our findings in FGFR2, in **Chapter 5** [40], we investigate whether analogous C-terminal tail loss in its closely related FGFR3 paralog similarly drives oncogenicity. Intriguingly, we uncover that FGFR3 exhibits divergent behavior: unlike FGFR2, FGFR3 requires both E18 truncation and fusion to a partner gene retaining the expression of a dimerizing domain to effectively drive oncogenic signaling and tumorigenesis. We also functionally validate the oncogenic potential of previously uncharacterized clinical *FGFR3* gene fusions (FGFR3^{ΔE18}-ADD1, FGFR3^{ΔE18}-TNIP2, and FGFR3^{ΔE18}-JAKMIP1). All these fusions contain a self-dimerizing domain contributed by the fusion partner, confer resistance to HER2/EGFR tyrosine kinase inhibitors, yet remain sensitive to pharmacologic FGFR inhibition *in vitro*.

Whereas **Chapters 2-5** focus on elucidating the oncogenic mechanisms of FGFR signaling and their implications for targeted therapy, **Chapter 6** [41] broadens this scope to explore mechanisms of therapeutic resistance within the PI3K-AKT-mTOR pathway, one of the major signaling axis frequently co-opted by receptor tyrosine kinases such as FGFRs. Despite the clinical promise of mTOR inhibitors (mTORi) in breast cancer, treatment efficacy remains limited by intrinsic and acquired resistance. Through multiomic profiling of treatment-naïve and mTORi-resistant mouse invasive lobular carcinomas, we identify MYC activation as a key driver of resistance. MYC counteracts mTORi-induced translational suppression by sustaining ribosomal protein synthesis and tumor growth. Moreover, mTORi-resistant cells depend on MYC activity, as combined inhibition of mTOR and MYC restores therapeutic sensitivity. Importantly, we show that MYC activation correlates with poor response to the mTORi everolimus in metastatic breast cancer patients, establishing MYC as a

potential driver of clinical mTORi resistance that may stratify breast cancer patients for mTOR-targeted therapies.

Finally, the general discussion in **Chapter 7** reflects on the main findings presented in this thesis, highlights remaining challenges, and outlines directions for future research.

REFERENCES

1. Force, L.M. *et al.* (2025) The global, regional, and national burden of cancer, 1990–2023, with forecasts to 2050: a systematic analysis for the Global Burden of Disease Study 2023. *The Lancet* 406, 1565–1586
2. Bray, F. *et al.* (2024) Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians* 74, 229–263
3. Dillekås, H. *et al.* (2019) Are 90% of deaths from cancer caused by metastases? *Cancer Med* 8, 5574–5576
4. Mani, K. *et al.* (2024) Causes of death among people living with metastatic cancer. *Nat Commun* 15, 1519
5. Wheler, J. *et al.* (2014) Unique molecular landscapes in cancer: implications for individualized, curated drug combinations. *Cancer Res* 74, 7181–7184
6. Egas-Bejar, D. *et al.* (2014) Theranostic Profiling for Actionable Aberrations in Advanced High Risk Osteosarcoma with Aggressive Biology Reveals High Molecular Diversity: The Human Fingerprint Hypothesis. *Oncoscience* 1, 167–179
7. Priestley, P. *et al.* (2019) Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature* 575, 210–216
8. Martínez-Jiménez, F. *et al.* (2023) Pan-cancer whole-genome comparison of primary and metastatic solid tumours. *Nature* 618, 333–341
9. Adashek, J.J. *et al.* (2021) From Tissue-Agnostic to N-of-One Therapies: (R)Evolution of the Precision Paradigm. *Trends in Cancer* 7, 15–28
10. Subbiah, V. and Kurzrock, R. (2018) Challenging Standard-of-Care Paradigms in the Precision Oncology Era. *Trends Cancer* 4, 101–109
11. Mateo, J. *et al.* (2022) Delivering precision oncology to patients with cancer. *Nat Med* 28, 658–665
12. Jabbour, E. and Kantarjian, H. (2024) Chronic myeloid leukemia: 2025 update on diagnosis, therapy, and monitoring. *American Journal of Hematology* 99, 2191–2212
13. Piccart-Gebhart, M.J. *et al.* (2005) Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *N Engl J Med* 353, 1659–1672
14. Soria, J.-C. *et al.* (2018) Osimertinib in Untreated EGFR-Mutated Advanced Non-Small-Cell Lung Cancer. *N Engl J Med* 378, 113–125
15. Shaw, A.T. *et al.* (2020) First-Line Lorlatinib or Crizotinib in Advanced ALK-Positive Lung Cancer. *N Engl J Med* 383, 2018–2029
16. Drilon, A. *et al.* (2020) Efficacy of Selpercatinib in RET Fusion-Positive Non-Small-Cell Lung Cancer. *N Engl J Med* 383, 813–824
17. Drilon, A. *et al.* (2018) Efficacy of Larotrectinib in TRK Fusion-Positive Cancers in Adults and Children. *N Engl J Med* 378, 731–739
18. Jardim, D.L. *et al.* (2015) Impact of a Biomarker-Based Strategy on Oncology Drug Development: A Meta-analysis of Clinical Trials Leading to FDA Approval. *J Natl Cancer Inst* 107, djv253
19. Schwaederle, M. *et al.* (2015) Impact of Precision Medicine in Diverse Cancers: A Meta-Analysis of Phase II Clinical Trials. *J Clin Oncol* 33, 3817–3825
20. Schwaederle, M. *et al.* (2016) Precision Oncology: The UC San Diego Moores Cancer Center PREDICT Experience. *Mol Cancer Ther* 15, 743–752
21. Schwaederle, M. *et al.* (2016) Association of Biomarker-Based Treatment Strategies With Response Rates and Progression-Free Survival in Refractory Malignant Neoplasms: A Meta-analysis. *JAMA Oncol* 2, 1452–1459

22. Kim, C. and Prasad, V. (2015) Cancer Drugs Approved on the Basis of a Surrogate End Point and Subsequent Overall Survival: An Analysis of 5 Years of US Food and Drug Administration Approvals. *JAMA Intern Med* 175, 1992–1994
23. Davis, C. *et al.* (2017) Availability of evidence of benefits on overall survival and quality of life of cancer drugs approved by European Medicines Agency: retrospective cohort study of drug approvals 2009–13. *BMJ* 359, j4530
24. Salas-Vega, S. *et al.* (2017) Assessment of Overall Survival, Quality of Life, and Safety Benefits Associated With New Cancer Medicines. *JAMA Oncol* 3, 382–390
25. Prasad, V. *et al.* (2015) The Strength of Association Between Surrogate End Points and Survival in Oncology: A Systematic Review of Trial-Level Meta-analyses. *JAMA Intern Med* 175, 1389–1398
26. Wahida, A. *et al.* (2023) The coming decade in precision oncology: six riddles. *Nat Rev Cancer* 23, 43–54
27. Initial sequencing and analysis of the human genome | Nature[Online]. Available: <https://www.nature.com/articles/35057062>. [Accessed: 31-Oct-2025]
28. Schwarze, K. *et al.* (2020) The complete costs of genome sequencing: a microcosting study in cancer and rare diseases from a single center in the United Kingdom. *Genet Med* 22, 85–94
29. Ng, P.K.-S. *et al.* (2018) Systematic Functional Annotation of Somatic Mutations in Cancer. *Cancer Cell* 33, 450–462.e10
30. Kim, E. *et al.* (2016) Systematic Functional Interrogation of Rare Cancer Variants Identifies Oncogenic Alleles. *Cancer Discov* 6, 714–726
31. Han, H. *et al.* (2024) Functional annotation of the Hippo pathway somatic mutations in human cancers. *Nat Commun* 15, 10106
32. Tsang, Y.H. *et al.* (2016) Functional annotation of rare gene aberration drivers of pancreatic cancer. *Nat Commun* 7, 10500
33. Hayes, T.K. *et al.* (2024) Comprehensive mutational scanning of EGFR reveals TKI sensitivities of extracellular domain mutants. *Nat Commun* 15, 2742
34. Pahuja, K.B. *et al.* (2018) Actionable Activating Oncogenic ERBB2/HER2 Transmembrane and Juxtamembrane Domain Mutations. *Cancer Cell* 34, 792–806.e5
35. Tangermann, C. *et al.* (2025) Saturation mutagenesis identifies activating and resistance-inducing FGFR kinase domain mutations. *Nat Genet* DOI: 10.1038/s41588-025-02431-8
36. Muñíos, F. *et al.* (2021) In silico saturation mutagenesis of cancer genes. *Nature* 596, 428–432
37. Yemelyanenko, J. *et al.* (2026) FGFR rearrangements: oncogenic drivers and therapeutic targets. *Trends in Cancer* DOI: 10.1016/j.trecan.2026.01.010
38. Zingg, D. *et al.* (2022) Truncated FGFR2 is a clinically actionable oncogene in multiple cancers. *Nature* 608, 609–617
39. Zingg, D. *et al.* (2025) The C-terminal Kinase Domain-Binding and Suppression Motif Prevents Constitutive Activation of FGFR2. *Cancer Research* DOI: 10.1158/0008-5472.CAN-24-3349
40. Yemelyanenko, J. *et al.* (2025) C-terminal Truncation and Fusion Partner Determine Oncogenicity of FGFR3. *Cancer Res* DOI: 10.1158/0008-5472.CAN-24-2648
41. Bhin, J. *et al.* (2023) MYC is a clinically significant driver of mTOR inhibitor resistance in breast cancer. *Journal of Experimental Medicine* 220, e20211743

2

CHAPTER 2

FGFR REARRANGEMENTS: ONCOGENIC DRIVERS AND THERAPEUTIC TARGETS

Julia Yemelyanenko, Jos Jonkers# & Daniel Zingg#

co-corresponding author

Trends in Cancer, 2026 May; 12(5):461-476
<https://doi.org/10.1016/j.trecan.2026.01.010>

ABSTRACT

Genomic amplifications and hotspot mutations affecting fibroblast growth factor receptors (FGFRs) have long been recognized as oncogenic drivers across human cancers. However, recent studies have uncovered *FGFR* in-frame fusions and complex structural variants as an additional class of tumor driver alterations. Notably, the identification of *FGFR2* exon 18 truncations and the demonstration of their potent oncogenic competence have refined our understanding of *FGFR*-driven tumorigenesis and have impacted clinical trial design for FGFR-targeted agents. This review explores the biological and clinical implications of *FGFR* rearrangements. It covers their mechanisms in driving cancer, their potential as biomarkers to predict treatment response, and the emerging challenges and opportunities for FGFR-targeted therapy. Ultimately, a deeper understanding of *FGFR* rearrangements is critical for advancing precision oncology and improving patient benefit.

Highlights

- Fibroblast growth factor receptor (*FGFR*) rearrangements, including canonical in-frame fusions and noncanonical structural variants, function as oncogenic drivers in multiple cancers.
- Advances in next-generation sequencing technologies and the integration of multiomic diagnostics enable the detection of rare and complex noncanonical FGFR rearrangements that are expressed.
- Noncanonical FGFR rearrangements that involve noncoding or intergenic regions can generate functional oncogenic isoforms through pseudoexon formation or promoter/enhancer hijacking.
- Structural rearrangements that generate exon 18-truncated variants result in distinct functional consequences for FGFR2 versus FGFR3.
- *FGFR* rearrangements emerge as actionable biomarkers, guiding the use of FGFR-targeted therapy.

Hijacking growth: novel mechanisms underlying FGFR oncogenicity

Fibroblast growth factor (FGF) ligands and FGF receptors (FGFRs) were linked to cancer for the first time in the late 1980s and early 1990s, when genetic amplification and aberrant expression of FGFs and FGFRs were detected in various cancer tissues and cell lines [1–4]. Though these initial discoveries were made more than 3 decades ago, the full complexity of *FGFR* alterations in cancer has become apparent only through more recent large-scale oncogenomic studies, as well as the implementation of targeted tumor sequencing assays into routine clinical practice. A notable yet diverse class of previously underappreciated *FGFR* alterations is structural rearrangements that range from in-frame fusions to complex noncanonical events. Recent work has demonstrated that most of these rearrangements

can act as oncogenic drivers and consequently dictate therapeutic response to FGFR inhibitors. In particular, it has largely gone unnoticed that rearrangements in *FGFR2* and *FGFR3* typically truncate exon 18 (E18), which encodes the C-terminal tails of these receptor tyrosine kinases (RTKs) [5–7]. The E18-truncation event has turned out to be critical for FGFR2 and FGFR3 oncogenicity, but the full oncogenic competence of *FGFR3* rearrangements depends on distinct additional determinants [6,7]. These findings have redefined our understanding of FGFR-driven tumorigenesis and have impacted the design of clinical trials evaluating next-generation FGFR inhibitors.

Tumor sequencing studies, and in particular the ever-growing wealth of real-world oncogenomic data, will continue to identify novel *FGFR* rearrangements across cancer indications. However, their functional and clinical relevance will remain incompletely understood, highlighting the urgent need for systematic characterization to guide precision therapy more effectively. In this review, we explore the biology and clinical relevance of *FGFR* rearrangements in cancer. We examine how canonical and noncanonical rearrangements arise and contribute to tumorigenesis across various cancer types. We also discuss the oncogenic mechanisms that underlie these events and their utility as predictive biomarkers for FGFR-targeted therapies, focusing on approved small-molecule inhibitors, co-occurring alterations, and resistance mechanisms. Finally, we outline key challenges and future directions for the clinical translation of these insights into refined patient stratification and better treatment strategies.

Canonical FGF/FGFR signaling

Structural features and signaling players

FGFRs belong to the RTK family and comprise four highly conserved transmembrane receptors (FGFR1-FGFR4), as well as the structurally related FGFR-like 1 (FGFRL1, also known as FGFR5). FGFR1-4 share a common architecture consisting of three extracellular immunoglobulin like domains (D1-D3) that mediate ligand binding, a single-pass transmembrane domain, and an intracellular region containing a split tyrosine kinase domain essential for signal transduction, followed by a short C-terminal tail that contributes to downstream signaling regulation (**Figure 1A**, Key figure) [8–10]. Alternative splicing of exons encoding the D3 domain of FGFR1-3 generates the *b* and *c* isoforms that confer distinct ligand-binding and tissue specificities [8,9]. FGFRL1 retains sequence similarity to the extracellular and transmembrane domains of FGFR1-4; however, it only possesses a short intracellular tail that lacks the split tyrosine kinase domain, precluding trans-autophosphorylation and engagement in canonical FGFR signaling pathways [11,12]. Given the lack of kinase activity of FGFRL1, the focus of this review is exclusively on the four canonical FGFRs (FGFR1-4).

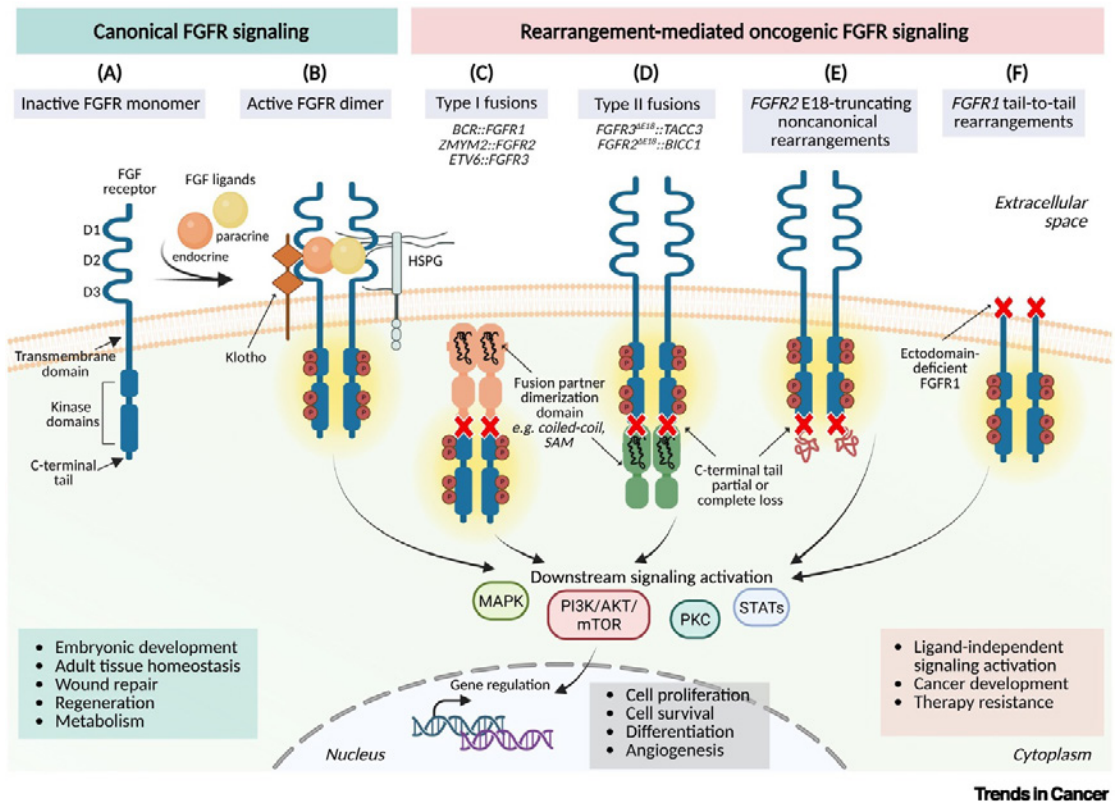


Figure 1. Overview of canonical versus rearrangement-mediated oncogenic FGFR signaling.

During canonical FGFR signaling, an inactive FGFR monomer (A) is bound by an FGF ligand together with a cofactor, either an HSPG (paracrine) or Klotho (endocrine), thereby inducing receptor dimerization and subsequent transphosphorylation of tyrosine residues within the kinase domains (B). These phosphotyrosines act as docking sites for key intracellular effectors, leading to activation of downstream transduction pathways, including MAPK, PI3K/AKT/mTOR, PKC, and STATs. The resulting transcriptional programs drive cell proliferation, survival, differentiation, and angiogenesis, which in healthy conditions support embryonic development, adult tissue homeostasis, wound repair, regeneration, and metabolic processes. In contrast, rearrangement-mediated oncogenic FGFR signaling (C–F) occurs independently of ligand binding and promotes constitutive signaling, oncogenic transformation, and therapy resistance. Type I fusions (C) retain the intracellular kinase domain but lose the extracellular and transmembrane regions, which are replaced by a 5' partner containing a dimerization domain, leading to altered subcellular localization. In type II fusions (D), the extracellular, transmembrane, and kinase domains are typically retained, whereas the C-terminal tail is truncated – commonly at intron 17 or exon 18 (E18) of the gene – and replaced by the fusion partner, which often contributes a dimerization domain. Canonical and noncanonical *FGFR2* E18-truncating rearrangements (E) disrupt the C-terminal tail of *FGFR2*, abolishing autoinhibitory motifs and thereby promoting unchecked kinase activity and downstream signaling. Noncanonical intragenic or “tail-to-tail” rearrangements within or near the *FGFR1* gene generate ectodomain-deficient (N-terminally truncated) *FGFR1* (F). Created in <https://BioRender.com>. AKT: protein kinase B; FGF: fibroblast growth factor; FGFRs: fibroblast growth factor receptors; HSPG: heparan sulfate proteoglycan; MAPK: mitogen-activated protein kinase; mTOR: mammalian target of rapamycin; PI3K: phosphoinositide 3-kinase; PKC: protein kinase C; SAM: sterile alpha motif; STAT: signal transducer and activator of transcription.

FGFR activation is initiated by the binding of an FGF ligand to an inactive FGFR monomer, in conjunction with heparan sulfate proteoglycan (HSPG) or Klotho cofactors (**Figure 1A**). This induces a conformational change that promotes receptor dimerization and subsequent transphosphorylation of the tyrosine residues within the kinase domains (**Figure 1B**). First, the A-loop tyrosines are phosphorylated, followed by those in the carboxy (C)-terminal tail, kinase insert and juxtamembrane regions [8]. The phosphorylated tyrosine residues on activated FGFRs serve as docking sites for key intracellular signaling mediators. These are mainly the adaptor protein FGFR substrate 2 α , which mediates the assembly of various signaling complexes, and the hydrolase phospholipase C γ , which hydrolyzes phosphatidylinositol 4,5-bisphosphate, generating the second messenger molecules diacylglycerol and inositol trisphosphate. These coordinated events lead to the activation of multiple downstream transduction pathways, including rat sarcoma (RAS) – mitogen-activated protein kinase (MAPK), phosphoinositide 3-kinase (PI3K) – protein kinase B (AKT) – mammalian target of rapamycin (mTOR), protein kinase C (PKC), and signal transducer and activator of transcription (STAT)-dependent signaling [8,10,13]. To date, 18 different secreted FGF ligands capable of activating FGFRs have been identified. The majority of FGFs are paracrine factors with broad but tissue-specific expression patterns (FGF1-10, FGF16-18, FGF20, and FGF22), whereas a subset of FGFs are secreted by endocrine organs and act systemically via circulation (FGF19, FGF21, and FGF23) [8,9]. The binding specificity between FGFs and their cognate receptors is determined by their sequence diversity and spatiotemporal expression patterns.

Physiological functions and regulatory control

FGF/FGFR signaling plays critical roles throughout embryonic development by regulating early mesoderm patterning, organogenesis, differentiation, and cell proliferation and migration [9,10,14]. FGF/FGFR signaling is also vital during various adult physiological processes, such as tissue homeostasis, wound repair, and regeneration across multiple organ systems, including bone, muscle, and the vasculature, by sustaining cellular proliferation and survival and modulating metabolic processes [10]. Given its broad physiological relevance, FGF/FGFR signaling must be precisely regulated in space, time, and magnitude. To this end, cells have evolved an extensive repertoire of mechanisms that collectively ensure tight control of FGFR activity to prevent excessive signaling. These include receptor endocytosis, inhibitory adaptor proteins, dephosphorylation of key tyrosines by phosphatases, as well as serine/threonine phosphorylations that serve as inhibitory cues, typically installed by downstream signaling kinases as a negative-feedback loop [9,10,14,15]. The diversity and partial redundancy of the different negative control mechanisms underscore the critical need to tightly control and attenuate the FGF/FGFR pathway activity. Notably, its dysregulation is central to various pathologies, such as skeletal disorders, chronic kidney disease, and cancer [8,10].

Aberrant FGF/FGFR signaling in cancer

Over the past 2 decades, major progress has been made in understanding the roles of the FGF/FGFR signaling pathway in tumor initiation, progression, and angiogenesis [16]. Beyond its oncogenic functions, aberrant FGF/FGFR signaling also contributes to resistance against chemotherapy, radiotherapy, and targeted therapies such as CDK4/6 inhibitors, endocrine therapy, and human epidermal growth factor receptor 2 (HER2)-directed agents, underscoring its multifaceted role in tumor biology and treatment response [17–22].

Today, we know that aberrant FGF/FGFR signaling activation is primarily caused by genetic alterations in the *FGFR* genes and deregulation of the release of FGF ligands by cancer cells, as well as stromal fibroblasts and macrophages [16,23]. Up to 7.1% of all human cancers harbor alterations in the *FGFR* genes, which occur across a broad spectrum of tumor types (**Figure 2A**) [24–28]. *FGFR* amplifications are most frequent (~5%), followed by activating mutations (~2%) and gene rearrangements (~0.5%) [24,25,27,28]. Both the prevalence and

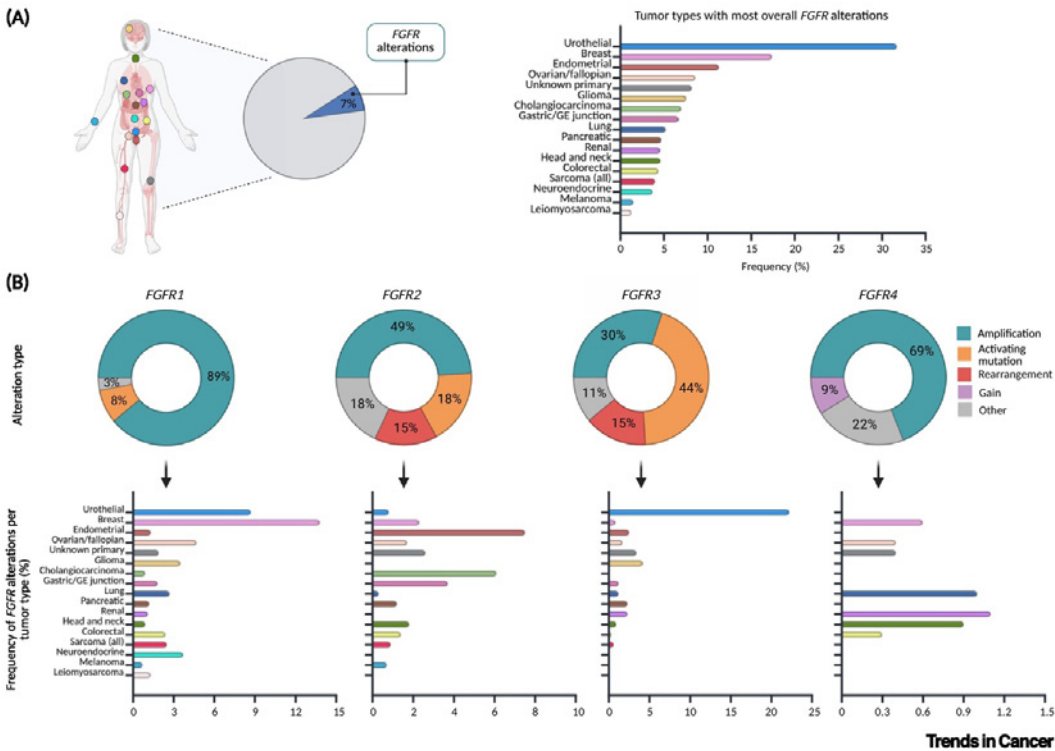


Figure 2. Landscape of *FGFR* genomic alterations across human cancers.

(A) Overall prevalence and tumor type distribution of *FGFR* genomic alterations. (B) Comparative alteration spectra across *FGFR1*-*FGFR4*, showing that each receptor exhibits its own characteristic mix of genomic events (amplification, activating mutation, rearrangement, gain, or other) and tumor type distributions. Data derived from [24]. Created in <https://BioRender.com>. GE: gastroesophageal.

the types of *FGFR* alterations vary significantly across tumor types (**Figure 2B**). For instance, oncogenic *FGFR* alterations are frequent in urothelial (~32%), breast (~18%), endometrial (~13%), and squamous lung cancers (~13%), whereas they are rare in soft tissue sarcomas, melanomas, and neuroendocrine tumors [24]. *FGFR3* hotspot mutations are the most common alteration type in urothelial carcinomas. Conversely, *FGFR1* amplifications and *FGFR2* mutations are prevalent in breast and squamous cell carcinomas and in endometrial cancer, respectively.

The functional impact of each *FGFR* alteration depends on both the specific genetic change and the tissue type affected, ultimately shaping tumor behavior and influencing patient response to FGFR-targeted therapies [16]. The growing catalog of *FGFR* rearrangements revealed by recent oncogenomic studies reflects the complexity of FGFR-driven oncogenesis. Yet, their biological and clinical significance remains only partially understood, warranting comprehensive functional characterization to inform precision oncology.

***FGFR* gene rearrangements in cancer**

Genomic rearrangements are large-scale structural alterations of the genome that disrupt the normal linear organization of the DNA. They include large insertions, duplications, deletions, inversions, and chromosomal translocations, which can span kilobases to megabases and affect gene function by duplicating or removing coding sequences, creating fusion genes, or altering regulatory elements. In addition to these chromosomal events, extrachromosomal DNA represents a distinct mode of genome reorganization in which circular DNA fragments carrying oncogenes or enhancers are excised from chromosomes, amplified independently, and can subsequently reintegrate into the genome to drive oncogene expression [29]. Rearrangements can occur within the same chromosome (intrachromosomal) or between different chromosomes (interchromosomal). At the molecular level, gene rearrangements typically arise from DNA double-strand breaks followed by faulty repair or replication processes, with the resulting structural variations largely shaped by the interplay between error-prone DNA repair pathways [i.e., **nonhomologous end joining** (see Glossary) and **microhomology-mediated end joining**], replication-associated errors, and the broader genomic context [30]. Catastrophic events such as **chromothripsis** and **chromoplexy**, which are now recognized as important cancer driver processes, further contribute to the complexity seen in cancer genomes [30,31].

Although *FGFR* rearrangements account for only 5-10% of all *FGFR* alterations, they have attracted considerable attention in recent years due to their distinct oncogenic mechanisms of action, clinical actionability, and relevance to emerging molecular classification and treatment paradigms [24,25,27,28].

Canonical in-frame FGFR fusions

The most well-studied *FGFR* rearrangements are gene fusions involving other protein-coding partners, particularly those occurring in-frame to produce functional chimeric transcripts. *FGFR* fusions can be classified into type I (**Figure 1C**) or type II (**Figure 1D**), based on whether the C- or N-terminal region of the *FGFR* is retained in the fusion protein [13,32].

Type I fusions, in which *FGFR* is the 3' partner, are more commonly associated with hematological malignancies and involve replacement of the N-terminal portion of *FGFR* – including the extracellular and transmembrane domains – by a fusion partner, leading to loss of the membrane anchor and consequent altered subcellular localization and dysregulated signaling (**Figure 1C**). The kinase domain becomes constitutively active in aberrant intracellular compartments, escaping normal ligand dependence, endocytic downregulation, and membrane-specific spatial feedback controls that normally restrain *FGFR* activation [15]. Examples of type I fusions include *BCR::FGFR1*, *ZMYM2::FGFR2* and *ETV6::FGFR3*.

In contrast, type II fusions, in which *FGFR* is the 5' fusion partner, are more common in solid tumors. Here, the *FGFR* extracellular, transmembrane, and kinase domains are typically retained, whereas the C-terminal tail is truncated – commonly at intron 17 (I17) or E18 of the gene – and replaced by the fusion partner (**Figure 1D**). Examples of type II fusions are *FGFR2::BICC1* and *FGFR3::TACC3*, which are established oncogenic drivers in intrahepatic **cholangiocarcinoma** (iCCA) and glioblastoma, respectively, among other tumor types [5–7,33]. In most type II *FGFR* fusions, the fusion partner contributes self-association domains such as coiled-coil (e.g., *FGFR3::TACC3*), sterile alpha motif (e.g., *FGFR2::BICC1*), Lissencephaly-1-homology (LisH) domain, IRSp53/MIM homology domain (IMD), and caspase domains, which promote ligand-independent dimerization or oligomerization through specific protein-protein interactions [5–7,13]. These fusions typically show ligand-independent and constitutive *FGFR* autophosphorylation and activation of downstream signaling pathways [6,7].

FGFR fusion-positive cancers usually contain few concurrent driver alterations [34–36] and are mutually exclusive with *FGFR* hotspot mutations and amplifications [6,7,37,38], suggesting that *FGFR* fusions function as primary oncogenic drivers and thus represent actionable targets. *FGFR* fusions have also emerged as mechanisms of acquired resistance to targeted therapy, such as EGFR inhibitors in nonsmall cell lung cancer (NSCLC) [20,34,39–41].

Tumor spectrum of FGFR fusions

FGFR fusions involve a broad spectrum of partner genes across a wide range of tumor types [6,7,24,25,27,28]. Type I *FGFR1* fusions, involving partners such as *ZMYM2*, *BCR*, *FOP1*,

and *CNTRL*, are associated with rare but aggressive hematological neoplasms [13,32]. In solid tumors, *FGFR1* fusions are much less common, with only sporadic cases reported in breast, lung, glioma, and salivary gland cancers [13,28,32]. *FGFR2* fusions, in particular type II fusions, are most common in iCCA and may occur in up to 50% of all iCCA cases, but they have also been recurrently observed in breast, lung, ovarian, and pancreatic cancers, among others [6,13,28]. To date, more than a hundred distinct *FGFR2* fusion partners have been identified, with *BICC1*, *TACC2*, *AHCYL1*, and *KIAA1217* being among the most frequent [6,13,28]. *FGFR3* type II fusions are frequently observed in urothelial, brain, lung, and breast cancers, with *TACC3* being the single most recurrent partner (~85%) [7]. Other partners, such as *TNIP2*, *ADD1*, and *JAKMIP1* have also been identified, albeit at much lower frequencies [7,13,24,27,28,42]. In contrast to *FGFR1-3*, *FGFR4* is rarely involved in gene fusions or other structural rearrangements, consistent with the low frequency of *FGFR4* alterations across cancers as compared with *FGFR1-3* (7% for *FGFR4* versus 49% for *FGFR1*, 19% for *FGFR2*, and 26% for *FGFR3*) [24]. The few reported *FGFR4* fusion events have been predominantly found in NSCLC and typically involve type I fusions with partners such as *ANO3*, *NSD1*, and *ZNF346* [24,27,28,34,42].

Noncanonical FGFR rearrangements

FGFR gene fusion studies have traditionally focused on protein-coding gene-gene rearrangements. This has left rearrangements involving noncoding regions of the genome largely unexplored, partly because of sequencing methods that target coding regions only, such as whole-exome sequencing or targeted gene panels. Rearrangements with noncoding regions have also been neglected because they have been considered incapable of generating chimeric transcripts. However, recent studies have demonstrated that structural alterations outside gene bodies can play important roles in tumor development and therapeutic response, in particular, in cancers that show aberrant kinase signaling [6,43,44]. The prevalence and potential significance of these events are underscored by the fact that up to 62% of all rearrangements found pan-cancer involve breakpoints in intergenic regions [43]. These rearrangements may drive tumorigenesis through two distinct mechanisms. First, gene-intergenic and intergenic-intergenic rearrangements can result in the repositioning of distant promoters or enhancers, boosting mRNA expression of oncogenes (known as promoter or enhancer “swapping” or “hijacking”) [43]. Second, intergenic sequences can give rise to pseudoexons that lead to the formation of functional transcripts by providing splice donor or acceptor sites and thereby maintaining reading frames or creating new junctions [45].

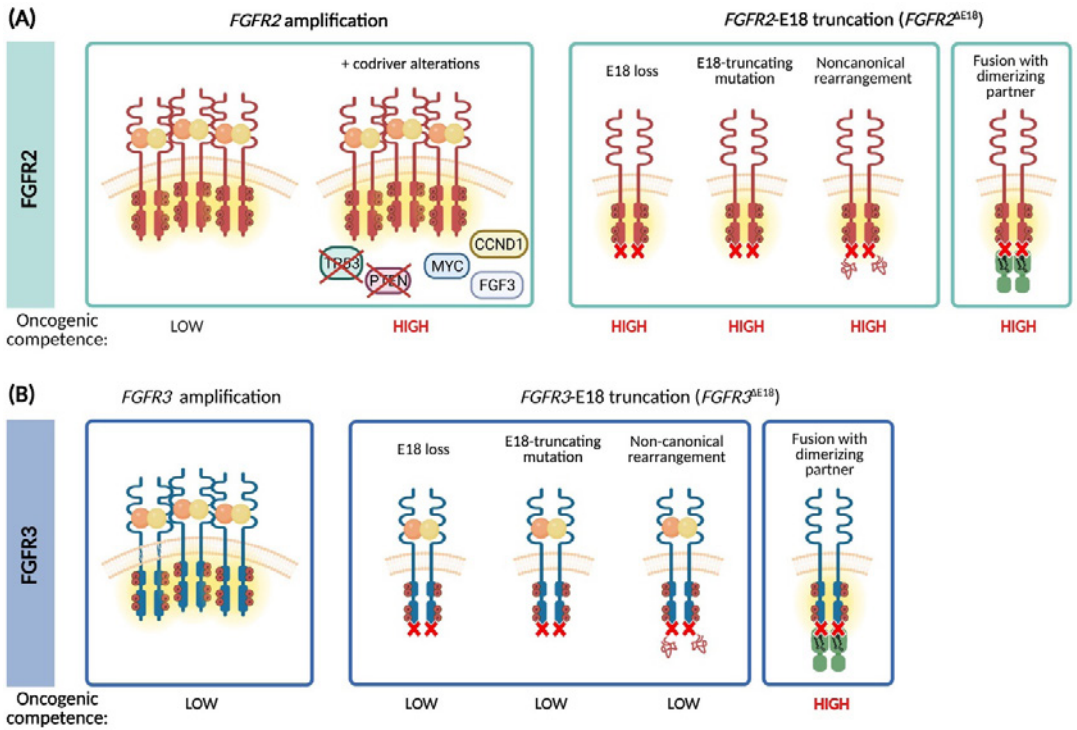
Next-generation sequencing (NGS) technologies have greatly improved the detection of rare and complex noncanonical rearrangements, with recent studies showing that these events are of particular clinical relevance for *FGFRs* [6,7,46]. For instance, noncanonical intragenic or “tail-to-tail” rearrangements within or near the *FGFR1* gene can generate ectodomain-deficient (N-terminally truncated) proteins that preserve kinase activity

(**Figure 1F**) and drive squamous cell lung cancer through ligand-independent dimerization [46]. Analysis of paired whole-genome sequencing (WGS) and RNA-sequencing data from a pan-cancer cohort also uncovered a substantial fraction of tumors harboring functionally relevant noncanonical rearrangements in E18 of *FGFR2* (**Figure 1E**) [6]. These include events in which the reading frame of the partner gene is unknown (potentially out of frame), the partner gene is out of strand, or the rearrangement involves intergenic or internal *FGFR* gene sequences instead of a partner gene.

Oncogenic mechanisms behind *FGFR* rearrangements

Truncation of *FGFR2* at E18 – its last exon encoding the C-terminal tail – acts as a potent oncogenic driver alteration in cancer [6]. Importantly, this tumor-driver potential is independent of a specific fusion partner, indicating that truncation or loss of E18 itself is the critical pathogenic event (**Figure 3A**). E18-truncated *FGFR2* (*FGFR2*^{ΔE18}) transcripts can result from different types of genomic alterations, including gene rearrangements or gene fusions, E1-17 partial amplifications, as well as E18 nonsense and frameshift mutations (**Figure 3A**). Notably, in a large real-world cohort of tumor sequencing profiles, among the 1,367 tumor samples that contained *FGFR2* alterations potentially generating E18-truncated transcripts, 44.6% were noncanonical rearrangements. Analysis of RNA-sequencing profiles from a subset of these cases confirmed the generation of functional *FGFR2*^{ΔE18} transcripts [6]. These findings suggest that noncanonical rearrangements that would have previously been classified as variants of unknown significance can give rise to functional oncogenic *FGFR2*^{ΔE18} isoforms, positioning them as bona fide tumor drivers and actionable targets for FGFR-targeted therapies. The tumor-driving effect of E18 loss in *FGFR2* was recently attributed to the disruption of a phenylalanine-serine motif that is located in the proximal part of the C-tail [47]. These two residues are essential for binding of the *FGFR2* C-tail to the kinase domain and for restraining kinase activity, and have therefore been termed the kinase domain-binding and suppression (KDBS) motif. Functional *in vitro* assays, together with *in vivo* mammary tumor modeling in mice, demonstrated that the C-terminal tail suppresses *FGFR2* oncogenicity through this KDBS motif and the concurrent regulation of endocytosis and the PI3K/ extracellular signal-regulated kinase (ERK) negative-feedback loop.

Perturbation of E18 of *FGFR3*, which encodes the *FGFR3* C-terminal tail akin to *FGFR2*, is also recurrent in cancer, resulting in the expression of *FGFR3*^{ΔE18} transcripts that lack (part of) E18 and thus have dysfunctional C-terminal tails [7]. However, in contrast to *FGFR2*, *FGFR3* requires both E18 truncation and fusion to a partner gene that retains the expression of a dimerizing domain for full oncogenic competence (**Figure 3B**). Interestingly, the KDBS motif is not fully conserved in *FGFR3* (phenylalanine-alanine instead of phenylalanine-serine), which might explain the differential impact of C-terminal truncation on *FGFR2* versus *FGFR3* oncogenicity and the selective pressure for *FGFR3* to preferably fuse with partners that contain dimerization domains. Notably, 95% of partner genes involved in



Trends in Cancer

Figure 3. Differential oncogenic competence of amplified full-length versus E18-truncated FGFR2/3.

(A) The oncogenic competence of full-length FGFR2 relies on cooperating codriver alterations such as TP53 or PTEN loss, or overexpression of MYC, CCND1, or FGF3. In contrast, E18-truncated variants promote tumorigenesis without requiring additional alterations. The key pathogenic event is disruption of E18, whether through exon loss, truncating mutations, noncanonical rearrangements, or fusion with a partner harboring a dimerizing domain. (B) For FGFR3, full oncogenic competence requires E18 truncation combined with fusion to a dimerizing partner. *FGFR3* amplification, E18 loss, E18-truncating mutations, and noncanonical rearrangements exhibit limited oncogenic activity and might depend on cooperating cancer driver mutations to enhance their oncogenic potential. Created in <https://BioRender.com>. CCND1: cyclin D1; FGF3: fibroblast growth factor receptor; MYC: v-myc myelocytomatosis viral oncogene homolog; PTEN: phosphatase and tensin homolog; TP53: tumor protein 53.

FGFR3^{ΔE18} fusions encode self-interacting domains, with the coiled-coil motif being the most frequent. TACC3 contains two coiled-coil domains that are typically preserved in *FGFR3*^{ΔE18}::TACC3 fusions. The *FGFR3* fusion partner oligomerization domains mediate constitutive, ligand-independent receptor dimerization, thus driving aberrant *FGFR3* kinase activation and downstream signaling [5–7,32]. Nevertheless, a sizable subset of all tumors harbor E18-truncating mutations (12.3%), potentially resulting in *FGFR3*^{ΔE18} transcripts. These alterations are devoid of fusion partners, suggesting that *FGFR3*^{ΔE18} might display oncogenic activity in specific comutational and/or cellular contexts [7].

Beyond partner-driven dimerization, *FGFR* fusions and noncanonical rearrangements can disrupt transcriptional or translational regulatory elements of the *FGFR* loci, thereby altering protein expression or stability. Additionally, they can introduce premature stop codons or delete negative regulatory regions (e.g., the KDBS motif in *FGFR2*) or extracellular domains, resulting in the loss of inhibitory signals and enhanced *FGFR* signaling [13,46,47]. Some rearrangements can also result in aberrant subcellular localization of the fusion protein, with retention in the cytoplasm rather than the plasma membrane, thereby disrupting normal signaling regulation (**Figure 1C**) [48].

***FGFR* rearrangements as predictive biomarkers for *FGFR* targeted therapy**

The implication of FGF/*FGFR* signaling in the pathogenicity of cancer has prompted the development of a range of *FGFR*-targeted therapeutic approaches, several of which have achieved regulatory approval for specific *FGFR*-altered tumor indications, and many more are currently being evaluated in clinical trials (**Table 1**).

Small-molecule FGFR inhibitors

The first generation of *FGFR* inhibitors consisted of multikinase small molecules that target the ATP-binding pocket of *FGFR1-4*, as well as other RTKs with highly structurally conserved tyrosine kinase domains (derazantinib, dovitinib, tasurgratinib, lenvatinib, lucitanib, nintedanib and, ponatinib) [83,84]. These molecules have shown considerable toxicity at the dosage required for full *FGFR* inhibition (likely due to multi-RTK inhibition), thus limiting their utility in clinical practice. This spurred the development of next-generation *FGFR* inhibitors, which have demonstrated improved potency and selectivity. They include pan-*FGFR* inhibitors (erdafitinib, rogaratinib, futibatinib, and resigratinib), as well as selective inhibitors of *FGFR1-3* (pemigatinib, infigratinib, fexagratinib, and zoligratinib), *FGFR2* (lirafugratinib), *FGFR3* (LOXO-435), and *FGFR4* (roblitinib and fisogatinib) [83–85]. Erdafitinib, pemigatinib, infigratinib, rogaratinib, and zoligratinib block *FGFR* kinase activity in an ATP-competitive and noncovalent manner, comparable to first-generation inhibitors. Conversely, futibatinib, resigratinib, lirafugratinib, roblitinib, and fisogatinib are covalent inhibitors that form irreversible bonds with specific cysteine residues within the ATP-binding pocket, resulting in stronger target engagement, enhanced selectivity, and less susceptibility to on-target resistance mutations [49,57,86].

To date, four *FGFR* inhibitors – three noncovalent (erdafitinib, infigratinib, and pemigatinib), and one covalent (futibatinib) – have received clinical approval for second- or later-line treatment in advanced solid tumors harboring specific *FGFR* gene alterations. Based on the phase 2 BLC2001 trial (NCT02365597), erdafitinib received accelerated FDA approval in 2019 for the treatment of *FGFR2/3* fusion- and *FGFR3* mutation-positive urothelial carcinoma [50,51]. The follow-up phase 3 THOR trial (NCT03390504) demonstrated that erdafitinib meaningfully improves objective response rates compared with standard-of-care

Table 1. Small-molecule FGFR inhibitors in clinical use or under development

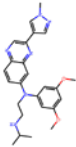
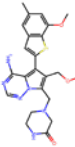
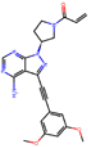
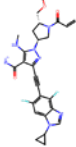
Inhibitor	Structure	Binding mode	Resistance-associated FGFR mutations ^a	Indication	Clinical status	Refs
<i>Pan-FGFR</i>						
Erdafitinib (JNJ-42756493)		Reversible	• FGFR2: N550K, V565F/L, C632Y, and D651Y • FGFR3: N540K, V555L/M, E587Q, and L608V	Urothelial carcinoma	Approved	[49–56]
				Glioma	Phase 2	
				Solid tumors	Phase 2	
Rogaratinib (BAY 1163877)		Reversible	• FGFR2: N550K and V565F/L	Urothelial carcinoma	Phase 1/2/3	[49]
				Squamous cell lung cancer	Phase 2	
				Sarcoma	Phase 2	
Futibatinib (TAS-120)		Covalent (binds FGFR2 C492)	• FGFR2: C492F and V565F Retained potency against: • FGFR2: N550D/H/K, V565I/L, E566A/G, M583I, L618V, K660M/N, and H683L	Cholangiocarcinoma	Approved	[49,54,57–61]
				Urothelial carcinoma	Phase 2	
				Colorectal cancer and other solid tumors	Phase 2	
				Bladder cancer	Phase 2	
				Advanced biliary cancer	Phase 3	
Resigristatinib (KIN-3248)		Covalent (binds FGFR2 C492)	• FGFR2: C492F Retained potency against: • FGFR2: N550H/K, V565F, and K660M • FGFR3: N540K, V555M, and K650M	Advanced solid tumors with FGFR2 and/or FGFR3 alterations	Phase 1b	[62]

Table 1. continued

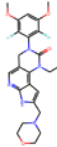
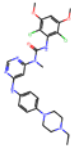
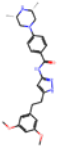
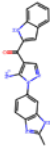
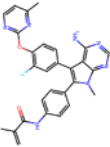
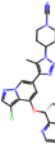
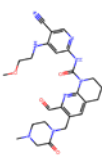
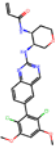
Inhibitor	Structure	Binding mode	Resistance-associated FGFR mutations ^a	Indication	Clinical status	Refs
FGFR1-3						
Pemigatinib (INCB-054828)		Reversible	• FGFR1: N546K and V561L/M	Cholangiocarcinoma	Approved	
			• FGFR2: M537I, N550D/H/K/T, V565F/I/L, E566A, L617V, K641R, and K659M	Myeloid/lymphoid neoplasms	Approved	[49,54,63–68]
			• FGFR3: N540K and V555L/M	Solid tumors Hematologic malignancies	Phase 1/2 Phase 1/2	
Infigratinib (BGJ398)		Reversible	• FGFR1: V561M • FGFR2: N550D/H/K, V565F/L, L618V, and K660M • FGFR3: V555M	Cholangiocarcinoma	Discontinued ^b	
			Retained potency against:			
			• FGFR2: M538I and H683L, which induce clinical resistance to zoligratinib •	Bladder cancer	Discontinued ^b	[49,57,69–72]
Fexagratinib (AZD4547)		Reversible	• FGFR1: V561M • FGFR2: I567S, N568H/T, V581L, E584G, S587L, K660R, and K678M	Solid tumors, lymphoma, and multiple myeloma Breast cancer	Phase 2 Phase 2	[71,73,74]
			• FGFR2: N550K, V565L, L618F/V, and K660M	Breast cancer	Phase 1b/2	[49,57]
Zoligratinib (Debio1347)		Reversible	Retained potency against:			
			• FGFR2: V565F			

Table 1. continued

Inhibitor	Structure	Binding mode	Resistance-associated FGFR mutations ^a	Indication	Clinical status	Refs
FGFR2						
Lirafugratinib (RLY-4008)		Covalent (binds FGFR2 C492)	• FGFR2: C492F - Retained potency against: • FGFR2: N550D/K/T, V565F/I/L, E566G, and K660M	Cholangiocarcinoma and other solid tumors	Phase 1/2	[54,75–77]
FGFR3						
LOXO-435 (LY3866288)		Reversible	Designed to preserve activity against resistance-associated FGFR3 mutations Retained potency against: • FGFR3: V555M	FGFR3-altered solid tumors	Phase 1	[78,79]
FGFR4						
Roblitinib (FGF401)		Covalent (binds FGFR4 C552)	• FGFR4: V550E - Retained potency against: • FGFR4: N535K	Hepatocellular carcinoma and other FGFR4-altered solid tumors	Phase 1/2	[80,81]
Fisogatinib (BLU-554)		Covalent (binds FGFR4 C552)	• FGFR4: V548L/M, V550E/L/M, C552F/R/W/Y, and A553P/V	Hepatocellular carcinoma	Phase 1/2	[82]

^aResistance-associated FGFR mutations:
FGFR1: N546K (molecular brake), V561L/M (gatekeeper).
FGFR2: C492F (covalent binding site, P-loop), N550D/H/K/T (molecular brake), V565F/I/L (gatekeeper), E566A/G (molecular brake), E584G (hinge), K660M/N/R (activation loop), K678M (activation loop).
FGFR3: N540K (molecular brake), V555L/M (gatekeeper), K650M (activation loop).
FGFR4: N535K (molecular brake), V550E/L/M (gatekeeper), C552F/R/W/Y (covalent binding site, hinge-1).
^bInfigratinib (BGJ398) FDA approval was voluntarily withdrawn by its sponsor in May 2024.

chemotherapy in *FGFR2/3*-altered urothelial carcinoma patients previously treated with an immune checkpoint inhibitor (ICI) [52], which led to its full approval in 2024. However, no differentiation versus the next-line ICI pembrolizumab was observed [53]. Pemigatinib [63,64], infigratinib [69], and futibatinib [57,58] also received accelerated FDA approval for the treatment of advanced iCCA harboring *FGFR2* fusions and/or rearrangements in 2020, 2021, and 2022, respectively. In 2022, pemigatinib was also approved for patients with myeloid/lymphoid neoplasms harboring *FGFR1* fusions [65,66].

Notably, in the phase 2 FIGHT-202 trial (NCT02924376) evaluating pemigatinib in CCA, objective and durable responses were observed exclusively in patients with *FGFR2* fusions or rearrangements involving the I17/E18 hotspot region of *FGFR2*, with an objective response rate of 35.5%. In contrast, no objective responses were observed in patients with *FGF/FGFR* alterations other than fusions/rearrangements or in those who were *FGF/FGFR* wild type [63]. A retrospective analysis of the FIGHT-202 trial data revealed that stratifying patients by *FGFR2* E18-truncating rearrangement type (in-frame fusions versus noncanonical rearrangements involving intergenic regions or frame-unknown events) had no impact on the trial outcome; both groups of patients responded to pemigatinib and achieved comparable **progression-free survival (PFS)** benefit, irrespective of the fusion/rearrangement type [6]. Preclinical data suggest that E18-truncating *FGFR2* fusions and rearrangements generate a constitutively active receptor that acts as a potent tumor driver, thereby conferring a strong dependence on FGFR2 signaling that translates into sensitivity to FGFR-targeted agents [6,87]. In contrast, wild-type *FGFR2* and *FGFR2* alterations that retain E18, such as amplifications, produce FGFR2 that remains ligand dependent and is not constitutively active [6]. As the oncogenic driver landscape of iCCA and other cancers suggests, tumors harboring non-E18-truncating *FGFR2* alterations are shaped by additional driver alterations and may therefore be less reliant on FGFR2 signaling [6,88,89]. Ultimately, this may account for the lack of objective responses to pemigatinib in patient cohorts that are devoid of I17/E18 hotspot-rearranged *FGFR2*, as observed in FIGHT-202 [63]. Together, these data suggest that *FGFR2* fusions and noncanonical rearrangements that involve E18 are predictive biomarkers for response to pemigatinib and, likely, for other FGFR-targeted therapies.

Conceivably, clinical outcomes in patients with select genomic *FGFR* alterations may also be improved by the use of FGFR inhibitors in earlier lines. It is noteworthy that two phase 3 trials in CCA evaluating infigratinib (NCT03773302) and pemigatinib (FIGHT-302, NCT03656536) in first-line treatment versus standard-of-care chemotherapy were terminated prematurely. In 2024, FDA approval of infigratinib was withdrawn by its sponsor due to challenges in recruiting patients, making continued commercial viability for the second-line indication unfeasible. Changes in the first-line standard-of-care for CCA also led to the discontinuation of FIGHT-302 in 2025. Efforts that are currently underway to advance next-generation FGFR-targeted medicines are centered on optimizing patient selection,

reducing class-related toxicities (i.e. hyperphosphatemia and ocular, nail, dermatologic, and gastrointestinal adverse events), and improving selectivity and safety profiles [85,90].

Other FGFR-targeted therapeutic approaches

While small-molecule FGFR inhibitors are the most clinically advanced, other modalities targeting the FGF/FGFR axis are also being explored. These include monoclonal antibodies (mAbs), antibody-drug conjugates (ADCs), ligand traps targeting FGFs, and DNA/RNA-based aptamers, among others (reviewed in detail by Katoh *et al.* [84]). Among these, bemarituzumab (FPA144), an FGFR2b-specific mAb engineered to enhance antibody-dependent cellular cytotoxicity, is the most clinically advanced. Bemarituzumab showed promising activity in a phase 2 trial in patients with FGFR2b-overexpressing gastric and gastroesophageal junction cancers [91], and the corresponding phase 3 trials are currently ongoing (NCT05052801 and NCT05111626). In addition, SIM0686 (SCR-A002), an FGFR2b-targeted and topoisomerase I inhibitor payload-conjugated ADC, has demonstrated potent antitumor activity in several preclinical models, including those that show bemarituzumab resistance [92], and is currently being evaluated in a phase 1 trial (NCT07050459). Collectively, these emerging strategies broaden the therapeutic options for FGFR-driven malignancies.

Comutational landscape

Extending the clinical utility of FGFR-targeted therapies will require a deeper understanding of the interplay between *FGFR* alterations and the broader comutational landscape to inform rational combination strategies. Notably, 94% of cancers that contain alterations affecting components of the FGF/FGFR axis harbor one or more additional pathogenic coalterations [89]. These most commonly affect *TP53*, as well as p53 pathway members, cell cycle regulators, other RTKs, and components of the MAPK and PI3K pathways [6,7,24,67,68,88,89,93,94].

Strikingly, different *FGFR* alteration types co-occur with distinct codriver alterations. Across cancers, and in breast cancer specifically, *FGFR2* hotspot mutations and *FGFR2* E18-truncating alterations, including fusions, functionally cooperate with *PIK3CA* hotspot mutations [6,94]. Conversely, *FGFR2* E18-truncating alterations and *FGFR3* fusions are mutually exclusive with *TP53* hotspot missense mutations, *PTEN* loss-of-function, and *MYC*, *CCND1*, and *FGF3/4/19* amplification, whereas these driver gene alterations co-occur with *FGFR2* amplifications [6,68,95]. Functional *in vivo* annotation showed that the oncogenicity of full-length *FGFR2* indeed depends on *TP53* or *PTEN* loss or overexpression of *MYC*, *CCND1*, or *FGF3*, whereas *FGFR2*^{ΔE18} variants promote tumor formation independent of any coalterations (**Figure 3A**) [6]. Interestingly, although *FGFR2* and *FGFR3* E18-truncating fusions/rearrangements are mutually exclusive with *TP53* mutations, the subset of patients double positive for *FGFR2/3* and *TP53* mutation status exhibit particularly low response rates to FGFR inhibitors, manifesting in poor PFS [67,68,95] and **overall survival** [64,89].

Taken together, integrating *FGFR* status information with computational profiles may guide combination treatment strategies to improve patient outcomes.

Preclinical studies have demonstrated convincing synergism between *FGFR* inhibitors and targeted approaches against PI3K [94,96–99], CDK4/6 [19] and EGFR/AXL [100]. In the clinic, *FGFR* targeted therapies are currently being evaluated in combination with PI3K inhibitors, CDK4/6 inhibitors, ICIs, chemotherapeutics, and DNA damage response-targeted agents. Initial reports suggest that such combinations may indeed yield superior clinical response rates compared with their respective monotherapies, particularly when combining *FGFR* inhibition with CDK4/6 blockade [89,101] or ICI [61,102]. Histology-agnostic clinical studies or **basket trials** could offer a valuable framework in this context, although the evaluation of larger patient cohorts will likely be necessary to more clearly define the therapeutic benefit and guide the integration of *FGFR*-based combinations into future precision oncology strategies.

Acquired resistance mechanisms

Despite the encouraging results of *FGFR* inhibitors in clinical trials, both primary and acquired resistance remain significant obstacles to achieving durable clinical benefit. To date, most insights into mechanisms of resistance come from studies in *FGFR2* fusion-positive iCCA [49,54,59,72,103,104] and *FGFR3*-altered urothelial carcinoma [55,56,70]. Recent basket trials have also uncovered resistance-associated mutations in indications other than iCCA or urothelial cancer, including NSCLC as well as pancreatic and gastroesophageal cancers [67]. In a recent study, Tangermann *et al.* used saturation mutagenesis in NCI-H1581 NSCLC cells to systematically interrogate all 11,520 possible point mutations across the kinase domains of *FGFR1-4* [105]. A total of 738 mutations conferring resistance to pemigatinib and futibatinib were identified, capturing 97% of the resistance mutations observed in clinical trials and substantially expanding the catalog of potential resistance mechanisms that may emerge in patients.

Across *FGFR1-FGFR3*, the most frequent mechanisms of acquired resistance to noncovalent ATP-competitive *FGFR* inhibitors involve secondary mutations in the **gatekeeper** and **molecular brake residues** within the *FGFR* kinase domain, which lead to steric hindrance of inhibitor binding and destabilization of the inactive kinase conformation, respectively [84] (Table 1). In a considerable fraction of cases, multiple polyclonal intratumoral mutations have been observed [49,55,67,72,106]. Pathway reactivation through the upregulation or mutation of other RTKs or downstream signaling components, including RAS/rapidly accelerated fibrosarcoma (RAF)/ERK and PI3K/AKT/mTOR, represents alternative mechanisms underlying both intrinsic and acquired resistance to *FGFR* inhibitors [49,55,74,84].

Low clinically achievable drug concentrations due to adverse effects can create a permissive environment for the outgrowth of resistant clones, as observed with the FGFR inhibitor infigratinib [57]. Next-generation isoform-selective FGFR2 and FGFR3 inhibitors may be dosed at higher concentrations, resulting in enhanced efficacy and reduced emergence of resistance. Moreover, covalent FGFR inhibitors retain efficacy in *FGFR*-driven cancers that are resistant to noncovalent ATP-competitive FGFR inhibitors via acquired kinase domain mutations [107]. For example, the FGFR3-selective inhibitor LOXO-435 (LY3866288) has shown promising initial results in *FGFR3*-altered advanced solid tumors [78], with preclinical data indicating efficacy against the *FGFR3* V555M gatekeeper mutation [79]. Similarly, the covalent pan-FGFR inhibitor futibatinib has demonstrated efficacy in four iCCA patients with *FGFR2* fusions who developed resistance to the reversible ATP-competitive inhibitors infigratinib and zoligratinib [57]. Another promising example is lirafugratinib (RLY-4008), a highly selective, covalent FGFR2 inhibitor that retains antitumor activity against a subset of *FGFR2* resistance mutations in iCCA [75–77].

Concluding remarks and future perspectives

FGFR rearrangements, particularly in-frame gene fusions, have emerged as actionable oncogenic drivers in a subset of tumor indications. Several FGFR inhibitors have achieved regulatory approval based on their clinical benefits and have established *FGFR* rearrangements as predictive biomarkers. Nevertheless, major challenges remain (see **Outstanding questions**). Not all *FGFR* alterations carry equal clinical relevance; therefore, optimizing patient selection for FGFR-targeted therapies requires systematic evaluation of rare canonical and noncanonical rearrangements, which are increasingly uncovered through advances in NGS. Functional characterization of these *FGFR* rearrangements is essential to inform the rational design of next-generation targeted therapies, as only a subset act as true oncogenic drivers. Currently, integrating WGS with matched transcriptomic and/or proteomic data provides the most reliable means of identifying functionally expressed drivers, including those generated through noncanonical rearrangements involving pseudoexon, frameshift, or antisense events. However, robust computational frameworks capable of distinguishing true oncogenic driver events from passenger alterations need to be further developed and implemented into routine clinical practice. Current machine-learning models based on background breakpoint distribution analysis represent a promising start, but further refinement is needed, particularly through the incorporation of genomic features (e.g., replication timing, chromatin state) across all tissue types [45]. In parallel, the influence of co-occurring genetic alterations on *FGFR*-driven oncogenesis and therapeutic response needs to be further investigated. Finally, acquired resistance to FGFR inhibitors continues to pose a significant obstacle, underscoring the need for next-generation agents designed to prevent or overcome resistance, as well as to mitigate adverse events.

OUTSTANDING QUESTIONS

- What are the key determinants of oncogenicity in fibroblast growth factor receptor (FGFR) rearrangements?
- Which FGFR rearrangements act as standalone oncogenic drivers versus those that require cooperating co-driver alterations to achieve full oncogenicity?
- How do co-occurring (epi)genetic alterations influence *FGFR*-driven oncogenesis, therapeutic response, and resistance evolution?
- How can we systematically assess the clinical significance of rare canonical and non-canonical *FGFR* rearrangements?
- Which biomarkers can be robustly implemented to reliably predict patient response and resistance to FGFR-targeted therapies across tumor types?
- How should biomarker testing be standardized to ensure consistent detection of *FGFR* rearrangements across tumor types and clinical laboratories?
- How should basket studies be designed to effectively evaluate FGFR-targeted therapies across tumor types and (rare) *FGFR* variants?

GLOSSARY

- **Basket trial:** a clinical study design that evaluates the efficacy of a targeted therapy across multiple cancer types that share the same molecular alteration, rather than being restricted to a single tissue of origin. Patients are “pooled” into molecularly defined baskets (e.g., specific *FGFR* mutations or fusions).
- **Cholangiocarcinoma (CCA):** malignant tumor that originates from the epithelial cells of the bile ducts, occurring either within the liver (intrahepatic; iCCA) or along the extrahepatic biliary tree.
- **Chromoplexy:** complex DNA rearrangement event characterized by interdependent and balanced chromosomal translocations involving multiple chromosomes, creating chains of rearrangements that disrupt and fuse several genes simultaneously.
- **Chromothripsis:** catastrophic mutational process in which one or more chromosomes are shattered into tens to hundreds of fragments in a single cellular event, followed by error-prone DNA repair mechanisms that reassemble the fragments in random order, resulting in clustered complex chromosomal rearrangements.
- **Gatekeeper residue:** conserved amino acid within the kinase domain that sits at the entry of a hydrophobic pocket adjacent to the ATP-binding site. Its mutation can modify pocket accessibility and thereby alter kinase activity and sensitivity to reversible ATP-competitive inhibitors.
- **Microhomology-mediated end joining:** error-prone DNA double-strand break repair pathway that aligns and joins broken DNA ends using short microhomologous sequences (typically 5-25 base pairs), usually resulting in deletions at the repair site.

- **Molecular brake residues:** conserved amino acid residues within the kinase domain that stabilize its inactive conformation and limit basal kinase activity; their mutation leads to increased kinase activation.
- **Nonhomologous end joining:** error-prone DNA double-strand break repair mechanism that directly ligates the broken DNA ends together without requiring an homologous template.
- **Overall survival:** length of time from the start of treatment or randomization in a clinical trial until death from any cause, considered the gold standard endpoint for measuring treatment benefit in oncology.
- **Progression-free survival (PFS):** length of time during and after treatment in a clinical trial that a patient lives with the disease without it worsening, measured from treatment start until disease progression or death from any cause, and is commonly used as a surrogate endpoint to assess the effectiveness of targeted therapies in oncology trials.

ACKNOWLEDGMENTS

We apologize to those authors whose work could not be cited due to space constraints. Research at the Netherlands Cancer Institute is supported by international grants from the Dutch Cancer Society (KWF) and the Dutch Ministry of Health, Welfare and Sport. Research in the Jonkers laboratory is funded by the Onco Institute and the Dutch Cancer Society (KWF grant 12894 to J.J. and D.Z.)

DECLARATION OF INTERESTS

D.Z. is an employee of Genentech, Inc. and holds equity in Roche. The other authors declare no competing interests.

REFERENCES

1. Shing, Y. *et al.* (1985) Angiogenesis is stimulated by a tumor-derived endothelial cell growth factor. *J Cell Biochem* 29, 275–287
2. Baird, A. *et al.* (1986) Molecular characterization of fibroblast growth factor: distribution and biological activities in various tissues. *Recent Prog Horm Res* 42, 143–205
3. Presta, M. *et al.* (1986) Purification from a human hepatoma cell line of a basic fibroblast growth factor-like molecule that stimulates capillary endothelial cell plasminogen activator production, DNA synthesis, and migration. *Mol Cell Biol* 6, 4060–4066
4. Adnane, J. *et al.* (1991) BEK and FLG, two receptors to members of the FGF family, are amplified in subsets of human breast cancers. *Oncogene* 6, 659–663
5. Wu, Y.M. *et al.* (2013) Identification of targetable FGFR gene fusions in diverse cancers. *Cancer discovery* 3, 636–647
6. Zingg, D. *et al.* (2022) Truncated FGFR2 is a clinically actionable oncogene in multiple cancers. *Nature* 608, 609–617
7. Yemelyanenko, J. *et al.* (2025) C-terminal Truncation and Fusion Partner Determine Oncogenicity of FGFR3. *Cancer Res* DOI: 10.1158/0008-5472.CAN-24-2648
8. Beenken, A. and Mohammadi, M. (2009) The FGF family: biology, pathophysiology and therapy. *Nat Rev Drug Discov* 8, 235–253
9. Ornitz, D.M. and Itoh, N. (2015) The Fibroblast Growth Factor signaling pathway. *WIREs Dev Biol* 4, 215–266
10. Xie, Y. *et al.* (2020) FGF/FGFR signaling in health and disease. *Sig Transduct Target Ther* 5, 1–38
11. Wiedemann, M. and Trueb, B. (2000) Characterization of a Novel Protein (FGFRL1) from Human Cartilage Related to FGF Receptors. *Genomics* 69, 275–279
12. Trueb, B. (2011) Biology of FGFRL1, the fifth fibroblast growth factor receptor. *Cell Mol Life Sci* 68, 951–964
13. Chen, L. *et al.* (2021) Fibroblast growth factor receptor fusions in cancer: opportunities and challenges. *Journal of Experimental & Clinical Cancer Research* 40, 345
14. Thisse, B. and Thisse, C. (2005) Functions and regulations of fibroblast growth factor signaling during embryonic development. *Developmental Biology* 287, 390–402
15. Szybowska, P. *et al.* (2021) Negative Regulation of FGFR (Fibroblast Growth Factor Receptor) Signaling. *Cells* 10, 1342
16. Babina, I.S. and Turner, N.C. (2017) Advances and challenges in targeting FGFR signalling in cancer. *Nature Reviews Cancer* 17, 318–332
17. Turner, N. *et al.* (2010) FGFR1 amplification drives endocrine therapy resistance and is a therapeutic target in breast cancer. *Cancer Res* 70, 2085–2094
18. Hanker, A.B. *et al.* (2017) HER2-Overexpressing Breast Cancers Amplify FGFR Signaling upon Acquisition of Resistance to Dual Therapeutic Blockade of HER2. *Clin Cancer Res* 23, 4323–4334
19. Formisano, L. *et al.* (2019) Aberrant FGFR signaling mediates resistance to CDK4/6 inhibitors in ER+ breast cancer. *Nat Commun* 10, 1373
20. Zhou, Y. *et al.* (2020) FGF/FGFR signaling pathway involved resistance in various cancer types. *J Cancer* 11, 2000–2007
21. Mao, P. *et al.* (2020) Acquired FGFR and FGF Alterations Confer Resistance to Estrogen Receptor (ER) Targeted Therapy in ER+ Metastatic Breast Cancer. *Clin Cancer Res* 26, 5974–5989
22. Servetto, A. *et al.* (2021) Nuclear FGFR1 Regulates Gene Transcription and Promotes Antiestrogen Resistance in ER+ Breast Cancer. *Clin Cancer Res* 27, 4379–4396

23. Clayton, N.S. *et al.* (2017) Fibroblast growth factor-mediated crosstalk in cancer etiology and treatment. *Dev Dyn* 246, 493–501
24. Helsten, T. *et al.* (2016) The FGFR landscape in cancer: Analysis of 4,853 tumors by next-generation sequencing. *Clinical Cancer Research* 22, 259–267
25. Krook, M.A. *et al.* (2020) Pan-cancer analysis of FGFR1-3 genomic alterations to reveal a complex molecular landscape. *JCO* 38, 3620–3620
26. Zuo, W. *et al.* (2020) Landscape of FGF/FGFR Alterations in 12,372 Chinese Cancer Patients. *Journal of Cancer* 11, 6695–6699
27. Sun, Y. *et al.* (2020) A comprehensive pan-cancer study of fibroblast growth factor receptor aberrations in Chinese cancer patients. *Annals of Translational Medicine* 8, 1290–1290
28. Murugesan, K. *et al.* (2022) Pan-tumor landscape of fibroblast growth factor receptor 1-4 genomic alterations. *ESMO Open* 7, 100641
29. Bailey, C. *et al.* (2024) Origins and impact of extrachromosomal DNA. *Nature* 635, 193–200
30. Willis, N.A. *et al.* (2015) Deciphering the Code of the Cancer Genome: Mechanisms of Chromosome Rearrangement. *Trends Cancer* 1, 217–230
31. Krupina, K. *et al.* (2024) Scrambling the genome in cancer: causes and consequences of complex chromosome rearrangements. *Nat Rev Genet* 25, 196–210
32. Krook, M.A. *et al.* (2021) Fibroblast growth factor receptors in cancer: genetic alterations, diagnostics, therapeutic targets and mechanisms of resistance. *Br J Cancer* 124, 880–892
33. Singh, D. *et al.* (2012) Transforming fusions of FGFR and TACC genes in human glioblastoma. *Science* 337, 1231–1235
34. Qin, A. *et al.* (2019) Detection of Known and Novel FGFR Fusions in Non–Small Cell Lung Cancer by Comprehensive Genomic Profiling. *Journal of Thoracic Oncology* 14, 54–62
35. Stein, L. *et al.* (2024) FGFR2-fusions define a clinically actionable molecular subset of pancreatic cancer. *NPJ Precis Oncol* 8, 207
36. Li, I. *et al.* (2025) Fibroblast growth factor receptor alterations and resistance mechanisms in the treatment of pediatric solid tumors. *cdr* 8, N/A–N/A
37. Wang, R. *et al.* (2014) FGFR1/3 tyrosine kinase fusions define a unique molecular subtype of non-small cell lung cancer. *Clinical Cancer Research* 20, 4107–4114
38. Hoskins, E.L. *et al.* (2025) Prevalence and biological impact of clinically relevant gene fusions in head and neck cancers. *npj Precis. Onc.* 9, 221
39. Ou, S.-H.I. *et al.* (2017) Emergence of FGFR3-TACC3 fusions as a potential by-pass resistance mechanism to EGFR tyrosine kinase inhibitors in EGFR mutated NSCLC patients. *Lung Cancer* 111, 61–64
40. Schrock, A.B. *et al.* (2018) Receptor Tyrosine Kinase Fusions and BRAF Kinase Fusions are Rare but Actionable Resistance Mechanisms to EGFR Tyrosine Kinase Inhibitors. *Journal of Thoracic Oncology* 13, 1312–1323
41. Raphael, A. *et al.* (2022) FGFR Fusions as an Acquired Resistance Mechanism Following Treatment with Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitors (EGFR TKIs) and a Suggested Novel Target in Advanced Non-Small Cell Lung Cancer (aNSCLC). *Journal of Clinical Medicine* 11, 2475
42. Gu, W. *et al.* (2021) Comprehensive identification of FGFR1-4 alterations in 5 557 Chinese patients with solid tumors by next-generation sequencing. *Am J Cancer Res* 11, 3893–3906
43. Yun, J.W. *et al.* (2020) Dysregulation of cancer genes by recurrent intergenic fusions. *Genome Biology* 21, 166
44. Yao, Y. *et al.* (2022) Characterizing kinase intergenic-breakpoint rearrangements in a large-scale lung cancer population and real-world clinical outcomes. *ESMO Open* 7, 100405

45. Seo, E. *et al.* (2025) Decoding Genomic Rearrangements in Cancer: Path to Novel Driver Discovery. *Molecules and Cells* DOI: 10.1016/j.mocell.2025.100272
46. Malchers, F. *et al.* (2023) Somatic rearrangements causing oncogenic ectodomain deletions of FGFR1 in squamous cell lung cancer. *J Clin Invest* 133, e170217
47. Zingg, D. *et al.* (2025) The C-terminal Kinase Domain-Binding and Suppression Motif Prevents Constitutive Activation of FGFR2. *Cancer Research* DOI: 10.1158/0008-5472.CAN-24-3349
48. Gallo, L.H. *et al.* (2015) Functions of Fibroblast Growth Factor Receptors in cancer defined by novel translocations and mutations. *Cytokine & Growth Factor Reviews* 26, 425–449
49. Wu, Q. *et al.* (2024) Landscape of Clinical Resistance Mechanisms to FGFR Inhibitors in FGFR2-Altered Cholangiocarcinoma. *Clin Cancer Res* 30, 198–208
50. Loriot, Y. *et al.* (2019) Erdafitinib in Locally Advanced or Metastatic Urothelial Carcinoma. *New England Journal of Medicine* 381, 338–348
51. Siefker-Radtke, A.O. *et al.* (2022) Efficacy and safety of erdafitinib in patients with locally advanced or metastatic urothelial carcinoma: long-term follow-up of a phase 2 study. *The Lancet Oncology* 23, 248–258
52. Loriot, Y. *et al.* (2023) Erdafitinib or Chemotherapy in Advanced or Metastatic Urothelial Carcinoma. *New England Journal of Medicine* 389, 1961–1971
53. Siefker-Radtke, A.O. *et al.* (2024) Erdafitinib versus pembrolizumab in pretreated patients with advanced or metastatic urothelial cancer with select FGFR alterations: cohort 2 of the randomized phase III THOR trial☆. *Annals of Oncology* 35, 107–117
54. Facchinetti, F. *et al.* (2024) Understanding and Overcoming Resistance to Selective FGFR Inhibitors across FGFR2-Driven Malignancies. *Clin Cancer Res* 30, 4943–4956
55. Facchinetti, F. *et al.* (2023) Resistance to Selective FGFR Inhibitors in FGFR-Driven Urothelial Cancer. *Cancer Discovery* 13, 1998–2011
56. Guercio, B.J. *et al.* (2023) Clinical and Genomic Landscape of FGFR3-Altered Urothelial Carcinoma and Treatment Outcomes with Erdafitinib: A Real-World Experience. *Clin Cancer Res* 29, 4586–4595
57. Goyal, L. *et al.* (2019) TAS-120 Overcomes Resistance to ATP-Competitive FGFR Inhibitors in Patients with FGFR2 Fusion-Positive Intrahepatic Cholangiocarcinoma. *Cancer Discov* 9, 1064–1079
58. Goyal, L. *et al.* (2023) Futibatinib for FGFR2-Rearranged Intrahepatic Cholangiocarcinoma. *N Engl J Med* 388, 228–239
59. Goyal, L. *et al.* (2025) A model for decoding resistance in precision oncology: acquired resistance to FGFR inhibitors in cholangiocarcinoma. *Annals of Oncology* 36, 426–443
60. Sootome, H. *et al.* (2020) Futibatinib Is a Novel Irreversible FGFR 1-4 Inhibitor That Shows Selective Antitumor Activity against FGFR-Deregulated Tumors. *Cancer Res* 80, 4986–4997
61. Koshkin, V.S. *et al.* (2024) 1965MO Phase II study of futibatinib plus pembrolizumab in patients (pts) with advanced/metastatic urothelial carcinoma (mUC): Final analysis of efficacy and safety. *Annals of Oncology* 35, S1136–S1137
62. Balasooriya, E.R. *et al.* (2024) The irreversible FGFR inhibitor KIN-3248 overcomes FGFR2 kinase domain mutations. *Clin Cancer Res* 30, 2181–2192
63. Abou-Alfa, G.K. *et al.* (2020) Pemigatinib for previously treated, locally advanced or metastatic cholangiocarcinoma: a multicentre, open-label, phase 2 study. *Lancet Oncol* 21, 671–684
64. Vogel, A. *et al.* (2024) An open-label study of pemigatinib in cholangiocarcinoma: final results from FIGHT-202☆. *ESMO Open* 9, 103488
65. Gotlib, J. *et al.* (2021) A Phase 2 Study of Pemigatinib (FIGHT-203; INCB054828) in Patients with Myeloid/Lymphoid Neoplasms (MLNs) with Fibroblast Growth Factor Receptor 1 (FGFR1) Rearrangement (MLN FGFR1). *Blood* 138, 385
66. Verstovsek, S. *et al.* (2025) Pemigatinib for Myeloid/Lymphoid Neoplasms with FGFR1 Rearrangement. *NEJM Evid* 4, EVIDoa2500017

67. Rodón, J. *et al.* (2024) Pemigatinib in previously treated solid tumors with activating FGFR1–FGFR3 alterations: phase 2 FIGHT-207 basket trial. *Nat Med* DOI: 10.1038/s41591-024-02934-7
68. Silverman, I.M. *et al.* (2021) Clinicogenomic Analysis of FGFR2-Rearranged Cholangiocarcinoma Identifies Correlates of Response and Mechanisms of Resistance to Pemigatinib. *Cancer Discov* 11, 326–339
69. Javle, M. *et al.* (2021) Infigratinib (BGJ398) in previously treated patients with advanced or metastatic cholangiocarcinoma with FGFR2 fusions or rearrangements: mature results from a multicentre, open-label, single-arm, phase 2 study. *Lancet Gastroenterol Hepatol* 6, 803–815
70. Pal, S.K. *et al.* (2018) Efficacy of BGJ398, a Fibroblast Growth Factor Receptor 1–3 Inhibitor, in Patients with Previously Treated Advanced Urothelial Carcinoma with FGFR3 Alterations. *Cancer Discov* 8, 812–821
71. Cowell, J.K. *et al.* (2017) Mutation in the FGFR1 tyrosine kinase domain or inactivation of PTEN is associated with acquired resistance to FGFR inhibitors in FGFR1-driven leukemia/lymphomas. *Int J Cancer* 141, 1822–1829
72. Goyal, L. *et al.* (2017) Polyclonal Secondary FGFR2 Mutations Drive Acquired Resistance to FGFR Inhibition in Patients with FGFR2 Fusion-Positive Cholangiocarcinoma. *Cancer Discov* 7, 252–263
73. Ryan, M.R. *et al.* (2019) The FGFR1 V561M Gatekeeper Mutation Drives AZD4547 Resistance through STAT3 Activation and EMT. *Mol Cancer Res* 17, 532–543
74. Kas, S.M. *et al.* (2018) Transcriptomics and transposon mutagenesis identify multiple mechanisms of resistance to the FGFR inhibitor AZD4547. *Cancer Research* 78, 5668–5679
75. Subbiah, V. *et al.* (2023) RLY-4008, the First Highly Selective FGFR2 Inhibitor with Activity across FGFR2 Alterations and Resistance Mutations. *Cancer Discov* 13, 2012–2031
76. Schram, A.M. *et al.* (2023) Abstract IA006: Clinical activity of lirafugratinib (RLY-4008), a highly selective FGFR2 inhibitor, in patients with advanced FGFR2-altered solid tumors: The ReFocus study. *Mol Cancer Ther* 22, IA006
77. Schönherr, H. *et al.* (2024) Discovery of lirafugratinib (RLY-4008), a highly selective irreversible small-molecule inhibitor of FGFR2. *Proceedings of the National Academy of Sciences* 121, e2317756121
78. Iyer, G. *et al.* (2025) A first-in-human phase 1 study of LY3866288 (LOXO-435), a potent, highly isoform-selective FGFR3 inhibitor (FGFR3i) in advanced solid tumors with FGFR3 alterations: Initial results from FORAGER-1. *Journal of Clinical Oncology* DOI: 10.1200/JCO.2025.43.5_suppl.662
79. Ballard, J.A. *et al.* (2021) Abstract P141: Preclinical characterization of LOX-24350, a highly potent and isoform-selective FGFR3 inhibitor. *Mol Cancer Ther* 20, P141
80. Fairhurst, R.A. *et al.* (2020) Discovery of Roblitinib (FGF401) as a Reversible-Covalent Inhibitor of the Kinase Activity of Fibroblast Growth Factor Receptor 4. *J Med Chem* 63, 12542–12573
81. Weiss, A. *et al.* (2019) FGF401, A First-In-Class Highly Selective and Potent FGFR4 Inhibitor for the Treatment of FGF19-Driven Hepatocellular Cancer. *Mol Cancer Ther* 18, 2194–2206
82. Hatlen, M.A. *et al.* (2019) Acquired On-Target Clinical Resistance Validates FGFR4 as a Driver of Hepatocellular Carcinoma. *Cancer Discov* 9, 1686–1695
83. Katoh, M. (2019) Fibroblast growth factor receptors as treatment targets in clinical oncology. *Nat Rev Clin Oncol* 16, 105–122
84. Katoh, M. *et al.* (2024) FGFR-targeted therapeutics: clinical activity, mechanisms of resistance and new directions. *Nat Rev Clin Oncol* DOI: 10.1038/s41571-024-00869-z
85. Subbiah, V. and Verstovsek, S. (2023) Clinical development and management of adverse events associated with FGFR inhibitors. *Cell Rep Med* 4, 101204
86. Qu, L. *et al.* (2022) Structural insights into the potency and selectivity of covalent pan-FGFR inhibitors. *Commun Chem* 5, 5
87. Zingg, D. *et al.* (2025) The C-Terminal Kinase Domain–Binding and Suppression Motif Prevents Constitutive Activation of FGFR2. *Cancer Res* 85, 3234–3257

88. Kendre, G. *et al.* (2023) Charting co-mutation patterns associated with actionable drivers in intrahepatic cholangiocarcinoma. *Journal of Hepatology* 78, 614–626
89. Uehara, Y. *et al.* (2022) Targeting the FGF/FGFR axis and its co-alteration allies. *ESMO Open* 7, 100647
90. Chen, L. (2025) Next-generation isoform-selective fibroblast growth factor receptor inhibitors. *Trends in Pharmacological Sciences* 46, 1091–1104
91. Wainberg, Z.A. *et al.* (2024) Bemarituzumab as first-line treatment for locally advanced or metastatic gastric/gastroesophageal junction adenocarcinoma: final analysis of the randomized phase 2 FIGHT trial. *Gastric Cancer* 27, 558–570
92. Fu, Y. *et al.* (2025) Abstract 2964: SCR-A002, a novel FGFR2b-targeting ADC for solid tumors. *Cancer Res* 85, 2964
93. Apfelbaum, A.A. *et al.* (2025) A diverse landscape of FGFR alterations and co-mutations suggests potential therapeutic strategies in pediatric low-grade gliomas. *Nat Commun* 16, 7018
94. Juric, D. *et al.* (2025) PI3K α inhibitor and degrader inavolisib can co-opt FGFR2 to enhance responses in patients with PIK3CA-mutated solid tumors and in preclinical models. *Clin Cancer Res* DOI: 10.1158/1078-0432.CCR-25-1459
95. D'Souza, N. *et al.* (2024) Computational patterns in FGFR3-altered non-muscle invasive bladder cancer influence recurrence and progression. *Urologic Oncology: Seminars and Original Investigations* 42, S55
96. Packer, L.M. *et al.* (2017) PI3K Inhibitors Synergize with FGFR Inhibitors to Enhance Antitumor Responses in FGFR2mutant Endometrial Cancers. *Mol Cancer Ther* 16, 637–648
97. Holzhauser, S. *et al.* (2019) In vitro antitumor effects of FGFR and PI3K inhibitors on human papillomavirus positive and negative tonsillar and base of tongue cancer cell lines. *Oncol Lett* 18, 6249–6260
98. Holzhauser, S. *et al.* (2020) Targeting Fibroblast Growth Factor Receptor (FGFR) and Phosphoinositide 3-kinase (PI3K) Signaling Pathways in Medulloblastoma Cell Lines. *Anticancer Research* 40, 53–66
99. Lukoseviciute, M. *et al.* (2024) Combined targeted therapy with PI3K and CDK4/6, or FGFR inhibitors show synergistic effects in a neuroblastoma spheroid culture model. *Biomed Pharmacother* 177, 116993
100. Nakamura, R. *et al.* (2024) Triple combination therapy comprising osimertinib, an AXL inhibitor, and an FGFR inhibitor improves the efficacy of EGFR-mutated non-small cell lung cancer. *Cancer Letters* 598, 217124
101. Liu, K. *et al.* (2025) Significant Response to Palbociclib Plus Lenvatinib as Second-line Treatment for CDKN2A/2B Deletion Intrahepatic Cholangiocarcinoma: A Case Report. *J Clin Transl Hepatol* 13, 169–172
102. Zhang, Z. *et al.* (2023) Case Report: Persistent response to combination therapy of pemigatinib, chemotherapy, and immune checkpoint inhibitor in a patient with advanced intrahepatic cholangiocarcinoma. *Front. Immunol.* 14
103. Krook, M.A. *et al.* (2019) Tumor heterogeneity and acquired drug resistance in FGFR2-fusion-positive cholangiocarcinoma through rapid research autopsy. *Cold Spring Harb Mol Case Stud* 5, a004002
104. Krook, M.A. *et al.* (2020) Efficacy of FGFR Inhibitors and Combination Therapies for Acquired Resistance in FGFR2-Fusion Cholangiocarcinoma. *Mol Cancer Ther* 19, 847–857
105. Tangermann, C. *et al.* (2025) Saturation mutagenesis identifies activating and resistance-inducing FGFR kinase domain mutations. *Nat Genet* DOI: 10.1038/s41588-025-02431-8
106. Goyal, L. *et al.* (2025) Genomic correlates of response and resistance to the irreversible FGFR1-4 inhibitor futibatinib based on biopsy and circulating tumor DNA profiling. *Annals of Oncology* 36, 414–425
107. Chen, L. (2025) Next-generation isoform-selective fibroblast growth factor receptor inhibitors. *Trends Pharmacol Sci* 46, 1091–1104

3

CHAPTER 3

TRUNCATED *FGFR2* IS A CLINICALLY ACTIONABLE ONCOGENE IN MULTIPLE CANCERS

Daniel Zingg*, Jinhyuk Bhin*, **Julia Yemelyanenko***, Sjors M. Kas*, Frank Rolfs, Catrin Lutz, Jessica K. Lee, Sjoerd Klarenbeek, Ian M. Silverman, Stefano Annunziato, Chang S. Chan, Sander R. Piersma, Timo Eijkman, Madelon Badoux, Ewa Gogola, Bjørn Siteur, Justin Sprengers, Bim de Klein, Richard R. de Goeij-de Haas, Gregory M. Riedlinger, Hua Ke, Russell Madison, Anne Paulien Drenth, Eline van der Burg, Eva Schut, Linda Henneman, Martine H. van Miltenburg, Natalie Proost, Huiling Zhen, Ellen Wientjens, Roebi de Bruijn, Julian R. de Ruiter, Ute Boon, Renske de Korte-Grimmerink, Bastiaan van Gerwen, Luis Féliz, Ghassan K. Abou-Alfa, Jeffrey S. Ross, Marieke van de Ven, Sven Rottenberg, Edwin Cuppen, Anne Vaslin Chessex, Siraj M. Ali, Timothy C. Burn, Connie R. Jimenez, Shridar Ganesan#, Lodewyk F. A. Wessels# & Jos Jonkers#

*equal contribution

#co-corresponding author

Nature, 2022 Aug;608(7923):609-617

<https://doi.org/10.1038/s41586-022-05066-5>

ABSTRACT

Somatic hotspot mutations and structural amplifications and fusions that affect fibroblast growth factor receptor 2 (encoded by *FGFR2*) occur in multiple types of cancer [1]. However, clinical responses to FGFR inhibitors have remained variable [1–9], emphasizing the need to better understand which *FGFR2* alterations are oncogenic and therapeutically targetable. Here we apply transposon-based screening [10,11] and tumour modelling in mice [12,13], and find that the truncation of exon 18 (E18) of *Fgfr2* is a potent driver mutation. Human oncogenomic datasets revealed a diverse set of *FGFR2* alterations, including rearrangements, E1–E17 partial amplifications, and E18 nonsense and frameshift mutations, each causing the transcription of E18-truncated *FGFR2* (*FGFR2*^{ΔE18}). Functional in vitro and in vivo examination of a compendium of *FGFR2*^{ΔE18} and full-length variants pinpointed *FGFR2*-E18 truncation as single-driver alteration in cancer. By contrast, the oncogenic competence of *FGFR2* full-length amplifications depended on a distinct landscape of cooperating driver genes. This suggests that genomic alterations that generate stable *FGFR2*^{ΔE18} variants are actionable therapeutic targets, which we confirmed in preclinical mouse and human tumour models, and in a clinical trial. We propose that cancers containing any *FGFR2* variant with a truncated E18 should be considered for FGFR-targeted therapies.

MAIN

FGFR2 is a receptor tyrosine kinase (RTK) that consists of an extracellular ligand-binding domain, intracellular tyrosine kinase domains and a carboxy (C)-terminal tail relevant for receptor activity fine-tuning [14]. In human cancers, *FGFR2* can be affected by hotspot mutations and structural variants, namely fusions and amplifications [1], some of which produce truncated FGFR2 isoforms [15–19]. *FGFR2* structural variants have been considered to be oncogenic and actionable due to the resulting overexpression and increased stabilization of the receptor [2,20–22]. However, in patients with cancer with such structural variants, ATP-competitive small-molecule inhibitors targeting FGFRs have produced inconsistent clinical benefits [1–9]. A better understanding of the determinants defining the oncogenicity and clinical actionability of *FGFR2* structural variants is therefore critical for precise matching of cancer patients to FGFR-targeted therapies.

A *SB* screen identified *Fgfr2*^{ΔE18}

Sleeping Beauty (*SB*) transposon-based insertional mutagenesis screening has revealed potential tumour drivers in mice through transcriptional activation and/or truncation of target genes, and identified *Fgfr2* as a top candidate driver in mammary tumorigenesis [10,11] (Fig. 1a,b). Mapping of the *SB* insertions in *Fgfr2* showed strong enrichment for insertions in intron 17 (I17; Fig. 1c and Extended Data Fig. 1a). Analysis of RNA sequencing (RNA-seq) data of *SB* tumours revealed that *Fgfr2*-I17 insertions enforce splicing of *Fgfr2*-E17 into the transposon. This led to *Fgfr2* transcripts that lacked E17–E18 spanning reads, and expression of *Fgfr2*^{ΔE18} was confirmed by reverse transcription with quantitative PCR (RT-qPCR; Fig. 1d,e and Extended Data Fig. 1b–d). Tumours with *SB* insertions in the 5' region of *Fgfr2* either contained a second *SB* insertion in I17 or contained rearrangements (REs) in *Fgfr2*-I17, producing gene fusions [11] and therefore also expressing *Fgfr2*^{ΔE18} (Extended Data Fig. 1d). E18 of both mouse and human *FGFR2* encodes the C terminus of this RTK [14]. We observed an overall upregulation of *Fgfr2* transcripts in tumours with *SB* insertions (Extended Data Fig. 1e), suggesting a loss of regulatory elements that are presumably encoded by the *Fgfr2* 3'-untranslated region (3'-UTR; ref. [21]) and/or positive oncogenic selection of C-terminally truncated FGFR2.

FGFR2^{ΔE18} variants in human cancer

To assess whether genomic alterations producing *FGFR2*^{ΔE18} occur in human cancers, we first analysed whole-genome sequencing (WGS) data of metastatic solid tumours from the Hartwig Medical Foundation (HMF; ref. [23]). Examination of structural variants affecting *FGFR2* in 2,112 HMF WGS profiles revealed a significant enrichment of RE breakpoints (BPs) in I17 (Fig. 1f and Extended Data Fig. 1f), coinciding with reported *FGFR2* fusion BPs [20–22]. Recurring chromosomal REs, such as breakage–fusion–bridge cycles, can produce focal *FGFR2* amplifications (*FGFR2*^{amp}; refs. [16,17,24]), which we observed in a fraction of tumours with *FGFR2* REs (Extended Data Fig. 1g). Some *FGFR2*-I17 REs implicated canonical in-frame

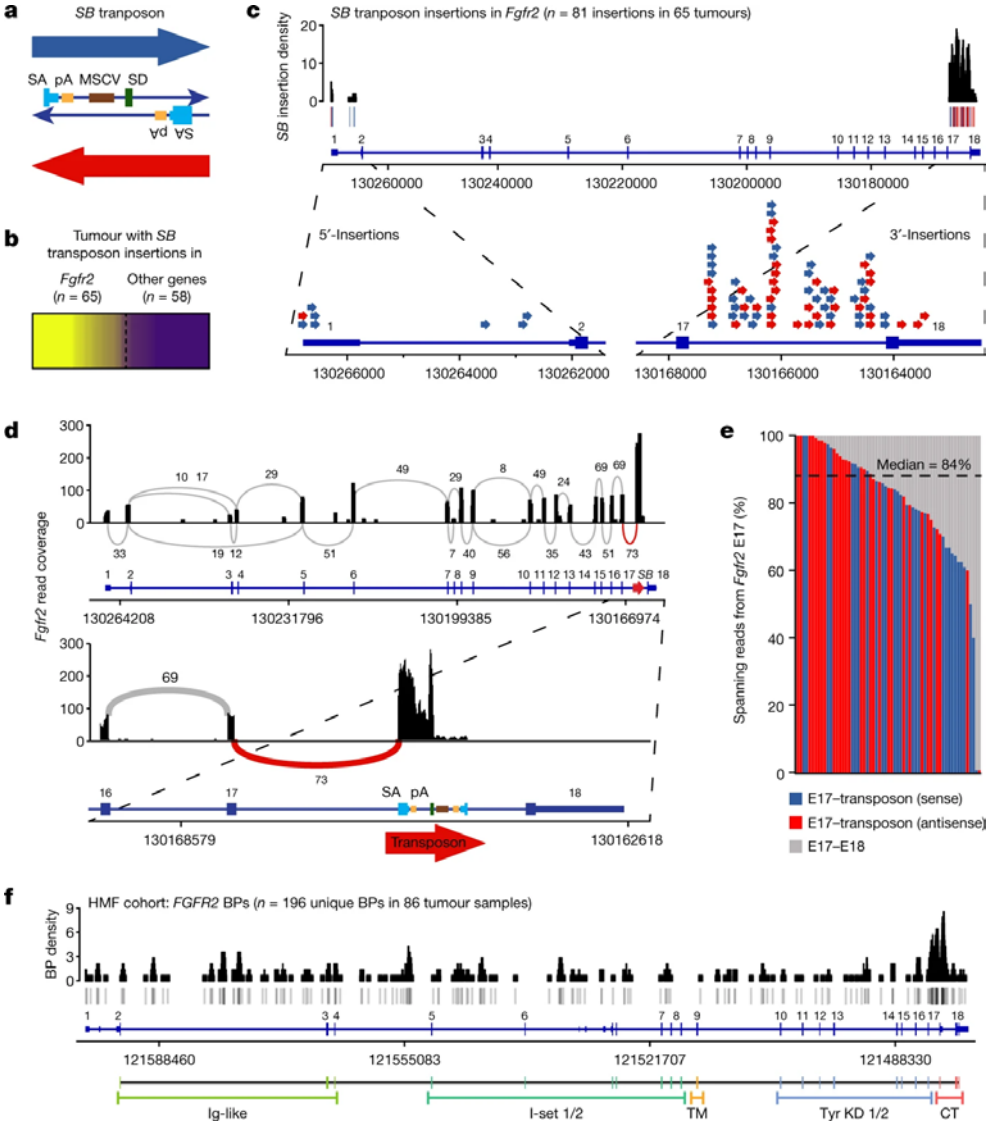


Fig. 1: SB transposon screen and WGS analysis identifies recurrent *FGFR2* E18 truncation.

a, Schematic of the SB transposon, which encodes splice acceptors (SA) followed by polyadenylation (pA) signals on both strands and a murine stem cell virus (MSCV) promoter followed by a splice donor (SD) on the plus strand. **b**, Mammary tumours with SB transposon insertions in *Fgfr2* as identified in an insertional mutagenesis screen [10]. The relative clonality of SB insertions in *Fgfr2* is shown by a colour gradient (yellow to purple, clonality of 1 to 0). SB insertions in *Fgfr2* were called for tumours with a *Fgfr2* relative insertion clonality of ≥ 0.25 . **c**, The SB transposon insertions found in *Fgfr2* (chromosome 7). The SB insertion density was calculated using a 500 bp sliding window. The blue bars/arrows show sense SB insertions; the red bars/arrows show antisense SB insertions. **d**, Sashimi plot showing *Fgfr2* read coverage and junction reads plotted as arcs with the indicated junction read counts of a tumour with an I17 antisense SB insertion. **e**, The ratio of spanning reads from *Fgfr2*-E17 to E18 versus SB transposon insertion in tumours with I17 SB insertions. $n = 31$ (sense) and $n = 30$ (antisense). **f**, BPs generating *FGFR2* (chromosome 10) genomic REs identified in 86 out of 2,112 analysed WGS profiles from the HMF cohort on metastatic solid tumours [23]. $n = 266$ (total) and $n = 196$ (unique) BPs. BP density was calculated using a 500 bp sliding window. The grey bars show BPs. Corresponding protein domains are indicated. CT, C terminus; TM, transmembrane; Tyr KD, tyrosine kinase domain.

fusions with 3'-partner genes, but we also found non-canonical REs in which the reading frame of the partner gene was undeterminable (frame unknown), the partner gene was out of strand, the partner sequence was derived from intergenic space or the REs occurred internally in *FGFR2* (Extended Data Fig. 1h and Supplementary Table 1).

We next analysed oncogenomic data from Foundation Medicine (FMI) derived from 249,570 targeted tumour sequencing assays for the occurrence of *FGFR2* alterations. Across cancers, we identified 1,367 samples with alterations potentially producing *FGFR2*^{ΔE18} (0.55% incidence). These were mutually exclusive to samples with *FGFR2*^{amp} ($n=838$, 0.34% incidence) or *FGFR2* missense hotspot mutations (*FGFR2*^{hotspot}; $n=978$, 0.39% incidence; Fig. 2a and Extended Data Fig. 2a). Alterations potentially perturbing E18 were made up of 55.4% *FGFR2*-I17/E18 in-frame fusions. The remaining 44.6% were classified as variants of unknown significance and comprised *FGFR2*-I17/E18 frame-unknown, out-of-strand, intergenic space and internal REs, some of which coincided with focal amplifications (Fig. 2a, Extended Data Fig. 2b and Supplementary Table 2). Across the *FGFR2* REs, intrachromosomal REs and known fusion partner genes (for example, *BICC1*, *TACC2* and *ATE1*; refs. [20–22]) were enriched (Extended Data Fig. 2c,d). *FGFR2* RE partner genes encoded 337 unique proteins. Among these, the ability to self-interact was enriched as compared to the human proteome (Extended Data Fig. 2e–g). Nevertheless, 42.8% of all REs made use of a partner without evident self-interaction ability (Extended Data Fig. 2h). Other variants of unknown significance were *FGFR2*-E1–E17 partial amplifications, E18 splice-acceptor-site mutations and protein-truncating mutations significantly enriched in the coding sequence of *FGFR2*-E18 (Fig. 2a and Extended Data Fig. 2i–k). The identified *FGFR2*^{ΔE18} variants were most frequent in cholangiocarcinoma, but we also found considerable frequencies of in-frame fusions and especially structural variants of unknown significance in gastroesophageal and breast cancer (Extended Data Fig. 2l).

Expression of *FGFR2*^{ΔE18} in human cancer

To validate expression of *FGFR2*^{ΔE18} variants, we analysed RNA-seq profiles matched to the HMF WGS samples. In the majority of the cases in which RNA-seq data was available, the predicted *FGFR2* RE types were robustly expressed; REs with intergenic space produced *FGFR2* transcripts terminating in intergenic region (IGR) pseudoexons encoding splice acceptors, a coding sequence and stop codons (Extended Data Figs. 1h and 3a–c). We also observed splicing to an alternative *FGFR2*-E18, termed C3 (Extended Data Fig. 3b–e), which is located in I17 and encodes a single isoleucine followed by a 3'-UTR. Two more *FGFR2* isoforms make use of an alternative E18. The encoded C termini either overlap with the proximal part of the canonical *FGFR2* C terminus (C2) or are different to it (C4; refs. [16–18]). Thus, splicing to E18-C3 or E18-C4 generates *FGFR2* isoforms that encode dysfunctional C termini resembling E18 truncation (Extended Data Fig. 2k). In a few cases with IGR REs, we found *FGFR2* in-frame fusions at the RNA level. Reconstruction of derivative chromosomes

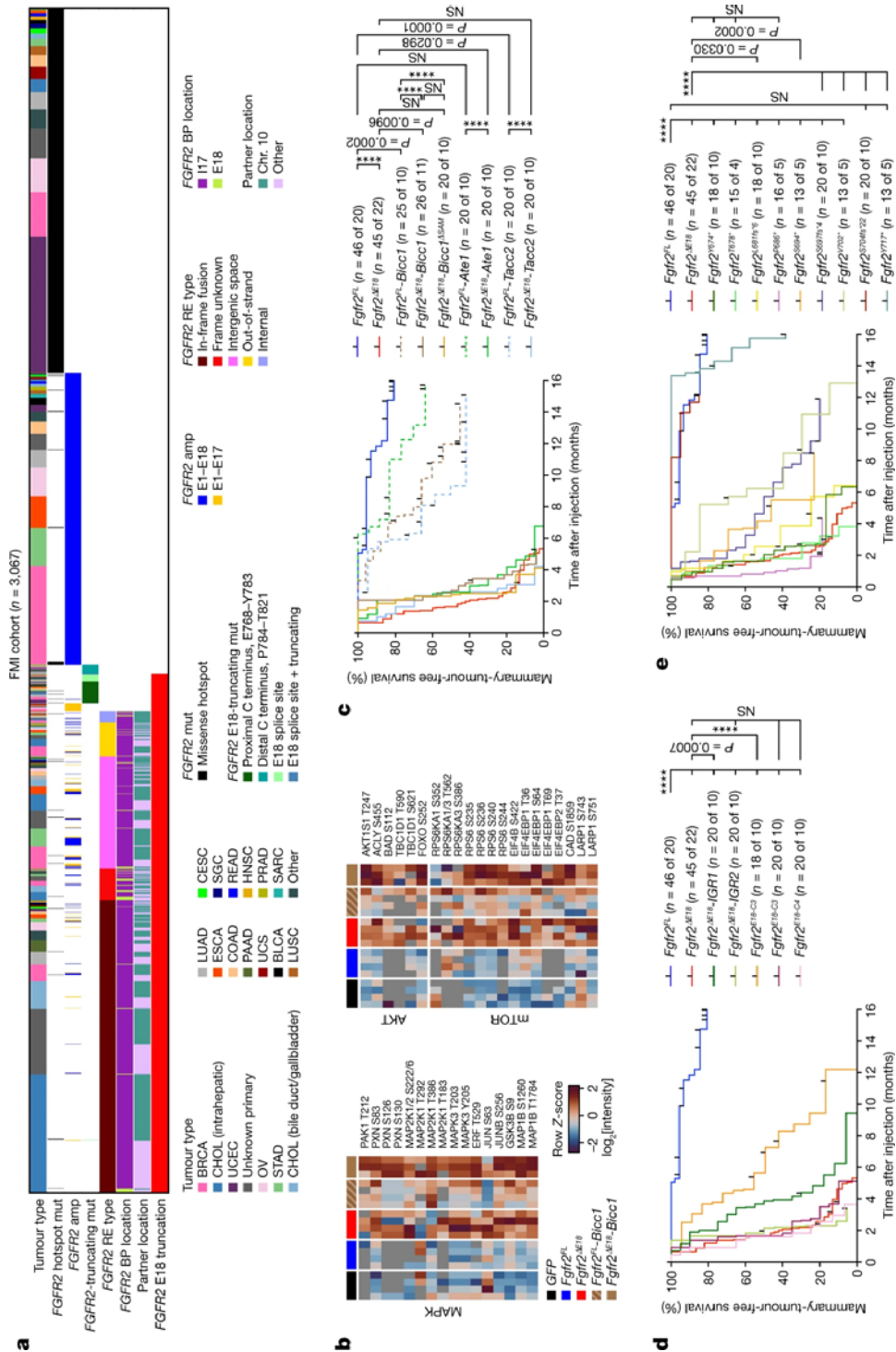


Fig. 2: Human *FGFR2* E18-truncating alterations are oncogenic drivers in mice.

a. Analysis of 3,067 samples (1.23% incidence) containing *FGFR2*-I17/E18 in-frame fusions ($n = 757$, 0.30% incidence), frame unknown REs ($n = 82$, 0.03% incidence), intergenic space REs ($n = 291$, 0.12% incidence), out-of-strand REs ($n = 88$, 0.04% incidence), internal REs ($n = 29$, 0.01% incidence), *FGFR2*-E18 splice-site mutations (mut; $n = 21$, 0.01% incidence), E18-truncating nonsense and frameshift mutations (proximal, $n = 59$, 0.02% incidence; distal, $n = 23$, 0.01% incidence), *FGFR2*-E1-E17 partial amplifications (amp; $n = 73$, 0.03% incidence), E1-E18 full-length amplifications ($n = 838$, 0.34% incidence), and/or *FGFR2* missense hotspot mutations affecting Ser252, Cys382, Asn549 or Lys659 ($n = 978$, 0.39% incidence) found in 249,570 pan-cancer diagnostic panel-seq profiles from FMI. BLCA, bladder urothelial carcinoma; BRCA, breast invasive carcinoma; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; CHOL, cholangiocarcinoma; chr, chromosome; COAD, colon adenocarcinoma; ESCA, oesophageal carcinoma; HNSC, head and neck squamous cell carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; chr, chromosome; COAD, colon adenocarcinoma; UCEC, uterine corpus endometrial carcinoma; PAAD, pancreatic adenocarcinoma; READ, rectum adenocarcinoma; SARC, sarcoma; SGC, salivary gland carcinoma; STAD, stomach adenocarcinoma; UCS, uterus carcinosarcoma. **b.** Global phosphoproteomic analysis of NMuMG cells expressing *GFP* or the indicated *Fgfr2* variants. Groups were compared in a pairwise manner using the robust kinase activity inference (RoKAI) tool, including two-tailed hypothesis testing on Z-scores and false-discovery rate (FDR) multiple-testing correction using the Benjamini-Hochberg method. Group-comparison fold change (FC) values of $-1.5 \geq FC \geq 1.5$ and $P < 0.05$ were considered. The heatmaps show phosphosites subselected from the RoKAI output and grouped into the indicated signalling pathways guided by RoKAI as colour-coded row Z-scores calculated from $\log_2$ -transformed intensity values. **c-e.** Kaplan-Meier curves showing mammary-tumour-free survival of female mice intraductally injected with lentiviruses encoding the indicated *Fgfr2* variants. Cohort counts (n) are injected mammary glands (MGs) per number of mice. The *Fgfr2^{FL}* and *Fgfr2^{E18}* curves in **c** are duplicated in **d** and **e**. P values were calculated using log-rank (Mantel-Cox) tests; **** $p < 0.0001$; NS, not significant ($P \geq 0.05$).

revealed complex *FGFR2* REs with several BPs that ultimately yielded in-frame fusions with protein-coding genes (Extended Data Fig. 3d,e).

Next, we performed hybrid-capture RNA-seq analysis of two tumour samples, which were diagnosed by FMI to contain structural variants of unknown significance in *FGFR2*. One sample contained an *FGFR2*-I17 RE with intergenic space and the other contained an *FGFR2*^{amp} involving E1–E17 only. RNA-seq profiling revealed a *FGFR2* in-frame fusion in the first tumour, whereas the second tumour showed high *FGFR2*-E1–E17 expression with splicing to E18-C3 (Extended Data Fig. 4a,b). Comprehensive analysis of The Cancer Genome Atlas (TCGA) RNA-seq data identified tumours expressing *FGFR2* in-frame fusions as well as non-canonical REs (Extended Data Fig. 4c–e). We found a few tumours containing *FGFR2*-I17 REs and concomitantly using E18-C3. However, a larger fraction of tumours used *FGFR2*-E18-C3 or *FGFR2*-E18-C4 in a mutually exclusive manner (Extended Data Fig. 4d–g and Supplementary Table 3). Taken together, we demonstrated that human tumours express diverse *FGFR2*^{ΔE18} transcripts derived from a variety of genomic alterations and alternative splicing events.

E18 loss is key to *FGFR2* oncogenicity

Previous research showed in vitro transforming abilities of C-terminally truncated *FGFR2* isoforms [17–19,25–27]. Our in vivo screening data and analyses of human oncogenomic datasets similarly suggested that exclusion of E18 is a critical determinant to render *FGFR2* REs oncogenic. To test this, we introduced mouse *Fgfr2*^{ΔE18} variants into mouse mammary epithelial cells. These were *Fgfr2*^{ΔE18} alone or fused to *Ate1*, *Bicc1*, *Tacc2*, two of the IGRs found in TCGA (Extended Data Fig. 4f,g), or the human E18-C2, E18-C3 or E18-C4 sequences, as well as *Fgfr2* bearing E18 nonsense and frameshift mutations. The corresponding controls were full-length (FL) *Fgfr2* (representing *FGFR2*^{amp}), *Fgfr2*^{FL} fusions, *Fgfr2*^{hotspot} variants and kinase-domain-dead *Fgfr2*^{K422R} variants (Extended Data Figs. 2k and 5a and Supplementary Table 4). Mass-spectrometry-based expression proteomics and phosphoproteomics revealed that overexpressed *Fgfr2*^{ΔE18} and *Fgfr2*^{ΔE18}-*Bicc1* both induced *FGFR2* signalling resulting in the activation of the MAPK and PI3K–AKT–mTOR pathways (Fig. 2b and Extended Data Fig. 5b–f). This depended on a functional *FGFR2* kinase domain, whereas the BICC1–SAM oligomerization domain [28] was dispensable for *Fgfr2*^{ΔE18}-*Bicc1* activity (Extended Data Fig. 5g,h). Comparably, all of the tested *Fgfr2*^{ΔE18} variants, including proximal E18-truncating mutations and hotspot *Fgfr2*^{C287R}, promoted colony formation in a 3D soft agar assay (Extended Data Fig. 6a). By contrast, overexpression of *Fgfr2*^{FL}, its fusion variants that retain E18, and distal E18-truncating mutations and the remaining *Fgfr2*^{hotspot} variants had limited potential to promote *FGFR2* signalling or soft agar colonies (Fig. 2b and Extended Data Figs. 5 and 6a).

Next, we evaluated the in vivo oncogenicity of *Fgfr2* variants using somatic delivery to mouse mammary glands through intraductal injection of lentiviruses [12,13]. Lineage

tracing using lentiviral *Fgfr2*-P2A-*cre* constructs and *mT/mG* female mice showed comparable mammary epithelial transduction rates and FGFR2 expression levels across the *Fgfr2* variants tested. However, only *Fgfr2*^{ΔE18} variants drove clonal expansion of the mammary epithelium, which depended on the FGFR2 kinase domain but not on the BICC1–SAM oligomerization domain (Extended Data Fig. 6b,c). To assess *Fgfr2*^{ΔE18} oncogenicity in mammary tumour models representative of different breast cancer subtypes—including invasive lobular carcinoma, a hallmark of which is E-cadherin loss [29]—we intraductally delivered *Fgfr2* variants to wild-type (WT) or *Wap-cre;Cdh1*^{F/F} mice. *Fgfr2*^{ΔE18} variants rapidly induced mammary tumours regardless of *Cdh1* mutation status (Fig. 2c,d and Extended Data Fig. 6d–g), and progressive truncation of *Fgfr2*-E18 gradually decreased tumour onset (Fig. 2e and Extended Data Figs. 6h and 7a,b). By contrast, mammary glands injected with *Fgfr2*^{FL} variants displayed no or slow tumorigenesis in WT and *Wap-cre;Cdh1*^{F/F} mice (Fig. 2c–e and Extended Data Figs. 6d–h and 7a,b). *Fgfr2*^{hotspot} variants were also non-tumorigenic, except for *Fgfr2*^{C287R}, which drove marked mammary tumour formation (Extended Data Fig. 7c,d). Furthermore, we generated genetically engineered mouse models (GEMMs) bearing Cre-inducible *Fgfr2*-IRES-*Luc* alleles (Extended Data Fig. 7e–i), in which *Wap-cre*-mediated induction of *Fgfr2*^{FL}-IRES-*Luc* had comparably little effect on mammary tumorigenesis. However, induction of *Fgfr2*^{ΔE18}-IRES-*Luc* led to increased mammary gland bioluminescence, which coincided with rapid and multifocal tumour formation in *Wap-cre;Cdh1*^{F/F};*Fgfr2*^{ΔE18}-IRES-*Luc* and *Wap-cre;Cdh1*^{F/F};*Fgfr2*^{ΔE18}-IRES-*Luc* females (Extended Data Fig. 7j–m). Histopathological evaluation of the mammary glands of *Fgfr2*^{FL} somatic models and GEMMs revealed mostly healthy tissue or low-grade lesions. By contrast, the majority of *Fgfr2*^{ΔE18} glands contained FGFR2-positive high-grade adenocarcinomas or E-cadherin-negative invasive lobular carcinomas or sarcomatoid tumours (Extended Data Fig. 8). Proteomic analyses of tumours induced by *Fgfr2* variants demonstrated consistent expression and phosphorylation of FGFR2 variants along with downstream signalling activities, which were distinct from the phosphoproteome of FGFR2-independent *K14-cre;Brca1*^{F/F};*Trp53*^{F/F};*(Mdr1a/b)*^{−/−} tumours (Extended Data Fig. 9a–c). Notably, MAPK and AKT–mTOR signalling pathways were particularly active in tumours driven by *Fgfr2*^{ΔE18} variants (Extended Data Fig. 9d,e). Together, these data establish that E18 truncation of *Fgfr2* is a bona fide tumour-driver alteration and the loss of the C terminus is a key determinant of FGFR2 oncogenicity.

FGFR2 oncogenicity depends on co-drivers

Compared with *Fgfr2*^{ΔE18}, our *in vivo* modelling efforts showed limited oncogenic competences of *Fgfr2*^{FL} and *Fgfr2*^{hotspot} variants in WT and *Cdh1*-deficient mammary glands. Yet, besides *FGFR2*^{ΔE18}, *FGFR2*^{amp} and *FGFR2*^{hotspot} made up considerable fractions of human *FGFR2* alterations. The oncogenic ability of specific *FGFR2* alterations might be affected by the tissue of origin as well as the mutational context. To examine possible cooperation between *FGFR2* variants and other genes, we analysed driver gene alterations diagnosed by FMI oncogenomic profiling and their incidence in *FGFR2*-altered cancers (Extended

Data Fig. 10a and Supplementary Table 2). *FGFR2^{ΔE18}*, *FGFR2^{amp}* and *FGFR2^{hotspot}* showed co-occurrences and mutual exclusivities with distinct sets of driver alterations (Extended Data Fig. 10b,c). The proportions of *FGFR2^{ΔE18}*, *FGFR2^{amp}* and *FGFR2^{hotspot}* varied across cancer entities, suggesting differential selections of *FGFR2* aberrations and concurrent driver alterations among tissues of origin (Extended Data Fig. 10d). We evaluated driver-gene enrichments among the three *FGFR2* alteration categories in a tumour-type-specific manner. In breast cancers with *FGFR2^{amp}*, *TP53* driver mutations, *MYC* amplifications, *PTEN* loss-of-function alterations, and *CCND1* and *FGF3/4/19* co-amplifications were significantly more enriched compared with the other classes of *FGFR2* aberration (Fig. 3a,b). Accordingly, *FGFR2^{amp}* showed co-occurrence with *TP53*, *PTEN* and *MYC* alterations in breast cancer (Fig. 3c). In several other cancer types, we also observed enrichments of *TP53* and *MYC* driver alterations in *FGFR2*-amplified cases (Extended Data Fig. 10e). By contrast, *FGFR2^{ΔE18}* and *FGFR2^{hotspot}* samples did not co-occur with these drivers (Fig. 3b,c and Extended Data Fig. 10e). This suggested that the oncogenic competence of full-length *FGFR2^{amp}* depends on specific cooperating driver genes.

We therefore combined lentiviral *Fgfr2^{FL}* with *cre* to delete floxed *Trp53* (*Trp53^f*) alleles and a single-guide RNA against *Pten* (sgPten) to disrupt the endogenous *Pten* locus. Intraductal delivery of *Fgfr2^{FL}-P2A-cre* or *sgPten-Fgfr2^{FL}-P2A-cre* lentiviruses into mammary glands of *Trp53^{f/f}* or *Trp53^{f/f};Cas9* mice, respectively, significantly increased *Fgfr2^{FL}* tumorigenicity. *Fgfr2^{FL}* became nearly as oncogenic as *Fgfr2^{ΔE18}* when *Trp53* and *Pten* were concomitantly lost, whereas *Fgfr2^{ΔE18}* oncogenicity was unaffected by the loss of *Trp53* and/or *Pten* (Fig. 3d and Extended Data Fig. 11a). Similarly, combinations of *Fgfr2^{FL}* with *Myc*, *Fgf3* and/or *Ccnd1* cDNAs into single lentiviral constructs cooperatively shortened tumour onset after intraductal delivery, with the latencies of the *Fgfr2^{FL}-T2A-Myc* and *Fgfr2^{FL}-T2A-Fgf3-P2A-Ccnd1* combinations matching *Fgfr2^{ΔE18}* single-driver latency. Notably, *Myc*, *Fgf3* and *Ccnd1* alone were effectively non-tumorigenic (Fig. 3e,f and Extended Data Fig. 11b,c). Evaluation of mammary glands containing *Fgfr2^{FL}* and co-driver alterations confirmed targeting or expression of the driver combinations and revealed high-grade tumours comparable to *Fgfr2^{ΔE18}*-driven lesions (Extended Data Fig. 11d,e). Thus, *Fgfr2^{FL}* oncogenicity relied on a cooperative oncogenomic network, whereas *Fgfr2^{ΔE18}* acted as a context-independent oncogene.

FGFR2^{ΔE18} tumours are sensitive to FGFRi

We next tested whether different *Fgfr2* variants were sensitive to the clinical FGFR inhibitors (FGFRi) AZD4547, pemigatinib, BGJ398 and debio-1347. Expression of *Fgfr2^{ΔE18}* variants and *Fgfr2^{C287R}* rendered mouse mammary epithelial cells highly sensitive to FGFR2 inhibition (Fig. 4a and Extended Data Fig. 12a,b). As a consequence, FGFRi suppressed both *Fgfr2^{ΔE18}*-variant-induced signalling and soft agar clonogenicity (Extended Data Figs. 5g,h and 6a). By contrast, cells expressing *Fgfr2^{FL}* and the remaining *Fgfr2^{hotspot}* variants were less sensitive to FGFRi (Fig. 4a and Extended Data Fig. 12a,b). We also orthotopically transplanted tumours

driven by *Fgfr2*^{ΔE18} variants and treated the recipient mice with AZD4547, which significantly suppressed tumour growth (Fig. 4b and Extended Data Fig. 12c,d).

To further investigate the connection between distinct *FGFR2* alterations and the FGFRi response in human tumour models, we analysed human cancer cell line pharmacogenomic datasets [30–32]. We evaluated the association between dose–responses to the FGFRi AZD4547 and PD173074 and genomic and transcriptomic features that potentially affect FGFR signalling, that is, *FGF3/4/19* amplification, *FGFR1–4* mutation, amplification, RE and expression, and use of *FGFR2*-E18-C3 (Extended Data Fig. 13a,b and Supplementary Table 5). Among these, *FGFR2/3* missense hotspot mutations, *FGFR2* expression and composite *FGFR* expression—a biomarker of FGFRi response [33]—modestly correlated with FGFRi sensitivity (Extended Data Fig. 13c–f), whereas cell lines with concurrent *FGFR2*^{amp} and E18 truncations through I17 REs and/or E18-C3 use exhibited high sensitivity to FGFRi (Extended Data Fig. 13a,g). Cell lines expressing E18-truncating *FGFR3* REs were less sensitive to FGFRi compared with cells with *FGFR2*^{ΔE18} (Extended Data Fig. 13b,g). On the basis of these correlates, we obtained the human SUM52PE, MFM-223, SNU-16, KATO-III and NCI-H716 cancer cell lines, each expressing amplified *FGFR2* variants [15–18,34,35]. These cells highly expressed full-length *FGFR2*^{E18-C1} but also E18-truncated transcripts, namely *FGFR2*^{E18-C3} and *FGFR2*-I17 REs, and were sensitive to FGFRi (Extended Data Fig. 14a–c). To functionally dissect the dependence on *FGFR2*^{E18-C1} versus E18-truncated transcripts, we used small interfering RNAs (siRNAs) targeting either shared or unique *FGFR2* exons (Extended Data Fig. 14d). Silencing of all *FGFR2* transcripts suppressed the growth of cell lines with *FGFR2*^{amp} (Extended Data Fig. 14e,f). Regardless of expression prevalence (Extended Data Fig. 14c), the growth of these cell lines could also be inhibited by specific silencing of E18-truncated *FGFR2* RE or E18-C3 transcripts. By contrast, siRNAs specifically targeting *FGFR2*^{E18-C1} only marginally suppressed cell line growth (Extended Data Fig. 14e,f). Importantly, in KATO-III cells mainly expressing *FGFR2*^{E18-C3} (Extended Data Fig. 14c), overexpression of *FGFR2*^{E18-C3} fully rescued silencing of any *FGFR2* transcripts, which depended on a functional FGFR2 kinase domain. However, full-length *FGFR2*^{E18-C1} was hardly able to rescue silencing of E18-truncated *FGFR2*^{E18-C3} (Extended Data Fig. 14g–i and Supplementary Table 4).

We next screened the CrownBio-HuPrime patient-derived xenograft (PDX) collection for the occurrence of genomic and transcriptomic alterations in the FGFR signalling pathway and enrolled the PDXs into a drug-intervention study using debio-1347 (Fig. 4c and Supplementary Table 6). All PDXs with *FGFR2/3*-I17 in-frame fusions as well as those with noncanonical *FGFR2* REs or E18-C3 use strongly responded to FGFR blockade (Fig. 4d and Extended Data Fig. 15a–c). Among the correlates with debio-1347 treatment response were *FGFR2* and composite *FGFR* expression (Extended Data Fig. 15d–f), but especially truncation of *FGFR2/3*-E18 exhibited substantial correlation with debio-1347-mediated growth inhibition (Extended Data Fig. 15g). Thus, human tumour models express

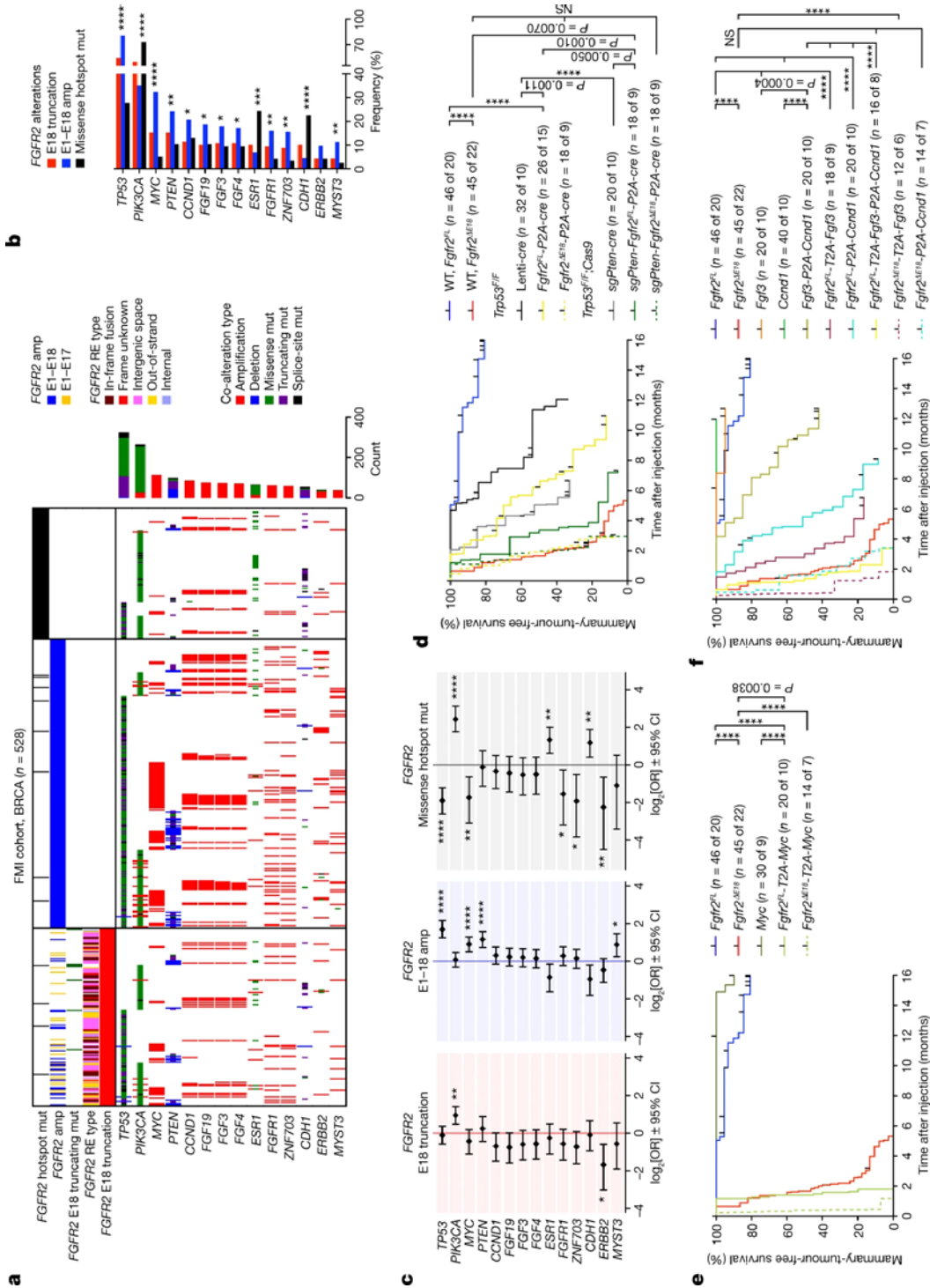


Fig. 3: The oncogenic competence of *FGFR2* alterations depends on co-occurring drivers.

a. Analysis of 528 breast cancer samples classified as either *FGFR2* E18-truncated ($n = 157$, 29.7% of total, 0.70% incidence), E1–E18 amplified ($n = 256$, 48.5% of total, 1.14% incidence) or missense hotspot mutant ($n = 115$, 21.8% of total, 0.51% incidence), and the top co-enriched tumour driver alterations found in 22,380 breast cancer profiles from FMI. **b.** Enrichment of the top tumour driver co-alterations in the indicated *FGFR2* alteration categories in the FMI breast cancer cohort. **c.** The odds ratios (ORs) of the top tumour driver co-alterations in the indicated *FGFR2* alteration categories (E18 truncation, $n = 157$; E1–E18 amplification, $n = 256$; missense hotspot mutation, $n = 115$) versus *FGFR2* WT samples ($n = 22,307$) of the FMI breast cancer cohort. Data are represented as log₂-transformed OR $\pm$ 95% confidence interval (CI). Co-occurrence, OR > 1; mutual exclusivity, OR < 1. *P* values were calculated using one-tailed proportion Z-tests (**b**) or two-tailed Fisher's exact tests (**c**) with FDR multiple-testing corrections using the Benjamini–Hochberg method (**b** and **c**). Sample sizes and statistical details for **b** and **c** are shown in Supplementary Table 2. **d–f.** Kaplan–Meier analysis of the mammary-tumour-free survival of *Trp53^{fl/fl}* and *Trp53^{fl/fl};Rosa26-Cas9* (**d**) or WT (**e,f**) female mice that were intraductally injected with lentiviruses encoding the indicated variants. Cohort counts (*n*) represent injected mammary glands (MGs) per number of mice. The *Fgfr2^{ΔE18}* curves in **d–f** are duplicates from Fig. 2c. *P* values were calculated using log-rank (Mantel–Cox) tests. **P* < 0.05; ***P* < 0.01; ****P* < 0.001; *****P* < 0.0001.



a, Half-maximum inhibitory concentration (IC_{50}) value quantifications of 2D-grown NMuMG cells expressing *GFP* or the indicated *Fgfr2* variants and treated with AZD4547 or pemigatinib for 4 days. Data are the mean of 5 independent experiments (*GFP*, *Fgfr2^{FL}*, *Fgfr2^{ΔE18}*) or 1 experiment (other *Fgfr2* variants). **b**, Kaplan–Meier analysis of mammary-tumour-specific survival of female syngeneic WT mice bearing mammary fat pad transplants derived from the indicated tumour donors and treated daily orally with vehicle or 12.5 mg per kg AZD4547 using an intermittent dosing regimen. *P* values were calculated using log-rank (Mantel–Cox) tests. **c**, Collection of PDX models ($n=36$) rank-ordered according to debio-1347 $\Delta T/\Delta C$ response ratios. *FGF/FGFR* copy number alteration and mutation data and RNA-seq profiles to analyse *FGF/FGFR* expression (exp) were obtained from CrownBio–HuPrime, and had been generated from non-treated PDXs. Composite *FGFR* expression was defined as high if normalized expression >3 . *FGFR2*-E18-C3 use and *FGFR* RE types were identified in RNA-seq profiles. GBM, glioblastoma multiforme; KIRC, kidney renal clear cell carcinoma; LIHC, liver hepatocellular carcinoma. **d**, Growth curves of the indicated PDXs engrafted in female BALB/c nude mice and treated daily orally with vehicle or debio-1347 (BR1115 and LI1050, 60 mg per kg; ES0042, GA0080 and GA3055, 80 mg per kg). $n=3$ mice per PDX model and treatment group. Data are mean $\pm$ s.d. *P* values were calculated using one-tailed two-way analysis of variance with FDR multiple-testing corrections using the two-stage step-up method from Benjamini, Krieger and Yekutieli.

and are dependent on E18-truncated *FGFR2* and *FGFR3* variants and are actionable by FGFR-targeted therapies.

These findings suggest that patients with cancer with any type of *FGFR2*^{ΔE18} variant might respond to FGFR2 targeting. We therefore re-examined FIGHT-202 (NCT02924376), a phase II trial of pemigatinib in patients with advanced cholangiocarcinoma. Patients with fusions or REs in the *FGFR2*-I17/E18 hotspot had an objective response rate of 35.5%, whereas those with other or no *FGF/FGFR* alterations had no response [9]. To determine which classes of *FGFR2* REs benefit from pemigatinib, we stratified individual patients according to their *FGFR2*^{amp} status and E18-damaging RE types, namely in-frame fusions, frame unknown REs and REs with IGRs. Patients with *FGFR2* in-frame fusions, frame unknown REs and IGR REs—independently of *FGFR2*^{amp} status—showed strong tumour responses to pemigatinib therapy (Fig. 5a), resulting in objective responses or stable disease in 80–100% of patients, irrespective of the diagnosed *FGFR2* RE type (Fig. 5b). As a consequence, although the patient cohort with no *FGF/FGFR* alterations quickly progressed during pemigatinib treatment (Fig. 5a–c), the four patient cohorts with non-amplified and amplified *FGFR2* in-frame fusions and non-canonical REs showed equally prolonged progression-free survival times (Fig. 5c). Taken together, *FGFR2*^{ΔE18}, generated by either in-frame fusions or other REs, is a clinically actionable oncogene in patients with cholangiocarcinoma and probably in patients with other types of cancer.

DISCUSSION

The C-terminal tail of FGFR2 encoded by E18 proposedly moderates RTK signalling [14]. The proximal part of the C terminus includes the RTK internalization motif [27], a tyrosine residue that is relevant for PI3K-pathway attenuation [25,36], and serine residues that bind to ERK1/2 and RSK2 to stimulate receptor endocytosis and suppress MAPK signalling [37,38]. The distal C terminus contains proline-rich motifs that bind to GRB2 and mitigate kinase domain activity [39,40]. We found that especially proximal C-terminal truncation is critical for oncogenic signal transduction, as evidenced by the gradually accelerated tumorigenesis observed for the *Fgfr2*^{E18-C2} variant and proximal E18-truncating mutations. Notably, C-terminally truncated isoforms of EGFR, HER2 and other RTKs have also shown elevated transforming activities [41–44]. Thus, the paradigm of C-terminal FGFR2 truncation identified here might be key to the pathogenicity of multiple RTKs.

FGFR2 amplification and fusion structural variants have been considered to be relevant tumour drivers owing to the consequential receptor overexpression and constitutive dimerization mediated by oligomerization domains in the fusion partners [2,20–22]. We identified that the tumour-driver potential of C-terminally truncated FGFR2 is independent of specific fusion partners, whereas full-length FGFR2 overexpression was marginally

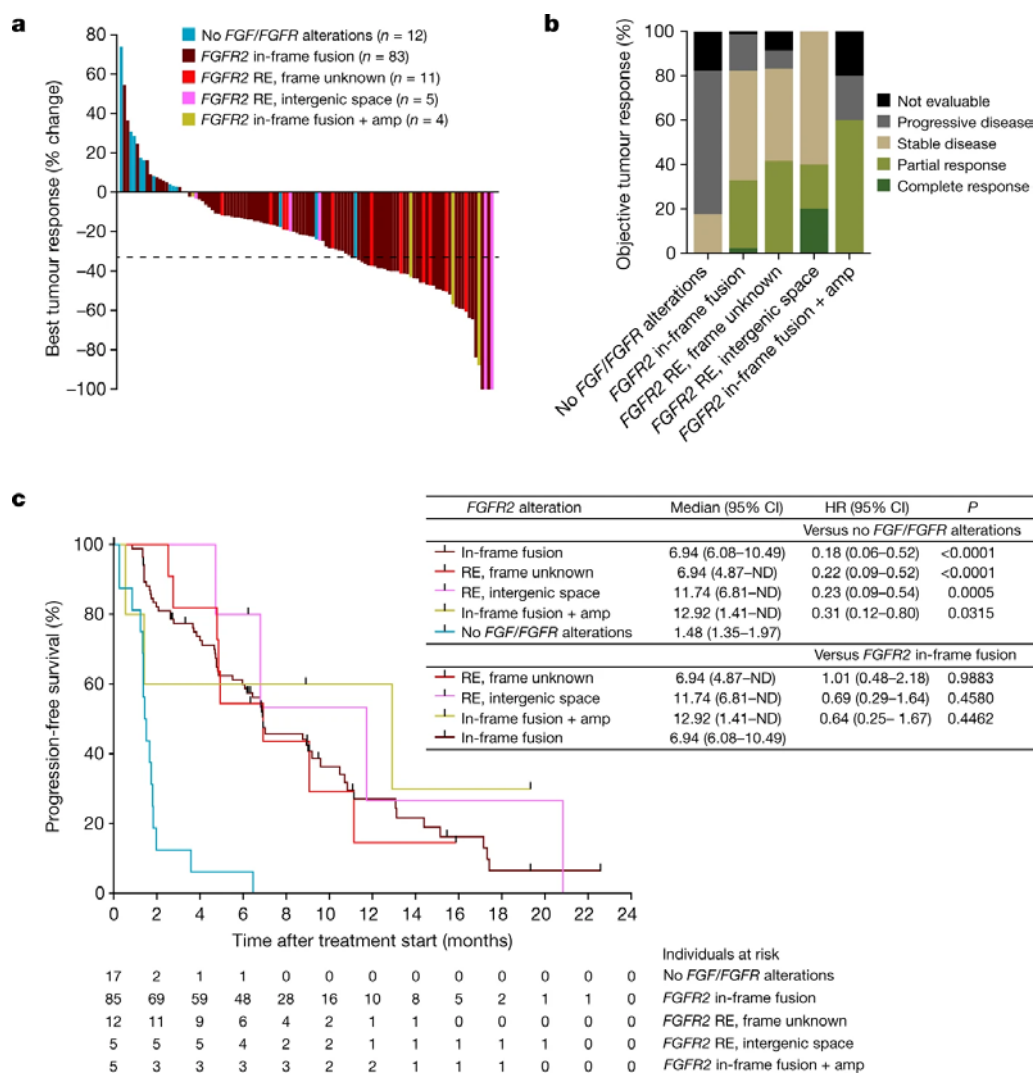


Fig. 5: Patients with cholangiocarcinoma respond to FGFR-targeted therapy irrespective of *FGFR2* RE type.

a, Centrally assessed best percentage change from the baseline in target lesion size of 115 (92%) of 125 individual patients with cholangiocarcinoma treated with pemigatinib, who had post-baseline scans. Data are from the FIGHT-202 study [9] and the coloured bars indicate *FGFR2*-117/E18 RE types and *FGFR2* amplification status as diagnosed by FoundationOne. **b**, Objective tumour responses observed in the FIGHT-202 study assessed according to the Response Evaluation Criteria in Solid Tumours v.1.1 (RECIST 1.1) and grouped according to *FGFR2* RE types/amplification status. No *FGF/FGFR* alterations ($n=17$), *FGFR2* in-frame fusion ($n=85$), frame unknown RE ($n=12$), intergenic space RE ($n=5$), in-frame fusion + amplification ($n=5$). 'Not evaluable' indicates that the patient was not evaluable for response using RECIST. **c**, Kaplan–Meier analysis of the progression-free survival of patients with cholangiocarcinoma treated with pemigatinib from the FIGHT-202 study and grouped according to *FGFR2* RE types/amplification status. Data are median $\pm$ 95% CI for each cohort, and log-rank hazard ratios (HR) $\pm$ 95% CI for the indicated comparisons are shown. P values were calculated using log-rank (Mantel–Cox) tests. ND, not defined.

tumorigenic in the absence of other driver alterations. The oncogenicity of *FGFR2^{amp}* might therefore depend on the ability to generate *FGFR2^{ΔE18}* transcripts, for example, through complex REs [16,17,24]. As shown in our study, *FGFR2^{ΔE18}* acts as a potent single-driver alteration. By contrast, the oncogenic competence of full-length *FGFR2* relied on co-drivers that may augment canonical FGFR2 signalling [45–48] and thereby phenocopy the strong signalling induction observed for C-terminally truncated FGFR2. In clinical trials, objective responses to FGFRi were scarce in patients with *FGFR2^{amp}* tumours [2–4]. Interestingly, tumours with overexpression of E18-truncated *FGFR2^{E18-C3}* responded particularly well to FGFR2 targeting [2]. In cohorts of mixed *FGFR* alterations, patients with *FGFR2*-E18-truncating fusions displayed favourable responses over patients with other *FGF/FGFR* alterations [5–9]. Thus, FGFRi efficacy might be dictated by the expression of single-driver *FGFR2^{ΔE18}* versus *FGFR2* alterations that depend on oncogenic co-drivers. In *FGFR*-aberrant cancers, *MYC* or *CCND1* amplifications can indeed confer resistance to FGFRi [48–50]. Combination therapies might therefore elevate the response rates in *FGFR2^{amp}* tumours, as proposed for FGFRi–CDK4/6i combination therapy in patients with breast cancer with *FGFR2* and *CCND1* amplifications [50].

Our findings have fundamental implications for the selection of patients for FGFR2 targeting therapies. Instead of considering patients on the basis of *FGFR2* mutation, fusion or amplification status alone, our data suggest that expression of oncogenic *FGFR2* transcripts and co-mutational landscapes should also be considered. Importantly, identifying cancers with structural variants or mutations that result in expression of *FGFR2^{ΔE18}* variants will be a highly relevant biomarker for FGFR-targeted therapeutics, and may substantially expand the number of cancer patients who may benefit from such therapy.

MATERIALS AND METHODS

Mouse models

GEMMs. The FVB/NRj *Wap-cre;Cdh1^{F/F}, Trp53^{F/F}, Trp53^{F/F};Rosa26-Cas9*, and *Rosa26-mT/mG* mouse strains were maintained at the NKI animal facility and PCR-genotyped as previously described [12,13,51–53]. To generate GEMMs bearing *Fgfr2*-IRES-Luc alleles, mouse *Fgfr2* (NM_201601.2) was isolated from a cDNA clone (MC221076, OriGene) using the primer sequences listed in Supplementary Table 7 amplifying *Fgfr2*-E1–E18 (FL) or *Fgfr2*-E1–E17 (ΔE18) and the sequences were verified and inserted with FseI-PmeI fragments into the *Frt-invCag*-IRES-Luc vector (shuttle vector). This resulted in the *Frt-invCAG-Fgfr2^{FL}*-IRES-Luc and *Frt-invCAG-Fgfr2^{ΔE18}*-IRES-Luc alleles. Flp-mediated integration of the shuttle vectors in *Wap-cre;Cdh1^{F/F};Col1a1^{frt/+}* GEMM-derived embryonic stem cell (ESC) clones (FVB/NRj background) and subsequent blastocyst injections of the modified ESCs were performed using the GEMM–ESC methodology [54]. Chimeric animals were mated with *Cdh1^{F/+}* and *Cdh1^{F/F}* mice on the FVB/NRj background to generate the experimental cohorts.

The *Col1a1*^{frt-invCAG-Fgfr2-IRES-Luc/+} and WT alleles were detected using standard PCR with an annealing temperature of 58°C using the primer sequences listed in Supplementary Table 8 generating the following PCR products: *Col1a1*^{frt-invCAG-Fgfr2-FL-IRES-Luc}, 585 bp; *Col1a1*^{frt-invCAG-Fgfr2-ΔE18-IRES-Luc}, 420 bp; and WT, 234 bp. Here, *Col1a1*^{frt-invCAG-Fgfr2-FL-IRES-Luc} and *Col1a1*^{frt-invCAG-Fgfr2-ΔE18-IRES-Luc} are referred to as *Fgfr2*^{FL}-IRES-Luc and *Fgfr2*^{ΔE18}-IRES-Luc, respectively. The GEMM cohorts were monitored weekly and mammary-tumour-free survival was scored (event) when the first palpable tumour was detected, whereas mice that did not develop any mammary tumours were censored. Tumour volume was measured in two dimensions using callipers as follows: volume = length × width² × 0.5.

Somatic mouse models. To somatically model *Fgfr2* variants in the mouse mammary gland, 6-week-old FVB/NRj WT, *Wap-cre;Cdh1*^{F/F}, *Trp53*^{F/F}, *Trp53*^{F/F};*Rosa26-Cas9* or *Rosa26-mT/mG* female mice were intraductally injected as previously described [12] with lentiviruses encoding *Fgfr2* variants in combination with *cre*, *Myc*, *Ccnd1*, *Fgf3* and/or a previously validated sgRNA targeting E7 of *Pten* (sgPten; refs. [12,13]). In brief, 20 μl of high-titre lentiviruses were injected into the fourth and/or the third mammary glands using a 34G needle. Lentiviral titres ranging from 2 × 10⁸ to 2 × 10⁹ transfection units (TU) per ml were used. The somatic model cohorts were monitored twice weekly and mammary-tumour-free survival was scored (event) for each injected mammary gland individually when palpable tumours were detected, whereas mammary glands that did not develop any tumours were censored. Tumour volume was measured in two dimensions using callipers as follows: volume = length × width² × 0.5.

AZD4547 intervention study. To allograft tumours, DMSO-preserved 1 mm³ tumour fragments derived from somatic *Fgfr2* models were orthotopically transplanted into the right mammary fat pad of 8-week-old syngeneic FVB/NRj female mice (Janvier Labs) as previously described [55]. The mice were twice weekly weighed and monitored for mammary tumour development and, as soon as tumours reached a volume of 62.5 mm³ (5 × 5 mm, measured in two dimensions using callipers; volume = length × width² × 0.5), the mice were randomly allocated to vehicle versus AZD4547 FGFRi treatment arms. The treatments were performed daily through oral gavage using vehicle (1% Tween-80 in demineralized water) or 12.5 mg per kg AZD4547 (AstraZeneca) according to a previously optimized intermittent dosing regimen [55]. Mice were euthanized 1 h after the last dosing.

General guidelines. For all mouse models, mammary-tumour-specific survival was scored when a single mammary tumour burden reached a volume of 1,500 mm³, the total mammary tumour burden reached a volume of 2,000 mm³ or the mice suffered from clinical signs of distress, such as respiratory distress, ascites, distended abdomen, rapid weight loss and severe anaemia, caused by primary tumour burden or metastatic disease. Mice that were euthanized due to other circumstances were censored. The maximal permitted disease end points were not exceeded in any of the experiments. Mammary glands were

collected and analysed for histological abnormalities. Sample sizes were determined using G*Power software (v.3.1; ref. [56]) and were large enough to measure the effect sizes. Tumour measurements and post mortem analyses were performed in a blinded manner. The mouse colony was housed in a certified animal facility under a 12 h–12 h light–dark cycle in a temperature- and humidity-controlled room set to 21 °C and 55% relative humidity. The mice were kept in individually ventilated cages, and food and water were provided ad libitum. All of the animal experiments were approved by the Animal Ethics Committee of the Netherlands Cancer Institute and performed in accordance with institutional, national and European guidelines for Animal Care and Use.

In vivo bioluminescence imaging

In vivo bioluminescence imaging of luciferase expression was performed as previously described [57] by intraperitoneally injecting 150 mg per kg beetle luciferin (E1601, Promega). Signal intensity was measured on the whole body of the mouse (excluding the head and tail) using an IVIS Spectrum In Vivo Imaging System (124262, PerkinElmer) operated by Living Image Software (v.4.5.2, PerkinElmer) and a size-fixed square. Signal intensity was quantified as flux (photons per second per cm² per steradian).

Histology and IHC

Tissues were formalin-fixed and paraffin-embedded (FFPE), sectioned and processed for haematoxylin and eosin (H&E) histochemical and immunohistochemistry (IHC) staining using routine procedures. For IHC staining, antigen retrieval was performed with citrate buffer (CBB999, ScyTek) at pH 6 (FGF3, FGFR2, PTEN) or Tris-EDTA at pH 9 (MYC, Cyclin D1, E-cadherin, P53). Sections were incubated with primary antibodies (Supplementary Table 9) overnight at 4 °C. Primary antibodies were labelled with the EnVision+ HRP Labelled Polymer Anti-Rabbit System (K4003, Dako), visualized with the Liquid DAB+ Substrate Chromogen System (K3468, Dako) and counterstained with haematoxylin. The antibodies used were independently validated by a certified pathologist by evaluation of IHC results in positive and negative biological control FFPE tissues to ensure specificity and sensitivity. Moreover, negative technical controls were performed by omission of the primary antibody in extra sections for a randomly selected subset of the samples. H&E and E-cadherin slides were used to classify mammary tumour lesion types according to the international consensus of mammary pathology [58]. IHC stains were quantitatively analysed by evaluating tumour cell-specific positivity using a histo-scoring system (0, negative; 1, weakly positive; 2, moderately positive; 3, strongly positive) or by calculating a histo (*H*)-score for each tumour defined as follows: $H\text{-score} = 1 \times (\text{the percentage of tumour cells with weak staining intensity}) + 2 \times (\text{the percentage of tumour cells with moderate staining intensity}) + 3 \times (\text{the percentage of tumour cells with strong staining intensity})$, resulting in a score between 0 and 300. All slides were reviewed and quantified by a comparative pathologist (S.K.) in a blinded manner. Slides were digitally processed

using a Panoramic 1000 whole-slide scanner (3DHISTECH) and captured using CaseViewer software (v.2.2.1, 3DHISTECH).

Isolation of MMECs

Primary mouse mammary epithelial cells (MMECs) were isolated from 10-week-old WT, *Fgfr2^{FL}-IRES-Luc* and *Fgfr2^{ΔE18}-IRES-Luc* female mice as previously described [53]. In brief, mammary glands were minced and digested with 4 mg ml⁻¹ collagenase A (11088793001, Roche) and 25 μg ml⁻¹ DNase I (DN25, Sigma-Aldrich) in Dulbecco's modified Eagle medium/nutrient mixture F-12 (DMEM/F-12, 31331, Thermo Fisher Scientific) containing 100 IU ml⁻¹ penicillin–streptomycin (15070, Thermo Fisher Scientific) for 1 h at 37 °C. Digests were passed through a 70 μm cell strainer that was prewetted with PBS containing 10% fetal bovine serum (FBS, S-FBS-EU-015, Serana) and 2 mM EDTA, and cells were cultured in DMEM/F-12 supplemented with 10% FBS, 100 IU ml⁻¹ penicillin–streptomycin, 5 ng ml⁻¹ epidermal growth factor (EGF, E4127, Sigma-Aldrich), 5 ng ml⁻¹ insulin (I0516, Sigma-Aldrich) and 10 μM Y-27632 (M1817, AbMole). Contaminating fibroblasts were removed from MMEC cultures by differential trypsinization. Confluent wells were transduced with adenoviral Ad5CMVCre particles (AdCre, 1 × 10⁹ TU ml⁻¹; University of Iowa Viral Vector Core) in the presence of 8 μg ml⁻¹ Polybrene (H9268, Sigma-Aldrich) for 48 h before subjection to subsequent assays. Switching of *Fgfr2-IRES-Luc* alleles was confirmed using 1 mg ml⁻¹ beetle luciferin and bioluminescence imaging with an Infinite M Plex plate reader (Tecan) operated with Tecan i-control software (v.3.9.1, Tecan).

FACS analysis of mammary glands

Mammary glands of *Rosa26-mT/mG* female mice injected with *Fgfr2-P2A-cre* lentiviruses were processed as described for MMEC isolation. Single cells were stained with the fluorescence-activated cell sorting (FACS)-validated BV650-conjugated anti-EPCAM antibody (1:100, 740559, BD Biosciences) in FACS buffer (PBS with 10% FBS and 2 mM EDTA), labelled with the LIVE/DEAD Fixable Violet Dead Cell Stain Kit (405 nm excitation, L34964, Thermo Fisher Scientific), fixed with BD Phosflow Fix Buffer I (557870, BD Biosciences) and permeabilized with BD Phosflow Perm Buffer III (558050, BD Biosciences), each for 30 min at 4 °C. Cells were incubated with primary antibodies (anti-FGFR2, 1:200, 11835, Cell Signaling Technology; anti-GFP, 1:200, ab6673, Abcam) overnight and subsequently with secondary antibodies for 1 h both in FACS buffer at 4 °C. Anti-FGFR2 and anti-GFP antibodies were validated for FACS using NMuMG cells overexpressing GFP or FGFR2 versus control cells negative for these proteins. Details of the antibodies used are provided in Supplementary Table 9. FACS was performed using the BD LSRFortessa Cell Analyzer (BD Biosciences) equipped with the BD FACSDiva Software (v.8.0.2, BD Biosciences) and with 405 nm (450/50, 670/30 pass filters), 488 nm (530/30 pass filters) and 638 nm (670/30 pass filters) lasers to measure 405-Live/Dead, BV650–EPCAM, EGFP–AF488 and FGFR2–AF647, respectively. Data were analysed using FlowJo (v.10.7.1, BD Biosciences).

Lentiviral vectors and virus production

The SIN.LV.SF, SIN.LV.SF-T2A-Puro, SIN.LV.SF-GFP-T2A-Puro lentiviral vectors and the SIN.LV.SF-Cre (Lenti-Cre), SIN.LV.SF-P2A-Cre and pGIN sgPten-P2A-Cre lentivectors all encoding improved *cre* with mammalian codon usage (derived from pBOB-CAG-iCRE-SD, 12336, Addgene) and the last-mentioned encoding a validated sgRNA targeting E7 of *Pten* (sgPten) were all previously described [10,12,13,59]. Mouse *Fgfr2* (NM_201601.2) and *Myc* (NM_010849.4) were isolated from cDNA clones (*Fgfr2*, MC221076, OriGene; *Myc*, 8861953, Source BioScience) using Q5 High-Fidelity DNA Polymerase (M0491S, New England Biolabs) and primers with AgeI–Sall overhangs amplifying *Fgfr2*^{FL} or *Fgfr2*^{ΔE18}, or BamHI–AgeI overhangs amplifying *Myc*. Amplicons were inserted into the SIN.LV.SF, SIN.LV.SF-T2A-Puro and SIN.LV.SF-P2A-Cre vectors, resulting in SIN.LV.SF-*Fgfr2*^{FL}, SIN.LV.SF-*Fgfr2*^{ΔE18}, SIN.LV.SF-*Fgfr2*^{FL}-T2A-Puro, SIN.LV.SF-*Fgfr2*^{ΔE18}-T2A-Puro, SIN.LV.SF-*Fgfr2*^{FL}-P2A-Cre, SIN.LV.SF-*Fgfr2*^{ΔE18}-P2A-Cre, and SIN.LV.SF-*Myc*. Custom-synthesized gBlocks gene fragments of mouse *Ccnd1* (NM_001379248.1), *Fgf3* (NM_008007.2), *Ate1*-E11–E12 (NM_001271343.1), *Bicc1*-E3–E21 (NM_001347189.1) and *Tacc2*-E15–E21 (NM_206856.4), which were the homologues of human *FGFR2* fusion partner genes (Extended Data Fig. 2h), as well as human full-length *FGFR2*^{E18-C1} (NM_000141.4), *FGFR2*^{E18-C2} (XM_017015921.2), *FGFR2*^{E18-C3} (NM_001144913.1), *FGFR2*^{E18-C4} (NM_001144915.2), the IGR1 sequence (identified in TCGA-A8-A08A; Extended Data Fig. 4f) and E18 sequences resulting from frameshift mutations (Extended Data Fig. 2k) were purchased (Integrated DNA Technologies). gBlocks gene fragments or parts thereof were assembled in the respective lentivectors using PCR amplification of backbone and insert(s) and the In-Fusion HD Cloning Plus kit (638911, Takara Bio) according to the manufacturer's recommendations. Point mutations in *FGFR2*, short deletions/insertions to generate gradual *Fgfr2* E18 truncations and introduction of the IGR2 sequence (identified in TCGA-BH-A203; Extended Data Fig. 4g) to *Fgfr2* were performed using the QuikChange Lightning Site-Directed Mutagenesis Kit (210519, Agilent). In-Fusion and site-directed mutagenesis primers were designed using SnapGene (v.5.2) and QuikChange Primer Design [60], respectively. All lentivectors were verified using Sanger sequencing. Concentrated lentiviral stocks were produced by transient co-transfection of four plasmids in HEK293T cells as previously described [61]. Viral titres were determined using the qPCR Lentivirus Titre Kit (LV900, Applied Biological Materials).

Cell culture

HEK293T cells (CRL-3216, ATCC) were cultured in Iscove's modified Dulbecco's medium (31980, Thermo Fisher Scientific) containing 10% FBS and 100 IU ml⁻¹ penicillin–streptomycin. MCF7 (HTB-22), MDA-MB-134-VI (HTB-23), MDA-MB-231 (HTB-26), NCI-H716 (CCL-251), NMuMG (CRL-1636), KATO-III (HTB-103), SNU-1 (CRL-5971) and SNU-16 (CRL-5974, all ATCC) as well as MFM-223 (98050130, ECACC) and SUM52PE (HUMANSUM-0003018, BioIVT) cells were cultured in DMEM/F-12 containing 10% FBS and 100 IU ml⁻¹ penicillin–streptomycin. All cell lines were previously authenticated by providers. No re-authentication was carried out for this study. To stably express the lentiviral GFP-T2A-Puro or FGFR2-T2A-Puro constructs,

NMuMG and KATO-III cells were transduced with lentiviral supernatants at equal TU per ml in the presence of $8\mu\text{g ml}^{-1}$ Polybrene for 24 h. Transduced cells were selected with $2\mu\text{g ml}^{-1}$ puromycin (A11138, Thermo Fisher Scientific) for 5 days and subsequently grown in DMEM/F-12 containing 10% FBS, 100 IU ml^{-1} penicillin–streptomycin, $1\mu\text{g ml}^{-1}$ puromycin and $10\mu\text{M}$ Y-27632. Overexpression of lentiviral constructs was verified using RT–qPCR (Supplementary Table 4). All cell lines were cultured in standard incubators at 37°C with 5% CO_2 and routinely tested for mycoplasma contamination using the MycoAlert Mycoplasma Detection Kit (LT07-218, Lonza).

Gene silencing using siRNA

Human cells were transfected with Silencer Select Negative Control 1 or 2 siRNAs (siCo#1 and #2, 4390844, 4390847, Thermo Fisher Scientific) or Silencer Select siRNAs designed with the GeneAssist Custom siRNA Builder (Thermo Fisher Scientific) to target shared exons (E5, E9, E15) among *FGFR2* isoforms (siFGFR2^{E5/E9/E15}), the 3'-UTR of E18-C1 of full-length *FGFR2* (siFGFR2^{E18-C1}), the 3'-UTR of E18-C3 of truncated *FGFR2*^{E18-C3} (siFGFR2^{E18-C3}) or the *FGFR2*-*COL14A1* fusion (siFGFR2-COL14A1). siFGFR2^{E5} targeted endogenous *FGFR2* transcripts as well as *FGFR2* transcripts derived from lentiviral constructs. All other siRNAs specifically targeted endogenous *FGFR2* transcripts, because the *FGFR2* cDNA sequences used in the lentivectors lacked 3'-UTRs and contained silent mutations in E9 and E15 to prevent the binding of siFGFR2^{E9/E15}. A list of the custom-designed siRNA sequences is provided in Supplementary Table 10. siRNA (50 nM) was used in combination with the jetPRIME transfection reagent (114-15, Polyplus Transfection) as previously described [62].

Drug-response curves

A total of 800 NMuMG or SNU-1 cells; 2,000 MCF7, MDA-MB-231, KATO-III, SNU-16, or SUM52PE cells; 3,000 MFM-223 or NCI-H716 cells; or 4,000 MDA-MB-134-VI cells per well were seeded in 96-well plates using DMEM/F-12 supplemented with penicillin–streptomycin and 10% FBS for human cell lines or 3% FBS for NMuMG. After 24 h, cells were treated with FGFRi for 4 days using vehicle (DMSO), AZD4547 (AstraZeneca), or pemigatinib (HY-109099), BGJ398 (HY-13311) or debio-1347 (HY-19957, all MedChemExpress) with a range of 0.1 nM to $100\mu\text{M}$. Usage of AZD4547, pemigatinib, BGJ398 and debio-1347 was previously described [63–66]. Cell viability was assayed using CellTiter-Blue Reagent (G808A, Promega) for 4 h and subsequently measuring fluorescence on the Infinite M Plex plate reader operated using the Tecan i-control software. Drug-response curves were modelled using [inhibitor] versus response with variable slope (four parameters) and least-squares regression in Prism (v.9.3.1, GraphPad Software).

2D colony-formation assay

A total of 5,000 KATO-III or MCF7 cells per well were seeded in six-well plates and, after 24 h, were treated with vehicle, 100 nM AZD4547 or 100 nM pemigatinib, or transfected with 50 nM siRNAs and cultured for 7 days. For KATO-III cells, six-well plates were precoated

with laminin using RAC-11P cells [67] as previously described [53]. Cells were stained with crystal violet as previously described [53] and the plates were imaged using the GelCount colony counter (Oxford Optronix).

96-well cell growth assay

A total of 800 SNU-1 cells; 1,500 MCF7, MFM-223, KATO-III, SNU-16 or SUM52PE cells; or 3,000 NCI-H716 cells per well were seeded in 96-well plates and, after 24 h, were treated with vehicle or 100 nM AZD4547, pemigatinib, BGJ398 or debio-1347, or transfected with 50 nM siRNAs. Cell density was assayed over 8 days on sister plates using the CellTiter-Blue Reagent for 4 h and the Infinite M Plex plate reader operated using the Tecan i-control software.

3D soft agar colony-formation assay

Six- or twelve-well plates were precoated with 2 ml or 1 ml of 0.6% low-gelling-temperature agarose (A9414, Sigma-Aldrich) by diluting 3% agarose solution (in PBS) in DMEM/F-12 supplemented with 3% FBS and penicillin–streptomycin. The bottom layer was solidified at 4°C for 30 min. NMuMG cells were passed through a 70 µm cell strainer and 20,000 or 10,000 single cells per well of the six- or twelve-well plates, respectively, were suspended in 2 ml or 1 ml of 0.35% low-gelling-temperature agarose in DMEM/F-12 supplemented with 3% FBS, penicillin–streptomycin and vehicle, 100 nM AZD4547 or 100 nM pemigatinib, and plated on top. The top layer was solidified at 4°C for 30 min before transferring the plates to the incubator. The plates were imaged after 15 days using the GelCount colony counter and anchorage-independent growth was quantified using the integrated GelCount colony counting platform (v.1.1.2, Oxford Optronix).

FACS analysis of cells

Cultured NMuMG cells were collected with 2 mM EDTA and passed through a 70 µm cell strainer prewetted with FACS buffer. Single cells were labelled with the LIVE/DEAD Fixable Violet Dead Cell Stain Kit and fixed, permeabilized, incubated with primary and secondary antibodies, and analysed as described for FACS analysis of mammary glands.

RNA isolation and RT-qPCR

RNA from frozen mammary tumour pieces was isolated as previously described [10,68]. Cultured cells were lysed (72 h after siRNA transfection in case of human cell lines) in buffer RLY (BIO-52079, Bioline) containing 1% 2-mercaptoethanol. Total RNA extraction and DNase treatment of samples was performed using the ISOLATE II RNA Mini Kit (BIO-52072, Bioline) according to manufacturer's guidelines. Purified RNA was quantified using the DS-11 Series Spectrophotometer/Fluorometer (DeNovix) and subjected to reverse transcriptase reaction using the Tetro cDNA Synthesis Kit (BIO-65042, Bioline) with oligo (dT)₁₈ primers (tumour pieces) or random hexamer primers (cells). qPCR was performed using the SensiFAST SYBR Hi-ROX Kit (BIO-92005, Bioline) and the QuantStudio 6 Flex Real-Time PCR System (4485691, Thermo Fisher Scientific) operated with the QuantStudio Real-Time PCR Software (v.1.7.2, Thermo Fisher Scientific). Primers used were designed using Primer-BLAST [69] and a list

of which is provided in Supplementary Table 11. Relative quantified cDNA was normalized using either mouse *Hprt* (tumour pieces) or *Usf1* (cells) or human *USF1* as the housekeeping transcript.

Protein isolation and western blotting

NMuMG cells were cultured in DMEM/F-12 starvation medium (0% FBS) for 48 h and treated with vehicle or 100 nM AZD4547 for 3 h. Cells were lysed in previously described RIPA buffer [53] containing Halt protease and phosphatase inhibitor cocktail (78440, Thermo Fisher Scientific). Protein concentrations were determined using the BCA Protein Assay Kit (23227, Thermo Fisher Scientific) and by measuring absorbance using the Infinite M Plex plate reader operated with the Tecan i-control software. Equal amounts of protein and the BlueEye Prestained Protein Marker (PS-104, Jena Bioscience) were separated on NuPAGE 4–12% Bis-Tris Mini Protein Gels (NP0323, NP0329, Thermo Fisher Scientific) and transferred overnight at 4 °C onto nitrocellulose membranes (88018, Thermo Fisher Scientific) in previously described transfer buffer [53]. The membranes were stained with Ponceau S solution (ab270042, Abcam) and imaged using Fusion FX (Vilber), blocked in 5% bovine serum albumin (BSA, A8022, Sigma-Aldrich) in PBS-T (0.05% Tween-20) and incubated with primary antibodies in 5% BSA in PBS-T overnight at 4 °C. The membranes were washed with PBS-T and incubated with secondary antibodies in 5% BSA in PBS-T for 1 h at room temperature. A list of the antibodies used (all validated for western blotting by the manufacturers) is provided in Supplementary Table 9. The membranes were washed in PBS-T and developed using SuperSignal West Pico PLUS Chemiluminescent Substrate or Femto Maximum Sensitivity Substrate (34580, 34095, Thermo Fisher Scientific). The membranes were imaged using Fusion FX operated with the Fusion FX7 Edge imaging system (v.18.05, Vilber), post-imaging processed with Photoshop 2022 (v.23.2.2, Adobe) using input levels and output levels, and band-intensities were measured with mean grey value in Fiji (v.1.0; ref. [70]). Protein band intensities were normalized to β -actin and phosphoprotein bands were further normalized to corresponding total protein bands and to FGFR2 intensity.

Proteomics of mouse cells and tumours

Sample preparation. Two (global phosphoproteomics) or three (phosphorylated-Tyr immunoprecipitation (p-Tyr IP) proteomics) 15 cm dishes of NMuMG cells expressing GFP or *Fgfr2* variants were collected in 3 ml urea lysis buffer [71]. Fresh-frozen samples of *Fgfr2* tumours collected in this study and *K14-cre;Brca^{f/f};Trp53^{f/f}* (KB1P) and *KB1P;Mdr1a/b^{-/-}* (KB1PM) tumours collected elsewhere [72] were mounted with Milli-Q H₂O and processed using a cryotome. Sections were collected to a final wet weight of up to 250 mg in urea lysis buffer (40x wet weight). Lysates were sonicated and cleared by centrifugation as previously described [73]. Protein concentrations were determined using the BCA Protein Assay Kit, and protein phosphorylation integrity was verified using western blotting and the p-Tyr-1000 antibody (8954, Cell Signaling Technology). To create a spectral library

for protein expression analysis, for each setting, a ten-band in-gel-digestion experiment was performed and SDS gels were processed as described previously [74]. Per cell lysate sample, 45 µg total protein was loaded. Furthermore, 45 µg total protein of a mouse liver lysate [71] was added. Tumour lysates were prepared in 6 pools consisting of 4–7 individual samples each, and 60 µg total protein was loaded per pool. For global phosphoproteomics and p-Tyr IP experiments, in-solution protein digestion of an equivalent of 500 µg total protein (p-Tyr IP cells, 5 mg; p-Tyr IP tumours, 4 mg) using trypsin and desalting with Oasis HLB 1 cm³ Vac Cartridge (186000383, Waters) was performed as previously described [71,73]. For global phosphoproteomic experiments, phosphopeptide enrichment was performed on the AssayMAP Bravo Platform (Agilent Technologies) using 5 µl Fe(III)-NTA immobilized metal affinity chromatography (IMAC) cartridges (G5496-60085, Agilent Technologies) starting from 200 µg desalted peptides in 0.1% trifluoroacetic acid and 80% acetonitrile. Phosphopeptides were eluted in 25 µl 5% NH₄OH/30% acetonitrile. Phosphopeptide enrichment for KB1P(M) tumours was performed using titanium dioxide beads as previously described [71]. IP of p-Tyr-containing peptides was performed using the PTMScan p-Tyr-1000 Kit (8803, Cell Signaling Technology) as previously described [73].

MS measurements. For *Fgfr2* samples, phosphopeptides were separated using the Ultimate 3000 nanoLC–mass spectrometry (MS)/MS system (Thermo Fisher Scientific) equipped with a 50 cm × 75 µm ID Acclaim Pepmap (C18, 1.9 µm) column. After injection, peptides were trapped at 3 µl min⁻¹ on a 10 mm × 75 µm ID Acclaim Pepmap trap at 2% buffer B (80% acetonitrile, 0.1% formic acid) and separated at 300 nl min⁻¹ in a 10–40% gradient of buffer B over 110 min at 35 °C. Eluting peptides were ionized at a potential of +2 kV into a Q Exactive HF mass spectrometer (Thermo Fisher Scientific) operated by Tune (v.2.11) and Xcalibur Software (v.4.3.73.11, OPTON-30965, both Thermo Fisher Scientific). Intact masses were measured at *m/z* 350–1,400 at a resolution of 120,000 (at *m/z* 200) in the Orbitrap system using an AGC target value of 3 × 10⁶ charges and a maxIT of 100 ms. The top 15 peptide signals (charge-states 2⁺ and higher) were submitted to MS/MS in the higher-energy collision cell (1.4 amu isolation width, 26% normalized collision energy). MS/MS spectra were acquired at a resolution of 15,000 (at *m/z* 200) in the Orbitrap system using an AGC target value of 1 × 10⁶ charges, a maxIT of 64 ms and an underfill ratio of 0.1%, resulting in an intensity threshold of 1.3 × 10⁵. Peptide separation for KB1P(M) samples was performed using a 40 cm × 75 µm (inner diameter) fused silica column custom packed with 1.9 µm 120 Å ReproSil Pur C18 aqua (Dr. Maisch). After injection, peptides were trapped at 6 µl min⁻¹ on a 10 mm × 100 µm (inner diameter) trap column packed with 5 µm 120 Å ReproSil Pur C18 aqua at 2% buffer B and separated at 300 nl min⁻¹ in a gradient of 10–40% buffer B over 90 min. The LC column was maintained at 50 °C using a pencil column heater (Phoenix S&T). Eluting peptides were ionized at a potential of +2 kV into a Q Exactive HF mass spectrometer operated by Tune and Xcalibur Software. Intact masses were measured at a resolution of 70,000 (at *m/z* 200) in the Orbitrap system using an AGC target value of 3 × 10⁶ charges. The top 10 peptide signals (charge-states 2⁺ and higher) were submitted to

MS/MS in the higher-energy collision cell (1.6 amu isolation width, 25% normalized collision energy). MS/MS spectra were acquired at a resolution of 17,500 (at m/z 200) in the Orbitrap system using an AGC target value of 1×10^6 charges, a maxIT of 80 ms and an underfill ratio of 0.1%, resulting in an intensity threshold of 1.3×10^4 . For *Fgfr2* and KB1P(M) samples, a dynamic exclusion was applied with a repeat count of 1 and an exclusion time of 30 s. For protein expression experiments, peptides (1 μ g total peptides, desalted) were separated and eluted as described for *Fgfr2* phosphopeptides. The data independent acquisition (DIA)-MS method consisted of an MS1 scan from 350 to 1,400 m/z at a resolution of 120,000 (AGC target of 3×10^6 and 60 ms injection time). For MS2, 24 variable-size DIA segments were acquired at 30,000 resolution (AGC target 3×10^6 and auto for injection time). The DIA-MS method starting at 350 m/z included one window of 35 m/z , 20 windows of 25 m/z , 2 windows of 60 m/z and one window of 418 m/z , which ended at 1,400 m/z . Normalized collision energy was set at 28. The spectra were recorded in centroid mode with a default charge state for MS2 set to 3⁺ and a first mass of 200 m/z . Spectral library data files were acquired with the same acquisition settings as for the phosphoproteomic experiments.

(Phospho)-peptide quantification and data analysis. For protein expression experiments, MS/MS spectra derived from data-dependent acquisition (DDA) mode of the in-gel digestion experiment were searched against the Swiss-Prot *Mus musculus* reference proteome (25,374 entries, canonical and isoforms, release 2021_10) using MaxQuant (v.2.0.3.0; refs. [75,76]) software with the default settings. Peptide identifications were propagated across samples with the match between runs (MBR) option enabled. The MaxQuant msms.txt file was used to generate a spectral library using Spectronaut software (v.15.4.210913, Biognosys). Spectra derived from single sample measurements in DIA mode were first analysed library-free in Spectronaut (directDIA) using the Biognosys factory settings to create a second spectral library. For the final search of DIA data in Spectronaut, both libraries were assigned using the default settings using the protein LFQ method set to MaxLFQ, imputation option switched off and an automatic normalization strategy. The Spectronaut report was further processed with R. Single-sample gene set enrichment analysis (ssGSEA) was performed by the GenePattern platform [77] using the ssGSEA module (v.10.0.11; ref. [78]) and Hallmark gene sets from MSigDB (v.7.0; ref. [79]). Missing values were imputed with a zero. For phosphoproteome experiments, phosphopeptide identification and quantification was performed as previously described [71] using the Swiss-Prot (*Fgfr2* samples) or the UniProt (KB1P(M) samples) *Mus musculus* reference proteomes (UniProt, 34,331 entries, canonical and isoforms, release 2015_06) and MaxQuant software. Phosphosites with a localization probability of <0.75 (class 1; ref. [80]) were discarded. The R package limma (v.3.52.1; ref. [81]) was used to perform differential expression analysis on class 1 phosphosite intensity data. For two-group comparisons, phosphosite intensity data were filtered for high data presence in at least one of the groups under comparison (cells, $\geq 75\%$; tumours, $\geq 50\%$). In the case of data presence in one group and absence in the other (phosphosite on/off behaviour), only observations with a very high data presence in the

'phosphosite on' group were allowed (cells, 100%; tumours, $\geq 90\%$). In these cases, missing values were imputed in the 'phosphosite off' group with a zero. Fold change values were determined using the mean of each treatment group and the antilog value was calculated. If downstream analysis did not allow the presence of duplicated phosphosite amino acid windows, the entry with the lowest *P* value was used. Phosphosite signature enrichment analysis (PTM-SEA; ref. [82]) was performed with the GenePattern platform [77] using a seven-amino-acid sequence flanking the phosphosite as an identifier and the mouse kinase/pathway definitions of PTMsigDB (v.1.9.0; ref. [82]) with the default settings. When PTM-SEA was performed following a two-group comparison, the rank metric was derived by multiplying the sign of FCs with the $-\log_{10}$ -transformed *P* values calculated by limma. When PTM-SEA was performed on single samples, duplicated phosphosite amino acid windows were filtered for entries with the highest row-sum of intensities over all of the samples. The samples were ranked using the phosphosite intensities and missing values were imputed with a zero. To assign probable upstream kinases to differentially regulated phosphosites, the robust kinase activity inference (RoKAI) tool [83] was used with the default settings and the UniProt *Mus musculus* reference proteome. RoKAI kinase and kinase target tables were shortlisted (cells, $\text{FDR} < 0.05$, number of substrates ≥ 3 ; tumours, number of substrates ≥ 2), assigned to significantly changed phosphosites ($-1.5 \geq \text{FC} \geq 1.5$, $P < 0.05$) and selected subsets of these phosphosites were visualized.

Analysis of *SB* transposon insertions in the mutagenesis screen

For the *SB* transposon insertional mutagenesis screen in ref. [10], mapping of *SB* insertions and calculation of insertion clonalities using next-generation sequencing of genomic DNA from *SB*-containing tumours was described in detail [10]. In brief, the relative clonality scores of *SB* insertions were calculated by normalizing each unique ligation score between genomic DNA and a *SB* cassette insertion to the highest ligation score within a given tumour sample. Then, each *SB* insertion was assigned a score between 0 (no insertion) and 1 (fully clonal insertion). Tumours with at least one relative insertion clonality score for *Fgfr2* of ≥ 0.25 were defined as tumours containing *SB* insertion(s) in *Fgfr2* ($n = 65$ tumours; total, $n = 123$ tumours).

Analysis of RNA-seq data from *SB* tumours

Published RNA-seq data generated from tumours of the *SB* transposon insertional mutagenesis screen [10] were used to derive *Fgfr2* gene- and exon-level expression as well as splice junction information. Gene fusions affecting *Fgfr2* in tumours with *SB* insertions were previously identified [11]. To quantify the expression of *SB* transposons in *Fgfr2*, customized fasta and gtf files were constructed for individual tumours by inserting the *SB* transposon sequence at the genomic position and according to its orientation as previously mapped [10]. Sequencing reads were then mapped on the basis of the customized fasta and gtf files using STAR (v.2.7.2; ref. [84]). Splice junctions between *Fgfr2*-E17 and the *SB* transposon were quantified using SJ.out.tab obtained from STAR alignment. To determine the usage

of *Fgfr2*-I17-inserted *SB* transposons as splice acceptors, the ratio of junction reads spanning from E17 to the *SB* transposon versus E18 was computed. Integrated Genomics Viewer (IGV, v.1.11.0; ref. [85]) was used to generate sashimi plots.

Analysis of WGS data from the HMF

WGS data on metastatic solid tumours were obtained from the HMF (data access request DR-138) through their Google cloud computing platform and analysed based on their bioinformatics pipeline (<https://github.com/hartwigmedical/pipeline5>) designed to detect all types of somatic alterations including structural variants and CNAs as previously described [23]. In brief, sequencing reads were mapped against the human reference genome GRCh37 using Burrows–Wheeler Alignment (BWA-MEM, v.0.7.5a; ref. [86]). Somatic structural variants were called with GRIDSS (v.1.8.0; ref. [87]) and CNAs and tumour purity were estimated using PURPLE (v.2.43; ref. [88]). Finally, LINX (v.1.9; ref. [88]) was performed to annotate events and to construct derivative chromosome structures. On the basis of the PURPLE output, samples containing structural variant BPs within the *FGFR2* genomic region were considered for further structural variant analyses ($n=266$ total BPs and 196 unique BPs in 86 tumour samples; Fig. 1f). To annotate structural variants, the location and orientation of both BP sides (*FGFR2* and its partner) were used to determine RE types. Among the RE partners, the gene encoding the longer protein sequence was used as RE classification backbone. The following RE types were defined: (1) in-frame fusion, both BP sides were located in the intronic regions of coding genes and the upstream and downstream exons adjacent to the BP were both in-frame (complete reading frame) or both BP sides were located in the exonic regions of coding genes and the fused sequence was in-frame; (2) frame unknown RE, both BP sides were located in the intronic regions of coding genes and either the upstream or the downstream exon adjacent to the BP was out of frame (incomplete reading frame), or one or both BP sides were located in exonic regions of coding genes and the fused sequence was out of frame. Any of these cases made the reading frame unpredictable (unknown). (3) RE with intergenic space, one BP side located to *FGFR2* and the other BP side located to a non-coding IGR; (4) out-of-strand RE, both BP sides were located in the coding regions of genes. The gene upstream to the BP (*FGFR2*) was supported by a sense-oriented read sequence, whereas the gene downstream to the BP was supported by an antisense-oriented read sequence; (5) internal RE, both BP sides were located within the genomic region of *FGFR2*; (6) unresolved REs, the gene upstream to the BP was supported by antisense-oriented read sequences or the REs contained single breakends. Unresolved REs were excluded, resulting in a refined list of samples containing *FGFR2* REs ($n=93$ REs in 55 tumour samples; Extended Data Fig. 1g,h). For the samples with multiple *FGFR2* REs, the relative allele frequency of each RE was computed using the ploidy level inferred by LINX. An I17/E18 RE allele frequency of >15% was used as a threshold to define samples with *FGFR2* REs causing E18 truncations (E18-truncating, $n=20$; others, $n=35$; Extended Data Fig. 1g,h and Supplementary Table 1). *FGFR2* copy number (CN) gains of >5 were defined as amplifications. Among the samples with *FGFR2* CN

segment BPs at I17, samples with E1–E17 CN (CN_{E1-E17}) > 5 and $CN_{E1-E17} - CN_{E18} > 2$ were defined as *FGFR2*-E1–E17 partially amplified. A few samples were expressing an *FGFR2* in-frame fusion gene based on RNA-seq, but showed discordant RE types in WGS. In these cases, in-depth annotation of the WGS data was performed using LINX to infer the plausible structures of derivative chromosomes constructed by complex RE events.

Analysis of RNA-seq data from the HMF

Raw RNA-seq data on metastatic solid tumours were obtained from the HMF (data access request DR-138). Sequencing reads were mapped to the human reference genome GRCh38 (Gencode v32 CTAT) using STAR (v.2.7.2; ref. [84]) with the recommended parameters to subsequently run STAR-Fusion (v.1.8.1; ref. [89]). STAR-Fusion was executed with chimeric alignment information (Chimeric.out.junction) obtained from STAR and GRCh38 Gencode v32 CTAT. Chimeric alignments from STAR and gene fusions from STAR-Fusion were inspected for RNA-seq alignments supporting the REs identified in WGS. For the samples with in-frame fusions identified in WGS, upstream and downstream exon numbers of the fusion gene inferred from STAR-Fusion were matched to the fusion found in WGS. For the samples with other types of REs identified in WGS, chimeric reads spanning the upstream exon and the downstream exon (out-of-frame REs), the downstream intergenic sequence (REs with intergenic space) or the downstream antisense gene sequence (out-of-strand REs) were mined from the 'Chimeric.out.junction' file and matched to the REs found in WGS. Genome coordinates were converted from GRCh37 to GRCh38 using UCSC Lift Genome Annotations (<https://genome.ucsc.edu/cgi-bin/hgLiftOver>). IGV (v.1.11.0; ref. [85]) was used to generate sashimi plots.

Analysis of RNA-seq data from TCGA

Among the 10,344 TCGA samples [90], we preselected samples potentially expressing *FGFR2*^{ΔE18} on the basis of several criteria: the presence of (1) *FGFR2* amplifications or (2) truncating mutations in *FGFR2*-E18, (3) shifts in CN segment values in *FGFR2*-I17, (4) a lack of *FGFR2*-E18 expression, (5) usage of *FGFR2*-E18-C3 or -E18-C4, and/or (6) previously annotated *FGFR2* fusions [91]. *FGFR2* amplification and mutation information was obtained from the cBioPortal [92]. CN segment files for CN break information and exon-level expression data were obtained from the NCI-GDC data portal (<https://portal.gdc.cancer.gov/>). Among the samples with *FGFR2* CN BPs at I17, samples with CN_{E1-E17} segment values ($\log_2[CN/2] > 0.3$ (typical GISTIC threshold for amplifications) and $CN_{E1-E17} - CN_{E18} > 0.3$ were defined as *FGFR2* E1–E17 partially amplified. To select samples with loss of *FGFR2*-E18 expression, E18 expression was normalized to the median expression of E1–E17. Tumour samples showing lower *FGFR2*-E18 expression compared with the minimum expression observed in TCGA normal tissue samples were selected. To evaluate E18-C3 and E18-C4 use, we obtained splice junction read counts from the NCI-GDC data portal. *FGFR2*-E17 to E18-C3 and E18-C4 spanning read counts were divided by total junction reads from *FGFR2*-E17 to calculate E18-C3 and E18-C4 use. Tumour samples showing higher *FGFR2*-E18-C3 or -E18-C4

use compared with the maximum usage observed in TCGA normal tissue samples were selected. In total, the selection process yielded 165 samples for which raw RNA-seq data were downloaded from the NCI-GDC data portal using TCGAbiolinks (v. 2.14.1; ref. [93]). Sequencing reads were mapped to the human reference genome GRCh38 (Ensembl v32 CTAT) using STAR (v.2.7.2; ref. [84]) with the recommended parameters to subsequently run STAR-Fusion (v.1.8.1; ref. [89]). STAR-Fusion was executed with chimeric alignment information (Chimeric.out.junction) derived from STAR to obtain high-confidence in-frame and out-of-frame (frameshift or fusion with non-coding RNA) gene fusions. STAR-Fusion uses only exon–exon spanning reads to detect gene fusions; we therefore used exon–intron/intron–intron spanning reads from the Chimeric.out.junction file to find non-canonical types of out-of-frame fusions applying several filtering steps. Chimeric spanning reads with *FGFR2* BPs were discarded, if we found (1) multiple chimeric alignments, (2) PCR duplicates and/or (3) mitochondrial/Immunoglobulin/HLA mapping. Out-of-frame REs were defined by either exon–exon spanning reads resulting in frameshift or fusion with non-coding RNA (STAR-Fusion) or exon–intron/intron–intron spanning reads (STAR chimeric alignments). Intergenic REs were defined by spanning reads between *FGFR2* and an IGR. Out-of-strand REs were defined by spanning reads between *FGFR2* and an antisense partner gene. REs with recurrent BP support were considered (spanning read count > 2). For the samples with multiple *FGFR2* REs, the relative expression of each RE was computed on the basis of the supporting junction read counts. An E17 junction read frequency of >15% was used as the threshold to define samples with *FGFR2*^{ΔE18} REs. IGV (v.1.11.0; ref. [85]) was used to generate sashimi plots.

Analyses of CCLE, CTRPv2 and GDSC pharmacogenomic datasets

Mutation, CN, gene expression, exon usage ratio and fusion data for cell lines of the Broad Institute Cancer Cell Line Encyclopedia (CCLE) were obtained from the CCLE data portal [30]. *FGFR2/3* missense hotspot mutations were selected in agreement with previous annotations [92,94] and, in the case of *FGFR2*, based on the FMI cohort (Extended Data Fig. 2a). Missense mutations affecting the following amino acids were considered to be hotspots: *FGFR2*, Ser252, Cys382, Asn549, Lys659; *FGFR3*, Arg248, Ser249, Tyr373, Lys650. CN data were obtained as $\log_2[\text{CN}/2]$ values, and $\log_2[\text{CN}/2] \geq 2$ was considered to be an amplification. *FGFR* fusion data (CCLE_Fusions_unfiltered_20181130.txt) were further cleaned by applying the following filters: (1) FPM > 0.1, (2) spanning fragment count ≥ 5 and (3) expression value RPKM ≥ 1 . *FGFR2/3* was considered to be E18-truncated if cell lines contained *FGFR2/3* fusions with I17 BPs or exhibited high *FGFR2*-E18-C3 use ($P < 0.01$ derived from Z-score normalization of exon usage ratio) among the samples with robust expression of *FGFR2*. To compute composite expression of FGF receptors, *FGFR1–4* expression was normalized by the geometric mean of each receptor among all of the samples and summed as previously described [33]. Drug-response data for AZD4547 and PD173074 were obtained from the Cancer Therapeutics Response Portal (CTRP) v2 deposited in the PharmacDB database [31,95] and from the Genomics of Drug Sensitivity in Cancer (GDSC) database [32],

respectively. Integrated area under the sigmoid-fit concentration-response curve values were used to evaluate the association between *FGF/FGFR* status and drug sensitivity.

Low-coverage WGS of human cancer cell lines

Genomic DNA from cultured cells was isolated using the ISOLATE II Genomic DNA Kit (BIO-52066, Bioline) according to the manufacturer's guidelines. Low-coverage WGS was performed as previously described [57]. Libraries were sequenced with 65 bp single reads using the HiSeq 2500 System with V4 chemistry (Illumina) operated by the HiSeq Control Software (v.2.2.68, Illumina). Sequencing reads were mapped to the human reference genome GRCh38 using BWA-MEM (v.0.7.5a; ref. [86]). Reads with mapping quality lower than 37 were excluded. The resulting alignments were analysed with QDNAseq (v.1.14.0) using sequence mappability and GC content correction and a bin size of 20,000 bp to generate segmented CN values [96].

RNA-seq analysis of human cancer cell lines

RNA-seq analysis of cultured cells was performed as previously described [68]. In brief, cells were lysed in Buffer RLT (79216, Qiagen) containing 1% 2-mercaptoethanol. Total RNA extraction was performed using the RNeasy Mini Kit (74104, Qiagen) according to the manufacturer's guidelines. The quality and quantity of RNA was assessed using the 2100 Bioanalyzer system and a Nano chip (Agilent). RNA samples with RIN > 8 were processed for polyA-stranded library preparation using the TruSeq RNA Library Prep Kit v2 (RS-122-2001/2, Illumina) according to the manufacturer's instructions, quality-checked with the 2100 Bioanalyzer system using a 7500 chip, and pooled equimolar into a 10 nM sequencing stock solution. Libraries were sequenced with 100 bp paired-end reads using the HiSeq 2500 System with V4 chemistry and operated by HiSeq Control Software. Sequencing reads were mapped to the human reference genome GRCh38 (Gencode v32 CTAT) using STAR (v.2.7.2; ref. [84]) with the recommended parameters to subsequently run STAR-Fusion (v.1.8.1; ref. [89]). Gene- and exon-level expression read counts were quantified by featureCounts (v.1.6.2; ref. [97]) on the basis of gene structures defined in GRCh38. Genes with CPM values greater than 1 in at least 10% of the total number of samples were considered expressed and used for downstream analysis. Read counts for expressed genes were normalized by trimmed mean of *M*-value (TMM) method using edgeR (v.3.26.6; refs. [98,99]). To detect *FGFR2* gene fusions and REs from RNA-seq, we followed the approach as described for TCGA RNA-seq analysis.

Hybrid-capture RNA-seq analysis of FFPE samples

Total RNA from FFPE samples was isolated using the RNeasy DSP FFPE Kit (73604, Qiagen) according to the manufacturer's guidelines. The quality and quantity of RNA was assessed using the Agilent TapeStation system and High Sensitivity D1000 Reagents (Agilent). A total of 20 ng of fragmented total RNA was used for Illumina-compatible cDNA library preparation. First, total RNA was used for reverse transcription and first-strand cDNA

synthesis. After end-repair and adapter ligation, cDNA sequences were selected for enrichment of exonic sequences using biotinylated target specific probes as provided in the TruSeq RNA Exome kit (Illumina). Standard RNA-seq libraries were generated using captured/exome-enriched cDNA. Purified cDNA sequences were amplified using barcoded primers for different samples. Purified libraries were quantified using Qubit Flex Fluorometer (Thermo Fisher Scientific) and sequenced with 2 × 150 bp configurations using the NextSeq 500 or the NovaSeq 6000 Systems (Illumina) operated by NextSeq (v.2.0.2) and NovaSeq (v.1.7.5) control software, respectively. STAR-Fusion (v.1.8.1; ref. [89]) and the human reference genome GRCh37 were used for RNA fusion detection with the default parameters. STAR (v.2.7.3a; ref. [84]) and RSEM (v.1.3.0; ref. [100]) were used for gene and transcript quantification using the default parameters. LeafCutter (v.0.2.9; ref. [101]) and STAR-produced bam files were used to examine intron excision counts for splicing variants. IGV (v.1.11.0; ref. [85]) was used to generate sashimi plots.

PDX models

Model selection and analysis. PDX models were previously characterized by Crown Bioscience [102] and are described in the HuPrime PDX collection (<https://www.crownbio.com/oncology/in-vivo-services/patient-derived-xenograft-pdx-tumor-models>). PDX models were selected on the basis of (1) *FGF3/4/19* amplification, (2) *FGFR2/3* missense hotspot mutations, (3) *FGFR1/2/3* amplification, (4) high expression of *FGFR1/2/3/4* and/or (5) expression of an *FGFR* fusion gene. The PDX models KI0551, LI0612 and LU1901 were included as controls, because each contained a *MET* oncogenic amplification potentially rendering tumours resistant to FGFRi [55,103]. CN and mutation data generated by whole-exome sequencing and raw RNA-seq data of the selected PDX models (Fig. 4c) were obtained from the CrownBio-HuPrime data portal. Sequencing data were derived from non-treated PDXs. Sequencing reads were mapped to the human (GRCh38 Gencode v32 CTAT) and mouse (mm10 Gencode M23) reference genomes to filter out mouse-derived reads using Disambiguate (v.2018.05.03-6; ref. [104]). The remaining human reads were analysed as described for the analysis of human cell line RNA-seq data, composite *FGFR1-4* expression was computed as described for the CCLE RNA-seq analysis, and *FGFR2* gene fusions and REs were detected as described for TCGA RNA-seq analysis. We also implemented fusions/REs previously annotated by Crown Biosciences and deposited in CrownBio-HuPrime into our analysis.

Debio-1347 intervention study. PDX fragments of 2–3 mm in diameter were injected subcutaneously into the right flank of 8-week-old female BALB/cAnNRj-*Foxn1^{nu/nu}* mice (HFK Bioscience and Shanghai Laboratory Animal Center), except for BL5001 and BL5002, for which 8-week-old female NOD.CB17-*Prkdc^{scid}*/NCrHsd mice (Envigo) were used. Mice were twice weekly weighed and monitored for tumour development and, as soon as tumours reached a volume of 200–250 mm³ (measured in two dimensions using callipers; volume = length × width² × 0.5), the mice were randomly allocated to vehicle versus debio-

1347 FGFRi treatment arms. The treatments were performed daily through oral gavage for 12–25 consecutive days using vehicle (1% Kollidon VA64 in demineralized water), 40 mg per kg debio-1347 (Debiopharm) and increased to 60 mg per kg during treatment (BL5001, BL5002), 60 mg per kg debio-1347 (BN2289, BL0597, BR0438, BR1115, CR3151, ES0136, ES0189, ES0204, ES0215, ES0218, ES2116, GA0114, LI0612, LI1035, LI1055, LU0755, PA1332) or 80 mg per kg debio-1347 (CR1428, ES0042, GA0080, GA0087, GA1224, GA3055, GL0720, HN0366, HN0696, HN1420, KI0551, LU1302, LU1380, LU1429, LU1901, LU2504, PA3013). The treatment response was determined by relative treatment-to-control ratios ($\Delta T/\Delta C$). ΔT and ΔC are the mean volume difference between last treatment day and initial treatment day of the treated and control groups, respectively. All of the animal procedures were conducted at a Crown Bioscience SPF facility. All of the procedures related to animal handling, care and the treatment in this intervention study were performed according to guidelines approved by the Institutional Animal Care and Use Committee (IACUC) of Crown Bioscience following the guidance of the Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC).

Analysis of CGP data from FMI

Hybrid-capture-based CGP. Comprehensive genomic profiling (CGP) was performed on FFPE tumour tissue or blood samples prospectively collected during routine clinical care. Testing was performed in a Clinical Laboratory Improvement Amendments-certified, College of American Pathologists-accredited, New York State-regulated reference laboratory (Foundation Medicine). Approval for this study, including a waiver of informed consent and a Health Insurance Portability and Accountability Act waiver of authorization, was obtained from the Western Institutional Review Board (protocol 20152817). For 217,017 tumour tissue specimens, DNA (>50 ng) was extracted from FFPE specimens and next-generation sequencing was performed by FoundationOne companion diagnostic testing using hybridization-captured, adapter-ligation-based libraries to high, uniform coverage (>500×) for all coding exons of 315 or 324 cancer-related genes plus selected introns of 28 or 36 genes that are frequently rearranged in cancer, as previously described [105]. A total of 26,289 samples were similarly assayed; DNA was sequenced for 406 genes and selected introns of 31 genes involved in REs, and RNA was sequenced for 265 genes [106]. For 6,264 liquid samples, plasma was isolated from 20 ml of peripheral whole blood and ≥20 ng of circulating tumour DNA was extracted to create adapted sequencing libraries for coding exons of 70 genes before hybrid-capture and sample-multiplexed sequencing [107]. Results were analysed for base substitutions, short insertions and deletions, CN gains or losses, and REs. The companion diagnostic tests included probes against all *FGFR2* exons and *FGFR2*-117.

Data analysis. FMI classifies *FGFR2* REs as fusions if the genomic BP is in the I17/E18 hotspot, if the predicted chimeric protein includes both an N terminus and a C terminus (in strand), and if the gene partner is either a previously described fusion partner (in-frame

or frame unknown) or a novel gene partner predicted to be in-frame with *FGFR2*. I17/E18 hotspot out-of-strand REs, any REs with a BP in intergenic space and any REs with a BP in E1–E17 of *FGFR2* were classified as REs. Here we reclassified *FGFR2* REs as described for WGS data analysis. In brief, REs were defined as in-frame fusions if the genomic BP was in I17/E18 and if the frame of the fusion partner was predictable and in-strand. Frame-unknown REs, out-of-strand REs and REs with a BP in intergenic space were classified as non-canonical REs. *FGFR2* amplifications were called if $\geq 80\%$ of the *FGFR2* targets were at an amplified CN (defined as $\geq 4 + \text{median ploidy of the sample}$). Differential CN gains of E18 targets < E1–E17 targets were defined as *FGFR2*-E1–E17 partial amplifications. In samples with *FGFR2* REs and co-amplification, low-level REs were discarded at a read threshold dependent on the amplification CN. *FGFR2*-E18-truncating nonsense and frameshift mutations were subgrouped into mutations affecting the proximal (E768–Y783) versus the distal (P784–T821) C terminus (encoded by E18) on the basis of the functional classifications of truncating mutations in this study (Fig. 2e and Extended Data Figs. 6a,h, 7a,b and 8a–c). I17/E18 in-frame fusions or non-canonical REs, E1–E17 partial amplifications, E18 splice-site mutations and/or proximal E18-truncating mutations were grouped as *FGFR2*-E18-truncating alterations. The four most common *FGFR2* missense mutations affecting Ser252, Cys382, Asn549 and Lys659 are referred to as hotspots throughout this study (Extended Data Fig. 2a), in agreement with previous annotations [92,94]. To establish the co-driver landscape of *FGFR2*-altered tumours, the top 30 driver genes concurrently altered (amplifications, deletions, and missense, truncating and splice-site mutations) in samples with *FGFR2* alterations (E18 truncations, E1–E18 full-length amplifications and/or missense hotspot mutations) were identified. The samples were grouped according to *FGFR2*-E18-truncating alterations, *FGFR2*-E1–E18 full-length amplifications or *FGFR2* missense hotspot mutations, and proportion Z-tests were used to identify co-driver genes significantly enriched in either of the 3 *FGFR2* alteration categories, both in the pan-cancer cohort as well as in the BRCA, CHOL, OV, COAD/READ, ESCA/STAD and LUAD/LUSC cohorts specifically. Fisher's exact tests were used to evaluate co-occurrence (odds ratio > 1) or mutual exclusivity (odds ratio < 1) of the co-driver genes in each of the 3 *FGFR2* alteration categories versus *FGFR2* WT samples in the pan-cancer cohort and in the BRCA cohort specifically.

Analysis of self-interacting capacity among *FGFR2* RE partners

The SLIPPER algorithm predicts the interaction capacities of proteins [108]. It has been trained with seven different proteome databases (DIP, IntAct, MINT, BioGRID, PDB, MatrixDB and 12D) to establish the SLIPPER Golden Standard Dataset of potentially self-interacting proteins [108]. On the basis of this dataset, the proportion of self-interactors among unique proteins encoded by *FGFR2* RE partner genes identified in the FMI dataset was calculated. The self-interacting ability of *FGFR2* RE partners was also evaluated using the SLIPPER algorithm [108] itself. To identify specific self-interacting domains in *FGFR2* RE partners, domain–domain interaction information was obtained from the 3did [109] and PPIDM

[110] databases, and domain enrichment analysis was performed with DAVID bioinformatic resources [111]. The Swiss-Prot *Homo sapiens* proteome (release 2021_04) was used as reference dataset for these analyses.

Re-examination of FIGHT-202 study

Details on the study design, eligibility criteria, and efficacy and safety findings of FIGHT-202 (NCT02924376), a phase II, open-label, multicentre, global study of pemigatinib in patients with previously treated advanced or metastatic cholangiocarcinoma, with or without *FGF/FGFR* alterations, were previously published [9]. Before entering screening for trial eligibility, the patients were either prescreened for *FGF/FGFR* status using FoundationOne or patients provided a commercial FoundationOne report or an *FGF/FGFR* status report based on local testing, the latter of which required retrospective central confirmation through FoundationOne. In FIGHT-202, *FGFR2* REs were classified (fusions versus REs) on the basis of the FoundationOne report and biomarker definition as described above. In this reanalysis of the FIGHT-202 oncogenomic data, we used the alteration data provided by FMI to classify *FGFR2* amplification status and *FGFR2* REs by frame only. Five patients who were classified as having fusions on their FoundationOne report had *FGFR2* amplifications in conjunction with fusions in the alteration data; four patients classified as having fusions on their FoundationOne report had an unknown frame in the structured data; and two patients who were classified as having REs on their FoundationOne report were classified as in-frame in the alteration data. These discrepancies are due to FoundationOne reporting rules and ongoing updates to the analysis and annotation pipeline used by FMI from the time of the original report to the time of the generation of the alteration data. Importantly, these changes do not affect the results from the primary efficacy cohort from FIGHT-202 but, rather, provide an alternative classification for this subset analysis.

Statistics and reproducibility

Data of in vitro and in vivo experiments were analysed using Prism (v.9.3.1, GraphPad Software). Genomic and proteomic data were analysed using R (v.3.6.3–4.1.2). In vitro experiments were independently repeated at least twice, and all attempts at replication were successful. Across these, data on $n \geq 3$ independent replica were collected. No sample size calculations were performed. Sample sizes of mouse cohorts and for ex vivo analyses thereof (FACS analyses, H&E and IHC analyses, proteomics, RNA-seq and RT-qPCR) were based on previous calculations [10] or determined using G*Power software (v.3.1; ref. [56]), and were large enough to measure the effect sizes. Data were reproducible across mice or batches analysed, and all attempts at replication were successful. Sample sizes in the FIGHT-202 trial were based on previous calculations [9] and were large enough to measure the effect sizes. The statistical tests and multiple-testing correction models used are described in the corresponding figure legends. $P < 0.05$ was considered to be statistically significant. Except for $P < 0.0001$ and $P \geq 0.05$, exact P values are always shown in the corresponding figure panels or, where indicated, in Supplementary Table 2.

ONLINE CONTENT

Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information, details of author contributions and competing interests, and statements of data and code availability are available at <https://doi.org/10.1038/s41586-022-05066-5>.

REFERENCES

1. Katoh, M. (2019) Fibroblast growth factor receptors as treatment targets in clinical oncology. *Nat Rev Clin Oncol* 16, 105–122
2. Pearson, A. *et al.* (2016) High-Level Clonal FGFR Amplification and Response to FGFR Inhibition in a Translational Clinical Trial. *Cancer Discov* 6, 838–851
3. Van Cutsem, E. *et al.* (2017) A randomized, open-label study of the efficacy and safety of AZD4547 monotherapy versus paclitaxel for the treatment of advanced gastric adenocarcinoma with FGFR2 polysomy or gene amplification. *Ann Oncol* 28, 1316–1324
4. Chae, Y.K. *et al.* (2020) Phase II Study of AZD4547 in Patients With Tumors Harboring Aberrations in the FGFR Pathway: Results From the NCI-MATCH Trial (EAY131) Subprotocol W. *J Clin Oncol* 38, 2407–2417
5. Taberner, J. *et al.* (2015) Phase I Dose-Escalation Study of JNJ-42756493, an Oral Pan-Fibroblast Growth Factor Receptor Inhibitor, in Patients With Advanced Solid Tumors. *J Clin Oncol* 33, 3401–3408
6. Javle, M. *et al.* (2018) Phase II Study of BGJ398 in Patients With FGFR-Altered Advanced Cholangiocarcinoma. *J Clin Oncol* 36, 276–282
7. Voss, M.H. *et al.* (2019) A Phase I, Open-Label, Multicenter, Dose-escalation Study of the Oral Selective FGFR Inhibitor Debio 1347 in Patients with Advanced Solid Tumors Harboring FGFR Gene Alterations. *Clin Cancer Res* 25, 2699–2707
8. Subbiah, V. *et al.* (2022) FIGHT-101, a first-in-human study of potent and selective FGFR 1-3 inhibitor pemigatinib in pan-cancer patients with FGF/FGFR alterations and advanced malignancies. *Ann Oncol* 33, 522–533
9. Abou-Alfa, G.K. *et al.* (2020) Pemigatinib for previously treated, locally advanced or metastatic cholangiocarcinoma: a multicentre, open-label, phase 2 study. *Lancet Oncol* 21, 671–684
10. Kas, S.M. *et al.* (2017) Insertional mutagenesis identifies drivers of a novel oncogenic pathway in invasive lobular breast carcinoma. *Nat Genet* 49, 1219–1230
11. de Ruiter, J.R. *et al.* (2017) Identifying transposon insertions and their effects from RNA-sequencing data. *Nucleic Acids Res* 45, 7064–7077
12. Annunziato, S. *et al.* (2016) Modeling invasive lobular breast carcinoma by CRISPR/Cas9-mediated somatic genome editing of the mammary gland. *Genes Dev* 30, 1470–1480
13. Annunziato, S. *et al.* (2019) Comparative oncogenomics identifies combinations of driver genes and drug targets in BRCA1-mutated breast cancer. *Nat Commun* 10, 397
14. Szybowska, P. *et al.* (2021) Negative Regulation of FGFR (Fibroblast Growth Factor Receptor) Signaling. *Cells* 10, 1342
15. Tannheimer, S.L. *et al.* (2000) Characterization of fibroblast growth factor receptor 2 overexpression in the human breast cancer cell line SUM-52PE. *Breast Cancer Res* 2, 311–320
16. Itoh, H. *et al.* (1994) Preferential alternative splicing in cancer generates a K-sam messenger RNA with higher transforming activity. *Cancer Res* 54, 3237–3241
17. Ueda, T. *et al.* (1999) Deletion of the carboxyl-terminal exons of K-sam/FGFR2 by short homology-mediated recombination, generating preferential expression of specific messenger RNAs. *Cancer Res* 59, 6080–6086
18. Sakaguchi, K. *et al.* (1999) Identification of a novel activated form of the keratinocyte growth factor receptor by expression cloning from parathyroid adenoma tissue. *Oncogene* 18, 5497–5505
19. Lorenzi, M.V. *et al.* (1996) FRAG1, a gene that potently activates fibroblast growth factor receptor by C-terminal fusion through chromosomal rearrangement. *Proc Natl Acad Sci U S A* 93, 8956–8961
20. Wu, Y.-M. *et al.* (2013) Identification of targetable FGFR gene fusions in diverse cancers. *Cancer Discov* 3, 636–647

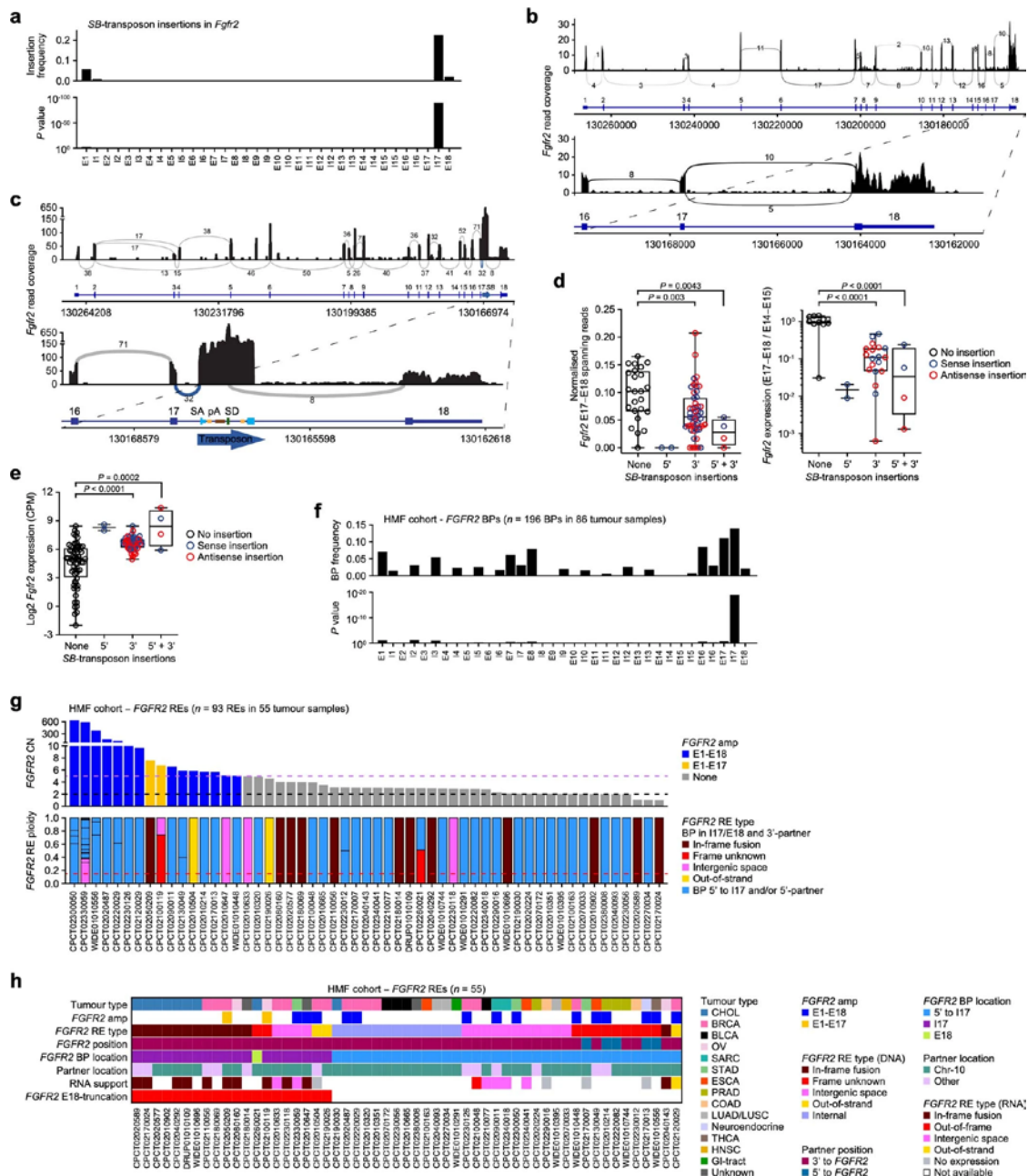
21. Jusakul, A. *et al.* (2017) Whole-Genome and Epigenomic Landscapes of Etiologically Distinct Subtypes of Cholangiocarcinoma. *Cancer Discov* 7, 1116–1135
22. Qin, A. *et al.* (2019) Detection of Known and Novel FGFR Fusions in Non-Small Cell Lung Cancer by Comprehensive Genomic Profiling. *J Thorac Oncol* 14, 54–62
23. Priestley, P. *et al.* (2019) Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature* 575, 210–216
24. Li, Y. *et al.* (2020) Patterns of somatic structural variation in human cancer genomes. *Nature* 578, 112–121
25. Lorenzi, M.V. *et al.* (1997) Ligand-independent activation of fibroblast growth factor receptor-2 by carboxyl terminal alterations. *Oncogene* 15, 817–826
26. Moffa, A.B. *et al.* (2004) Transforming potential of alternatively spliced variants of fibroblast growth factor receptor 2 in human mammary epithelial cells. *Mol Cancer Res* 2, 643–652
27. Cha, J.Y. *et al.* (2009) Aberrant receptor internalization and enhanced FRS2-dependent signaling contribute to the transforming activity of the fibroblast growth factor receptor 2 IIIb C3 isoform. *J Biol Chem* 284, 6227–6240
28. Rothé, B. *et al.* (2018) Crystal Structure of Bicc1 SAM Polymer and Mapping of Interactions between the Ciliopathy-Associated Proteins Bicc1, ANKS3, and ANKS6. *Structure* 26, 209–224.e6
29. Ciriello, G. *et al.* (2015) Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell* 163, 506–519
30. Ghandi, M. *et al.* (2019) Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* 569, 503–508
31. Seashore-Ludlow, B. *et al.* (2015) Harnessing Connectivity in a Large-Scale Small-Molecule Sensitivity Dataset. *Cancer Discov* 5, 1210–1223
32. Iorio, F. *et al.* (2016) A Landscape of Pharmacogenomic Interactions in Cancer. *Cell* 166, 740–754
33. Sánchez-Guixé, M. *et al.* (2022) High FGFR1-4 mRNA Expression Levels Correlate with Response to Selective FGFR Inhibitors in Breast Cancer. *Clin Cancer Res* 28, 137–149
34. Turner, N. *et al.* (2010) Integrative molecular profiling of triple negative breast cancers identifies amplicon drivers and potential therapeutic targets. *Oncogene* 29, 2013–2023
35. Mathur, A. *et al.* (2014) FGFR2 is amplified in the NCI-H716 colorectal cancer cell line and is required for growth and survival. *PLoS One* 9, e98515
36. Hart, K.C. *et al.* (2001) Identification of tyrosine residues in constitutively activated fibroblast growth factor receptor 3 involved in mitogenesis, Stat activation, and phosphatidylinositol 3-kinase activation. *Mol Biol Cell* 12, 931–942
37. Nadratowska-Wesolowska, B. *et al.* (2014) RSK2 regulates endocytosis of FGF receptor 1 by phosphorylation on serine 789. *Oncogene* 33, 4823–4836
38. Szybowska, P. *et al.* (2019) Cancer Mutations in FGFR2 Prevent a Negative Feedback Loop Mediated by the ERK1/2 Pathway. *Cells* 8, 518
39. Lin, C.-C. *et al.* (2012) Inhibition of basal FGF receptor signaling by dimeric Grb2. *Cell* 149, 1514–1524
40. Lin, C.-C. *et al.* (2020) Regulation of kinase activity by combined action of juxtamembrane and C-terminal regions of receptors [bioRxiv, 2020.10.01.322123](https://doi.org/10.1101/322123)
41. Khazaie, K. *et al.* (1988) Truncation of the human EGF receptor leads to differential transforming potentials in primary avian fibroblasts and erythroblasts. *EMBO J* 7, 3061–3071
42. Akiyama, T. *et al.* (1991) The transforming potential of the c-erbB-2 protein is regulated by its autophosphorylation at the carboxyl-terminal domain. *Mol Cell Biol* 11, 833–842
43. Woolford, J. *et al.* (1988) Activation of the feline c-fms proto-oncogene: multiple alterations are required to generate a fully transformed phenotype. *Cell* 55, 965–977
44. Niu, X.-L. *et al.* (2002) Deletion of the carboxyl terminus of Tie2 enhances kinase activity, signaling, and function. Evidence for an autoinhibitory mechanism. *J Biol Chem* 277, 31768–31773

45. Hung, K.L. *et al.* (2021) ecDNA hubs drive cooperative intermolecular oncogene expression. *Nature* 600, 731–736
46. Ota, S. *et al.* (2009) The role of senescence and prosurvival signaling in controlling the oncogenic activity of FGFR2 mutants associated with cancer and birth defects. *Hum Mol Genet* 18, 2609–2621
47. Hertzler-Schaefer, K. *et al.* (2014) Pten loss induces autocrine FGF signaling to promote skin tumorigenesis. *Cell Rep* 6, 818–826
48. Koziczak, M. *et al.* (2004) Blocking of FGFR signaling inhibits breast cancer cell proliferation through downregulation of D-type cyclins. *Oncogene* 23, 3501–3508
49. Liu, H. *et al.* (2017) c-Myc Alteration Determines the Therapeutic Response to FGFR Inhibitors. *Clin Cancer Res* 23, 974–984
50. Formisano, L. *et al.* (2019) Aberrant FGFR signaling mediates resistance to CDK4/6 inhibitors in ER+ breast cancer. *Nat Commun* 10, 1373
51. Derksen, P.W.B. *et al.* (2006) Somatic inactivation of E-cadherin and p53 in mice leads to metastatic lobular mammary carcinoma through induction of anoikis resistance and angiogenesis. *Cancer Cell* 10, 437–449
52. Derksen, P.W.B. *et al.* (2011) Mammary-specific inactivation of E-cadherin and p53 impairs functional gland development and leads to pleomorphic invasive lobular carcinoma in mice. *Dis Model Mech* 4, 347–358
53. Schipper, K. *et al.* (2019) Rebalancing of actomyosin contractility enables mammary tumor formation upon loss of E-cadherin. *Nat Commun* 10, 3800
54. Huijbers, I.J. *et al.* (2015) Using the GEMM-ESC strategy to study gene function in mouse models. *Nat Protoc* 10, 1755–1785
55. Kas, S.M. *et al.* (2018) Transcriptomics and Transposon Mutagenesis Identify Multiple Mechanisms of Resistance to the FGFR Inhibitor AZD4547. *Cancer Res* 78, 5668–5679
56. Faul, F. *et al.* (2009) Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods* 41, 1149–1160
57. Henneman, L. *et al.* (2015) Selective resistance to the PARP inhibitor olaparib in a mouse model for BRCA1-deficient metaplastic breast cancer. *Proc Natl Acad Sci U S A* 112, 8409–8414
58. Cardiff, R.D. *et al.* (2000) The mammary pathology of genetically engineered mice: the consensus report and recommendations from the Annapolis meeting. *Oncogene* 19, 968–988
59. Montini, E. *et al.* (2009) The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. *J Clin Invest* 119, 964–975
60. Novoradovsky, A. *et al.* (2005) Computational Principles of Primer Design for Site Directed Mutagenesis. 2005 NSTI Nanotechnology Conference and Trade Show - NSTI Nanotech 2005 Technical Proceedings
61. Follenzi, A. *et al.* (2000) Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences. *Nat Genet* 25, 217–222
62. Zingg, D. *et al.* (2018) EZH2-Mediated Primary Cilium Deconstruction Drives Metastatic Melanoma Formation. *Cancer Cell* 34, 69–84.e14
63. Zhang, J. *et al.* (2012) Translating the therapeutic potential of AZD4547 in FGFR1-amplified non-small cell lung cancer through the use of patient-derived tumor xenograft models. *Clin Cancer Res* 18, 6658–6667
64. Guagnano, V. *et al.* (2012) FGFR genetic alterations predict for sensitivity to NVP-BGJ398, a selective pan-FGFR inhibitor. *Cancer Discov* 2, 1118–1133
65. Nakanishi, Y. *et al.* (2014) The fibroblast growth factor receptor genetic status as a potential predictor of the sensitivity to CH5183284/Debio 1347, a novel selective FGFR inhibitor. *Mol Cancer Ther* 13, 2547–2558

66. Liu, P.C.C. *et al.* (2020) INCB054828 (pemigatinib), a potent and selective inhibitor of fibroblast growth factor receptors 1, 2, and 3, displays activity against genetically defined tumor models. *PLoS One* 15, e0231877
67. Delwel, G.O. *et al.* (1993) Expression and function of the cytoplasmic variants of the integrin alpha 6 subunit in transfected K562 cells. Activation-dependent adhesion and interaction with isoforms of laminin. *J Biol Chem* 268, 25865–25875
68. Boelens, M.C. *et al.* (2016) PTEN Loss in E-Cadherin-Deficient Mouse Mammary Epithelial Cells Rescues Apoptosis and Results in Development of Classical Invasive Lobular Carcinoma. *Cell Rep* 16, 2087–2101
69. Ye, J. *et al.* (2012) Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* 13, 134
70. Schindelin, J. *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9, 676–682
71. Rolfs, F. *et al.* (2021) Feasibility of Phosphoproteomics on Leftover Samples After RNA Extraction With Guanidinium Thiocyanate. *Mol Cell Proteomics* 20, 100078
72. Gogola, E. *et al.* (2018) Selective Loss of PARG Restores PARylation and Counteracts PARP Inhibitor-Mediated Synthetic Lethality. *Cancer Cell* 33, 1078-1093.e12
73. Beekhof, R. *et al.* (2019) INKA, an integrative data analysis pipeline for phosphoproteomic inference of active kinases. *Mol Syst Biol* 15, e8250
74. Böttger, F. *et al.* (2019) Tumor Heterogeneity Underlies Differential Cisplatin Sensitivity in Mouse Models of Small-Cell Lung Cancer. *Cell Rep* 27, 3345-3358.e4
75. Cox, J. and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 26, 1367–1372
76. Tyanova, S. *et al.* (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* 11, 2301–2319
77. Reich, M. *et al.* (2006) GenePattern 2.0. *Nat Genet* 38, 500–501
78. Barbie, D.A. *et al.* (2009) Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* 462, 108–112
79. Liberzon, A. *et al.* (2015) The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* 1, 417–425
80. Olsen, J.V. *et al.* (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* 127, 635–648
81. Ritchie, M.E. *et al.* (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43, e47
82. Krug, K. *et al.* (2019) A Curated Resource for Phosphosite-specific Signature Analysis. *Mol Cell Proteomics* 18, 576–593
83. Yilmaz, S. *et al.* (2021) Robust inference of kinase activity using functional networks. *Nat Commun* 12, 1177
84. Dobin, A. *et al.* (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21
85. Robinson, J.T. *et al.* (2011) Integrative genomics viewer. *Nat Biotechnol* 29, 24–26
86. Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760
87. Cameron, D.L. *et al.* (2017) GRIDSS: sensitive and specific genomic rearrangement detection using positional de Bruijn graph assembly. *Genome Res* 27, 2050–2060
88. Cameron, D.L. *et al.* (2019) GRIDSS, PURPLE, LINX: Unscrambling the tumor genome via integrated analysis of structural variation and copy number *bioRxiv*, 781013

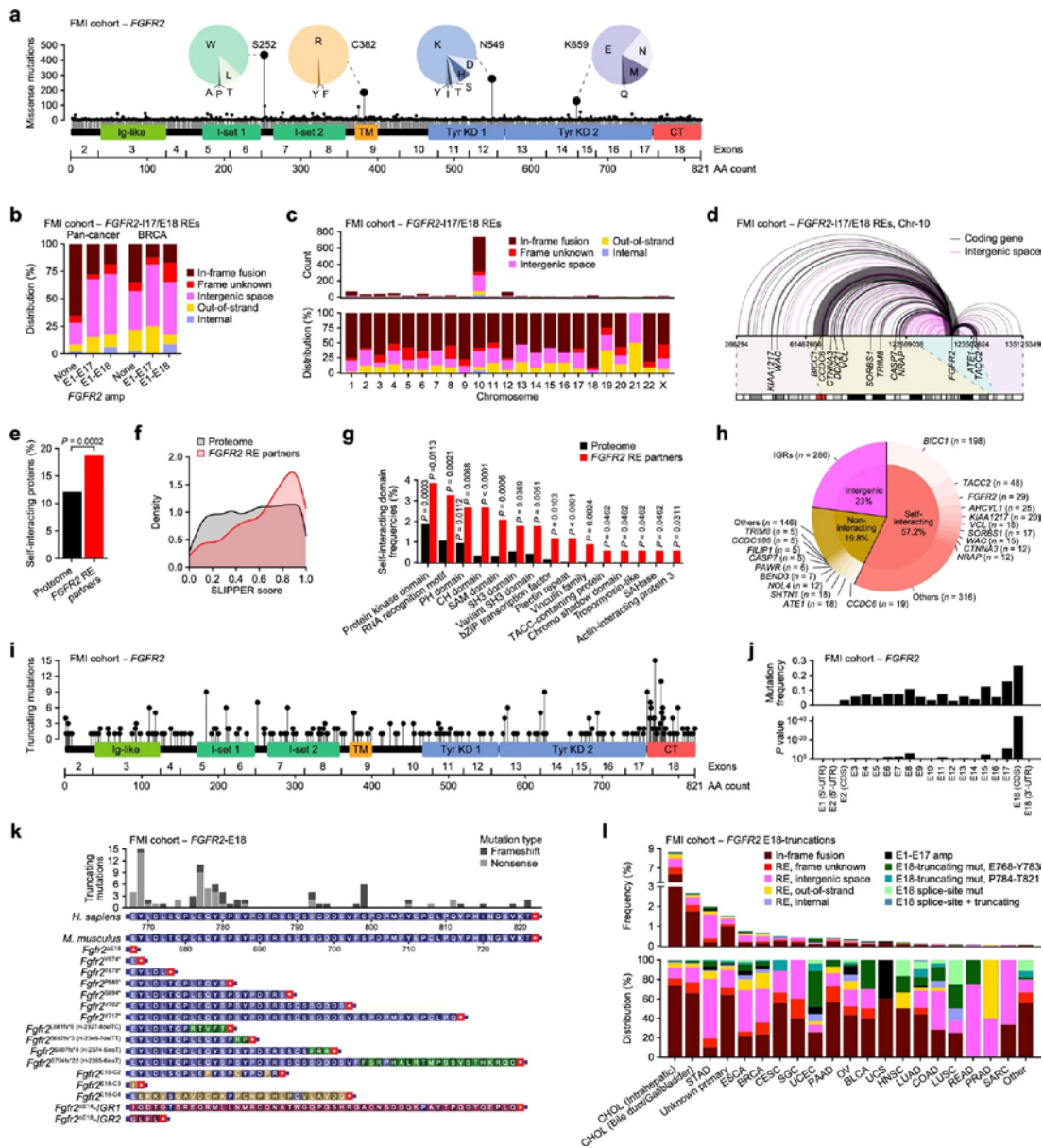
89. Haas, B.J. *et al.* (2019) Accuracy assessment of fusion transcript detection via read-mapping and de novo fusion transcript assembly-based methods. *Genome Biol* 20, 213
90. Cancer Genome Atlas Research Network *et al.* (2013) The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet* 45, 1113–1120
91. Gao, Q. *et al.* (2018) Driver Fusions and Their Implications in the Development and Treatment of Human Cancers. *Cell Rep* 23, 227–238.e3
92. Cerami, E. *et al.* (2012) The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2, 401–404
93. Colaprico, A. *et al.* (2016) TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res* 44, e71
94. Nakamura, I.T. *et al.* (2021) Comprehensive functional evaluation of variants of fibroblast growth factor receptor genes in cancer. *NPJ Precis Oncol* 5, 66
95. Smirnov, P. *et al.* (2018) PharmacDB: an integrative database for mining in vitro anticancer drug screening studies. *Nucleic Acids Res* 46, D994–D1002
96. Scheinin, I. *et al.* (2014) DNA copy number analysis of fresh and formalin-fixed specimens by shallow whole-genome sequencing with identification and exclusion of problematic regions in the genome assembly. *Genome Res* 24, 2022–2032
97. Liao, Y. *et al.* (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930
98. Robinson, M.D. and Oshlack, A. (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* 11, R25
99. Robinson, M.D. *et al.* (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140
100. Li, B. and Dewey, C.N. (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12, 323
101. Li, Y.I. *et al.* (2018) Annotation-free quantification of RNA splicing using LeafCutter. *Nat Genet* 50, 151–158
102. Chen, D. *et al.* (2019) In Vivo Pharmacology Models for Cancer Target Research. *Methods Mol Biol* 1953, 183–211
103. Kim, S.-M. *et al.* (2016) Activation of the Met kinase confers acquired drug resistance in FGFR-targeted lung cancer therapy. *Oncogenesis* 5, e241
104. Ahdesmäki, M.J. *et al.* (2017) Disambiguate: An open-source application for disambiguating two species in next generation sequencing data from grafted samplesF1000Research
105. Frampton, G.M. *et al.* (2013) Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. *Nat Biotechnol* 31, 1023–1031
106. He, J. *et al.* (2016) Integrated genomic DNA/RNA profiling of hematologic malignancies in the clinical setting. *Blood* 127, 3004–3014
107. Clark, T.A. *et al.* (2018) Analytical Validation of a Hybrid Capture-Based Next-Generation Sequencing Clinical Assay for Genomic Profiling of Cell-Free Circulating Tumor DNA. *J Mol Diagn* 20, 686–702
108. Liu, Z. *et al.* (2013) Proteome-wide prediction of self-interacting proteins based on multiple properties. *Mol Cell Proteomics* 12, 1689–1700
109. Stein, A. *et al.* (2005) 3did: interacting protein domains of known three-dimensional structure. *Nucleic Acids Res* 33, D413–417
110. Alborzi, S.Z. *et al.* (2021) PPIDomainMiner: Inferring domain-domain interactions from multiple sources of protein-protein interactions. *PLoS Comput Biol* 17, e1008844
111. Huang, D.W. *et al.* (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4, 44–57

EXTENDED DATA FIGURES



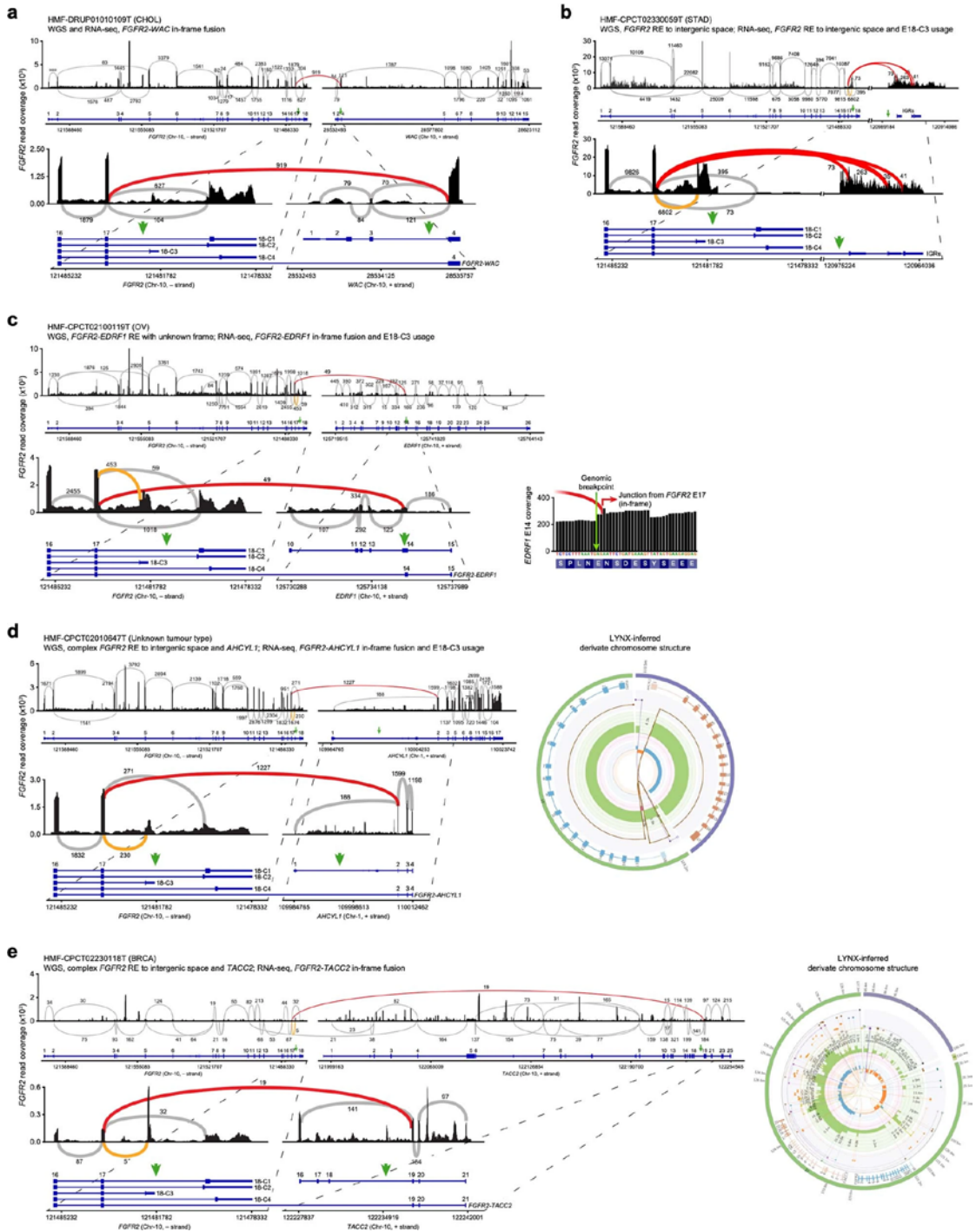
Extended Data Fig. 1: SB-insertions in *Fgfr2* and *FGFR2* REs found in the FMI cohort.

a, Normalized frequency (top panel) and enrichment significance (*P* values, bottom panel) of *Sleeping Beauty* (*SB*) transposon insertions (*n*=81 insertions in 65 tumours) in each *Fgfr2* exon (E) and intron (I) as identified in mammary tumours from a *SB*-transposon in vivo screen [10]. *SB* insertion frequency was normalized by the kilobase of feature (exon/intron) length and total number of *SB*-insertions. **b, c**, Sashimi plots showing *Fgfr2* read coverage and junction reads plotted as arcs with indicated junction read counts of tumours with no *SB*-insertion (**b**) and an I17 sense *SB* insertion (**c**) in *Fgfr2*. SA, splice acceptor; SD, splice donor; pA, polyadenylation signal. **d**, Left panel, counts of *Fgfr2*-E17–E18 spanning reads (counts per million, CPM) normalized to *Fgfr2* expression (CPM) in *SB* tumour RNA sequencing (RNA-seq) profiles (none, *n*=24; 5'-sense, *n*=2; 3'-sense, *n*=22; 3'-antisense, *n*=27; 5' + 3'-sense, *n*=2; 5' + 3'-antisense, *n*=2); right panel, RT-qPCR to quantify *Fgfr2*-E17–E18 over *Fgfr2*-E14–E15 expression in *SB*-tumours (none, *n*=10; 5'-sense, *n*=2; 3'-sense, *n*=8; 3'-antisense, *n*=11; 5' + 3'-sense, *n*=2; 5' + 3'-antisense, *n*=2; individual dots represent mean of 3 independent measurements). **e**, Expression of *Fgfr2* (CPM) in *SB* tumour RNA-seq profiles (none, *n*=64; 5'-sense, *n*=2; 3'-sense, *n*=30; 3'-antisense, *n*=28; 5' + 3'-sense, *n*=2; 5' + 3'-antisense, *n*=2). **f**, Normalized frequency (top panel) and enrichment significance (*P* values, bottom panel) of *FGFR2* genomic rearrangement (RE) breakpoints (BPs) identified in 2,112 whole-genome sequencing (WGS) profiles from the Hartwig Medical Foundation (HMF) cohort in each exon/intron. BP frequency was normalized by the kilobase of feature (exon/intron) length and total number of BPs. **g**, *FGFR2* copy numbers (CN, top panel) and RE ploidy frequencies (bottom panel) in samples with *FGFR2* REs. *FGFR2* BPs resulting in unresolved REs were excluded generating a refined list of REs (*n*=93 REs in 55 tumour samples). Dotted lines, black, normal CN; purple, amplified CN (> 5); red, RE ploidy frequency threshold (> 0.15) to call samples with E18-truncating *FGFR2* REs (E18-truncating, *n*=20; others, *n*=35). Amp, amplification. **h**, *FGFR2* RE types found in WGS profiles from HMF. RNA support indicates evidence for *FGFR2* REs in matching RNA-seq profiles. Empty fields, no RNA-seq data available. BLCA, bladder urothelial carcinoma; BRCA, breast invasive carcinoma; CHOL, cholangiocarcinoma; chr, chromosome; COAD, colon adenocarcinoma; ESCA, oesophageal carcinoma; GI, gastrointestinal; HNSC, head and neck squamous cell carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; OV, ovarian serous cystadenocarcinoma; PRAD, prostate adenocarcinoma; SARC, sarcoma; STAD, stomach adenocarcinoma; THCA, thyroid carcinoma [90]. Data in **d, e** are represented as median (centre line) $\pm$ interquartile range (IQR, 25th to 75th percentile, box) and $\pm$ full range (minimum to maximum, whiskers). *P* values were calculated with one-tailed binomial tests (**a, f**) or one-tailed one-way analysis of variance (ANOVA) and Tukey's multiple-testing corrections (**d, e**).



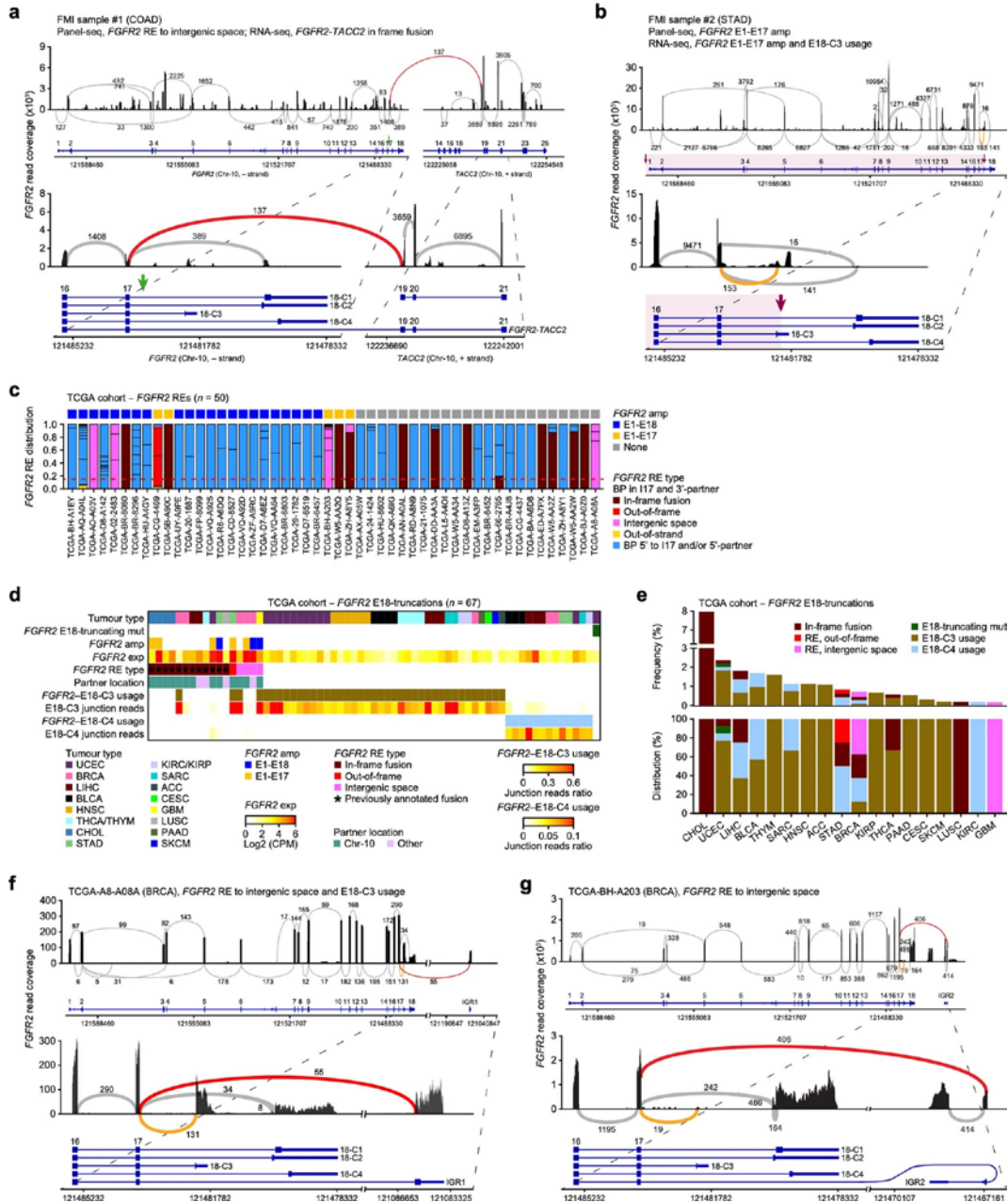
Extended Data Fig. 2: *FGFR2* alterations found in the HMF cohort.

a, Lollipop plot of *FGFR2* missense mutations identified in the Foundation Medicine (FMI) pan-cancer cohort (249,570 diagnostic hybrid-capture panel-seq profiles). The top four recurrent mutations (Ser 252, Cys 382, Asn 549, Lys 659) are referred to as hotspots in this study. **b**, Distribution of *FGFR2*-I17/E18 RE types in FMI samples with *FGFR2* normal CN, E1-E17 amp, and E1-E18 amp. **c**, Total numbers and distributions of *FGFR2*-I17/E18 RE types across chromosomes according to RE partner location. **d**, Linear chr-10 map depicting intrachromosomal *FGFR2*-I17/E18 REs. Thickness of arcs is proportional to the recurrence of the corresponding RE partners. Light / dark grey and red bars denote ideogram and centromere of chr-10. **e**, Percentage of unique proteins with self-interacting capacity among *FGFR2* RE partners ($n = 337$) versus the human proteome ($n = 20,385$). Based on the SLIPPER Golden Standard Dataset of self-interactors [108]. **f**, Distribution of self-interaction scores among *FGFR2* RE partners using the SLIPPER algorithm [108]. **g**, Enrichment of self-interacting protein domains among *FGFR2* RE partners using DAVID [111]. **h**, Recurrence of *FGFR2* RE partners grouped by presence of self-interacting domains. Full list of RE partners is disclosed in Supplementary Table 2. IGRs, intergenic regions. **i, j**, Lollipop plot (**i**) and normalized frequency (top panel) and enrichment significance (P values, bottom panel) (**j**) of *FGFR2*-truncating mutations identified in the FMI cohort. Mutation frequency was normalized by the kilobase of feature (exon/intron) length and total number of mutations. AA, amino acid; CDS, coding sequence; CT, C terminus; TM, trans-membrane; UTR, untranslated region. **k**, Distribution of *FGFR2*-E18-truncating mutations identified in the FMI cohort and corresponding cloned mouse *Fgfr2* variants representing most frequent human (H) *FGFR2*-E18 nonsense and frameshift (fs) mutations. C terminus sequences of cloned noncanonical E18-truncated *Fgfr2* (*Fgfr2*^{ΔE18}) variants are also displayed. IGR1 and IGR2 are based on TCGA-A8-A08A and TCGA-BH-A203 in Extended Data Fig. 5f, g. **l**, Frequencies (top panel) and distributions (bottom panel) per tumour type of E18-truncating *FGFR2* alterations found in the FMI cohort. CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; mut, mutation; PAAD, pancreatic adenocarcinoma; READ, rectum adenocarcinoma; SGC, salivary gland carcinoma; UCEC, uterine corpus endometrial carcinoma; UCS, uterus carcinosarcoma. P values were calculated with a one-tailed proportion z-test (**e**), one-tailed Fisher's exact tests (**g**), or one-tailed binomial tests (**j**).



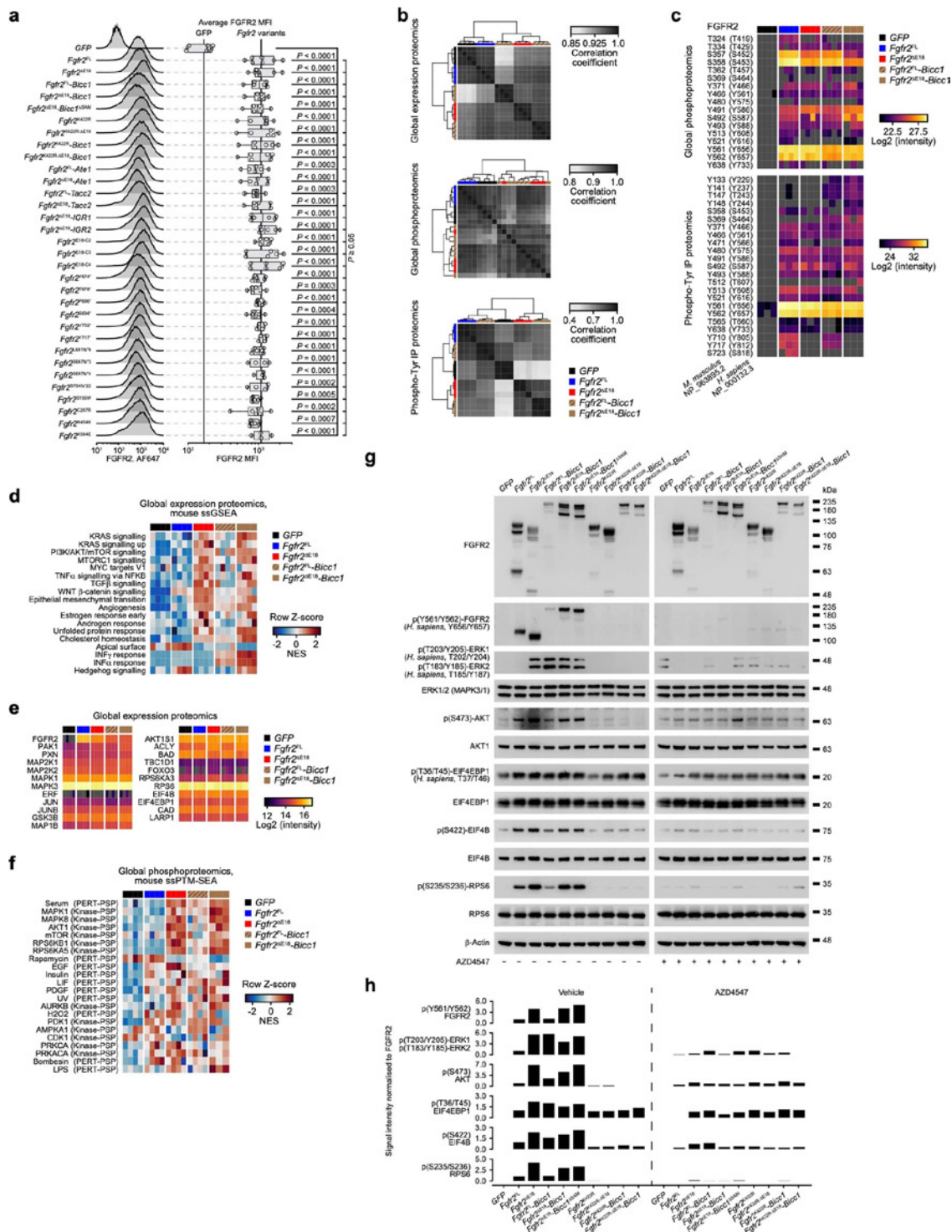
Extended Data Fig. 3: Expression of E18-truncating *FGFR2* variants in HMF samples.

a, Sashimi plot showing *FGFR2* read coverage and junction reads of the HMF sample DRUP01010109T (CHOL). *FGFR2*-*WAC* in-frame fusion identified with WGS and *FGFR2*-E17 to *WAC*-E4 junction confirmed with RNA-seq. **b**, Sashimi plot showing *FGFR2* read coverage and junction reads of the HMF sample CPCT02330059T (STAD). *FGFR2*-I17 RE to intergenic space identified with WGS and *FGFR2*-E17 to intergenic region (IGR) junctions and *FGFR2*-E18-C3 usage found with RNA-seq. **c**, Sashimi plot showing *FGFR2* read coverage and junction reads of the HMF sample CPCT02100119T (OV). *FGFR2*-*EDRF1* frame unknown RE identified with WGS and *FGFR2*-E17 to *EDRF1*-E14 in-frame junction and *FGFR2*-E18-C3 usage found with RNA-seq. **d**, Sashimi plot showing *FGFR2* read coverage and junction reads of the HMF sample CPCT02010647T (unknown tumour type). *FGFR2*-I17 RE to intergenic space identified with WGS and discordant *FGFR2*-*AHCYL1* in-frame fusion with *FGFR2*-E17 to *AHCYL1*-E2 junction and *FGFR2*-E18-C3 usage found with RNA-seq. **e**, Sashimi plot showing *FGFR2* read coverage and junction reads of the HMF sample CPCT02230118T (BRCA). *FGFR2*-I17 RE to intergenic space identified with WGS and discordant *FGFR2*-*TACC2* in-frame fusion with *FGFR2*-E17 to *TACC2*-E19 junction found with RNA-seq. Reconstructed derivative chromosomes using LINX [88] are displayed for CPCT02010647T (**d**) and CPCT02230118T (**e**) and depict complex *FGFR2* REs involving intergenic space and ultimately resolving to *AHCYL1*-E2 (**d**) and *TACC2*-E19 (**e**). Green arrows indicate BPs identified with WGS. E18-C1, canonical E18 of *FGFR2*²⁵; E18-C2/C3/C4, alternative *FGFR2*-E18.



Extended Data Fig. 4: Expression of E18-truncating *FGFR2* variants in FMI and TCGA samples.

a, Sashimi plot showing *FGFR2* read coverage and junction reads of the FMI sample #1 (COAD). *FGFR2*-I17 RE to intergenic space was diagnosed by FMI, and discordant *FGFR2*-*TACC2* in-frame fusion with *FGFR2*-E17 to *TACC2*-E19 junction was found with hybrid-capture RNA-seq. Green arrows indicate *FGFR2* BP identified with panel-seq. **b**, Sashimi plot showing *FGFR2* read coverage and junction reads of the FMI sample #2 (STAD). *FGFR2*-E1–E17 partial amp was diagnosed by FMI, and high *FGFR2* expression with few E17–E18 junction reads but E18–C3 usage was found with hybrid-capture RNA-seq. Purple arrows indicate partially amplified *FGFR2* region identified with panel-seq. **c**, *FGFR2* amp status and RE type distribution in samples with *FGFR2* REs ($n=50$) found in the pan-cancer cohort from The Cancer Genome Atlas (TCGA, $n=10,344$ samples). Dotted red line, RE read frequency threshold (> 0.15) to call samples expressing *FGFR2*^{ΔE18} REs ($n=17$). **d**, 67 samples (0.65% incidence) containing *FGFR2*^{ΔE18} in-frame fusions ($n=12$, 0.12% incidence), *FGFR2*^{ΔE18} non-canonical REs ($n=5$, 0.05% incidence), proximal *FGFR2*-E18-truncating mutations ($n=1$, 0.01% incidence), and cases with significant *FGFR2*-E18–C3 ($n=40$; E18–C3 usage only, $n=36$, 90% of total, 0.35% incidence; E18–C3 usage + RE, $n=4$, 10% of total, 0.05% incidence) and/or E18–C4 ($n=13$, 0.13% incidence) usage found in TCGA cohort. Asterisks mark previously annotated *FGFR2* in-frame fusions [91]. exp, expression; GBM, glioblastoma multiforme; KIRC, kidney renal clear cell carcinoma; KIRP, kidney renal papillary cell carcinoma; LIHC, liver hepatocellular carcinoma; SKCM, skin cutaneous melanoma; THYM, thymoma. **e**, Frequencies (top panel) and distributions (bottom panel) per tumour type of expressed *FGFR2*^{ΔE18} alterations found in TCGA cohort. **f, g**, Sashimi plots showing *FGFR2* read coverage and junction reads of TCGA-BRCA samples A8-A08A (**f**) and BH-A203 (**g**) with identified *FGFR2*-I17 REs to intergenic space and *FGFR2*-E18–C3 usage.



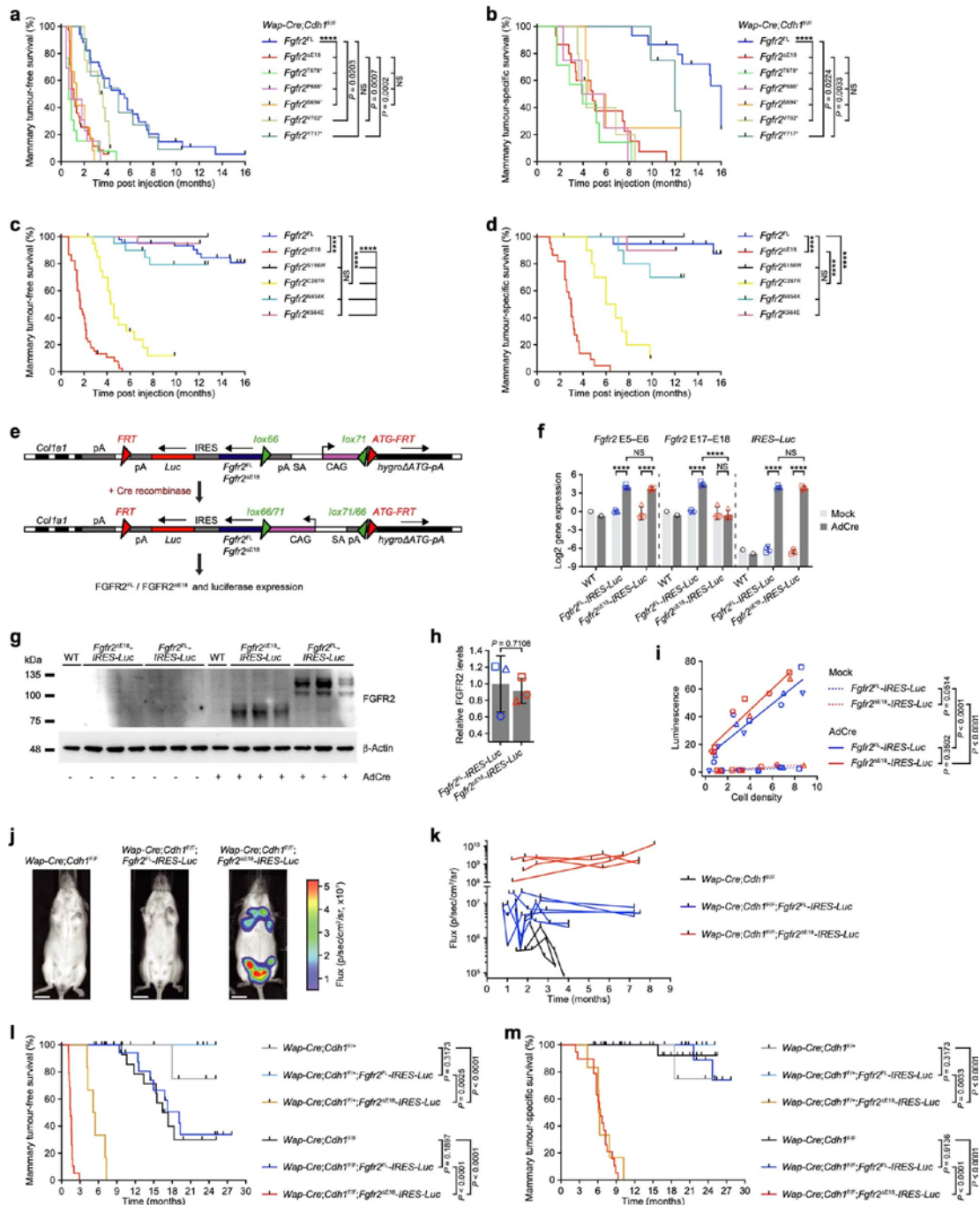
Extended Data Fig. 5: (Phospho)-proteomic analyses of NMuMG cells expressing *Fgfr2* variants.

a, Fluorescence-activated cell sorting (FACS) to analyse FGFR2 mean fluorescence intensity (MFI) in NMuMG cells expressing GFP or indicated *Fgfr2* variants. *Fgfr2^{FL}*, full-length (FL) *Fgfr2*. *Ate1*, *Bicc1*, and *Tacc2* correspond to the top-recurrent *ATE1*, *BICC1*, and *TACC2* fusion partner genes in Extended Data Fig. 2h. *Bicc1^{ΔSAM}* encodes BICC1 lacking its SAM oligomerisation domain. *Fgfr2^{K422R}* variants encode tyrosine kinase domain (KD)-dead FGFR2 variants. Truncated or alternative C-termini encoded by *IGR1/IGR2*, E18 C2/C3/C4, *Fgfr2^{1674*}*, *Fgfr2^{1678*}*, *Fgfr2^{2686*}*, *Fgfr2^{2694*}*, *Fgfr2^{1702*}*, *Fgfr2^{1717*}*, *Fgfr2^{1681fs*6}*, *Fgfr2^{5687fs*3}*, *Fgfr2^{5697fs*4}*, and *Fgfr2^{5704fs*22}* are displayed in Extended Data Fig. 2k. *Fgfr2^{5156W}*, *Fgfr2^{C287R}*, *Fgfr2^{N454K}*, and *Fgfr2^{K564E}* correspond to the human *FGFR2^{S252W}*, *FGFR2^{C382R}*, *FGFR2^{N549K}*, and *FGFR2^{K659E}* missense hotspot mutations in Extended Data Fig. 2a. Validation of overexpression of *Fgfr2* variants using RT-qPCR is in Supplementary Table 4. Data are represented as median (centre line) ± IQR (25th to 75th percentile, box) and ± full range (minimum to maximum, whiskers) of GFP, *n* = 6; *Fgfr2^{FL}*, *Fgfr2^{ΔE18}*, *n* = 7; *Fgfr2^{1702*}*, *Fgfr2^{1717*}*, *Fgfr2^{5687fs*3}*, *Fgfr2^{5697fs*4}*, *Fgfr2^{5704fs*22}*, *Fgfr2^{K564E}*, *n* = 4; other *Fgfr2* variants, *n* = 6 independent replica. *P* values were calculated with one-tailed one-way ANOVA and false discovery rate (FDR) multiple-testing correction using the two-stage step-up method from Benjamini, Krieger, and Yekutieli. For FACS gating strategy, see Supplementary Fig. 2a. **b**, Mass spectrometry-based proteomic data showing correlation of NMuMG cells expressing GFP or indicated *Fgfr2* variants for global protein expression, global phosphoproteomic analysis after enrichment with IMAC, and phospho-Tyr immunoprecipitation (IP)-enriched samples. Pearson's *R* correlation coefficients are depicted and heatmaps were clustered unsupervised. **c**, Heatmaps visualizing FGFR2 phosphosites identified in **(b)**. **d**, Single-sample gene set enrichment analysis (ssGSEA) based on hallmark gene sets from MSigDB [79] and the global protein expression dataset. Significant single-sample normalized enrichment scores (NES) were calculated using GSEA standard settings [77,78]. NES are visualised as colour-coded row Z-scores and depicted terms are based on *Fgfr2^{ΔE18}* versus *Fgfr2^{FL}* two-group comparisons using two-tailed unpaired Student's *t*-tests. Significant terms are shown (*P* < 0.05). **e**, Relative candidate protein expression levels corresponding to MAPK, AKT, and mTOR substrates displayed in Fig. 2b and based on the global protein expression. **f**, Single-sample phosphosite signature enrichment analysis (ssPTM-SEA) based on murine kinase/pathway definitions of PTMSigDB [82] and the global phosphoproteomic dataset. Significant single-sample NES were calculated using gene permutation (*n* = 1,000) and one-tailed permutation testing with FDR multiple-testing correction using the Benjamini-Hochberg method by applying PTM-SEA standard settings [82]. NES are visualised as colour-coded row Z-scores and depicted terms are based on *Fgfr2^{ΔE18}* versus GFP, *Fgfr2^{ΔE18}* versus *Fgfr2^{FL}*, and/or *Fgfr2^{FL}* versus GFP two-group comparisons using two-tailed unpaired Student's *t*-tests. Terms significant for either of the three two-group comparisons are shown (*P* < 0.05). **g**, Western blots showing expression and phosphorylation of indicated proteins in NMuMG cells expressing GFP or indicated *Fgfr2* variants and treated for 3 h with vehicle or 100 nM AZD4547. β-Actin was run on separate gels as sample processing control, and each blot was stained with Ponceau S to ensure equal loading of total protein. Blots stained with the same antibody were developed and recorded in parallel and subjected to equal post-imaging processing. For gel source data, see Supplementary Fig. 1a–g. **h**, Quantifications of relative phosphoprotein band intensities in **(g)** normalized to β-actin, corresponding total protein, and FGFR2 band intensities. Data in **g**, **h** represent 1 replica of 2 independent experiments.



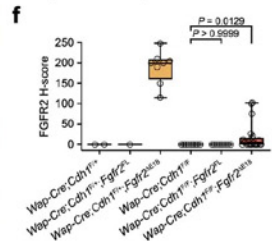
Extended Data Fig. 6: In vitro and in vivo oncogenic capacities of *Fgfr2* variants.

a, Representative images of 12-well plate wells and quantification of 3D soft agar colony formation assay using NMuMG cells expressing *GFP* or indicated *Fgfr2* variants and treated with vehicle, 100 nM AZD4547, or 100 nM pemigatinib for 15 days. Data are represented as mean $\pm$ standard deviation (s.d.) of *GFP*, *Fgfr2^{FL}*, *Fgfr2^{ΔE18}*, vehicle, *n* = 33; AZD4547, pemigatinib, *n* = 18 independent replica from 4 individual experiments. *Fgfr2^{FL}-Bicc1*, vehicle, *n* = 18; AZD4547, pemigatinib, *n* = 9 independent replica from 3 individual experiments. *Fgfr2^{ΔE18}-Bicc1*, *Fgfr2^{FL}-Ate1*, *Fgfr2^{ΔE18}-Ate1*, *Fgfr2^{FL}-Tacc2*, *Fgfr2^{ΔE18}-Tacc2*, *Fgfr2^{ΔE18}-IGR1*, *Fgfr2^{ΔE18}-IGR2*, *Fgfr2^{E18-C2}*, *Fgfr2^{E18-C3}*, *Fgfr2^{E18-C4}*, *Fgfr2^{T674*}*, *Fgfr2^{T678*}*, *Fgfr2^{L681S16}*, *Fgfr2^{P686*}*, *Fgfr2^{S694*}*, *Fgfr2^{S156W}*, *Fgfr2^{C287R}*, *Fgfr2^{N454K}*, vehicle, *n* = 12; AZD4547, pemigatinib, *n* = 6 independent replica from 2 individual experiments. *Fgfr2^{ΔE18}-Bicc1^{ΔSAM}*, *Fgfr2^{K422R}*, *Fgfr2^{K422R-ΔE18}*, *Fgfr2^{K422R}-Bicc1*, *Fgfr2^{K422R-ΔE18}-Bicc1*, *Fgfr2^{S687S13}*, *Fgfr2^{S697S14}*, *Fgfr2^{V702*}*, *Fgfr2^{S704S122}*, *Fgfr2^{V717*}*, *Fgfr2^{K564E}*, *n* = 6; AZD4547, pemigatinib, *n* = 3 independent replica from 1 experiment. **b, c**, FACS to quantify traced EGFP⁺ EpCAM⁺ epithelial cells (**b**) and their FGFR2 MFI (**c**). *Rosa26-mT/mG* female reporter mice were intraductally injected with lentiviruses encoding Cre or indicated *Fgfr2*-P2A-Cre variants resulting in Cre-mediated mT/mG allele switching, thus cell membrane-localized tdTomato (mT) expression was replaced by membrane-localized EGFP (mG) expression. Mammary glands (MGs) were subjected to FACS analysis at indicated timepoints post injection. Data are represented as mean $\pm$ s.d. and each data point represents a MG pool of an individual mouse. Analyses were done in batches of 1–2 mice of each *Fgfr2* variant and one timepoint. In (**b**), 1 week, uninjected MGs, *n* = 5; Cre, *n* = 7; other *Fgfr2* variants, *n* = 4; 3 weeks, all groups, *n* = 4; 6 weeks, uninjected MGs, Cre, *Fgfr2^{K422R}-P2A-Cre*, *Fgfr2^{K422R-ΔE18}-P2A-Cre*, *Fgfr2^{K422R}-Bicc1-P2A-Cre*, *Fgfr2^{K422R-ΔE18}-Bicc1-P2A-Cre*, *n* = 5; other *Fgfr2* variants, *n* = 6 mice. In (**c**), 1 week and 3 weeks, all groups, *n* = 3; 6 weeks, Cre, *Fgfr2^{FL}-P2A-Cre*, *Fgfr2^{ΔE18}-P2A-Cre*, *Fgfr2^{FL}-Bicc1-P2A-Cre*, *Fgfr2^{ΔE18}-Bicc1-P2A-Cre*, *Fgfr2^{ΔE18}-Bicc1^{ΔSAM}-P2A-Cre*, *n* = 4; other *Fgfr2* variants, *n* = 3 mice. For FACS gating strategy, see Supplementary Fig. 2b. **d, e**, Kaplan-Meier curves showing mammary tumour-specific survival of female wild-type (WT) mice intraductally injected with lentiviruses encoding indicated *Fgfr2* variants. *Fgfr2^{FL}*, *n* = 20; *Fgfr2^{ΔE18}*, *n* = 22; *Fgfr2^{FL}-Bicc1*, *Fgfr2^{ΔE18}-Bicc1^{ΔSAM}*, *Fgfr2^{FL}-Ate1*, *Fgfr2^{ΔE18}-Ate1*, *Fgfr2^{FL}-Tacc2*, *Fgfr2^{ΔE18}-Tacc2*, *Fgfr2^{ΔE18}-IGR1*, *Fgfr2^{ΔE18}-IGR2*, *Fgfr2^{E18-C2}*, *Fgfr2^{E18-C3}*, *Fgfr2^{E18-C4}*, *n* = 10; *Fgfr2^{ΔE18}-Bicc1*, *n* = 11 mice. *Fgfr2^{FL}* and *Fgfr2^{ΔE18}* curves in **a** are duplicated in **d, h, f, g**. **f, g**, Kaplan-Meier curves showing mammary tumour-free (**c**) and -specific (**d**) survival of female *Wap-Cre;Cdh1^{fl/fl}* mice intraductally injected with lentiviruses encoding indicated *Fgfr2* variants. *Fgfr2^{FL}*, *n* = 34 of 15; *Fgfr2^{ΔE18}*, *n* = 39 of 15; *Fgfr2^{FL}-Bicc1*, *n* = 19 of 10; *Fgfr2^{ΔE18}-Bicc1*, *n* = 21 injected MGs of 11 mice. **h**, Kaplan-Meier curves showing mammary tumour-specific survival of female wild-type (WT) mice intraductally injected with lentiviruses encoding indicated *Fgfr2* variants. *Fgfr2^{T674*}*, *Fgfr2^{L681S16}*, *Fgfr2^{S697S14}*, *Fgfr2^{S704S122}*, *n* = 10; *Fgfr2^{T678*}*, *n* = 4; *Fgfr2^{P686*}*, *Fgfr2^{S694*}*, *Fgfr2^{V702*}*, *Fgfr2^{V717*}*, *n* = 5 mice. *P* values were calculated with one-tailed two-way ANOVA and FDR multiple-testing corrections using the two-stage step-up method from Benjamini, Krieger, and Yekutieli (**a, c**), one-tailed Kruskal-Wallis tests and Dunn's multiple-testing corrections (**b**), or log rank (Mantel-Cox) tests (**b-h**). *****P* < 0.0001; NS, not significant (*P* ≥ 0.05).



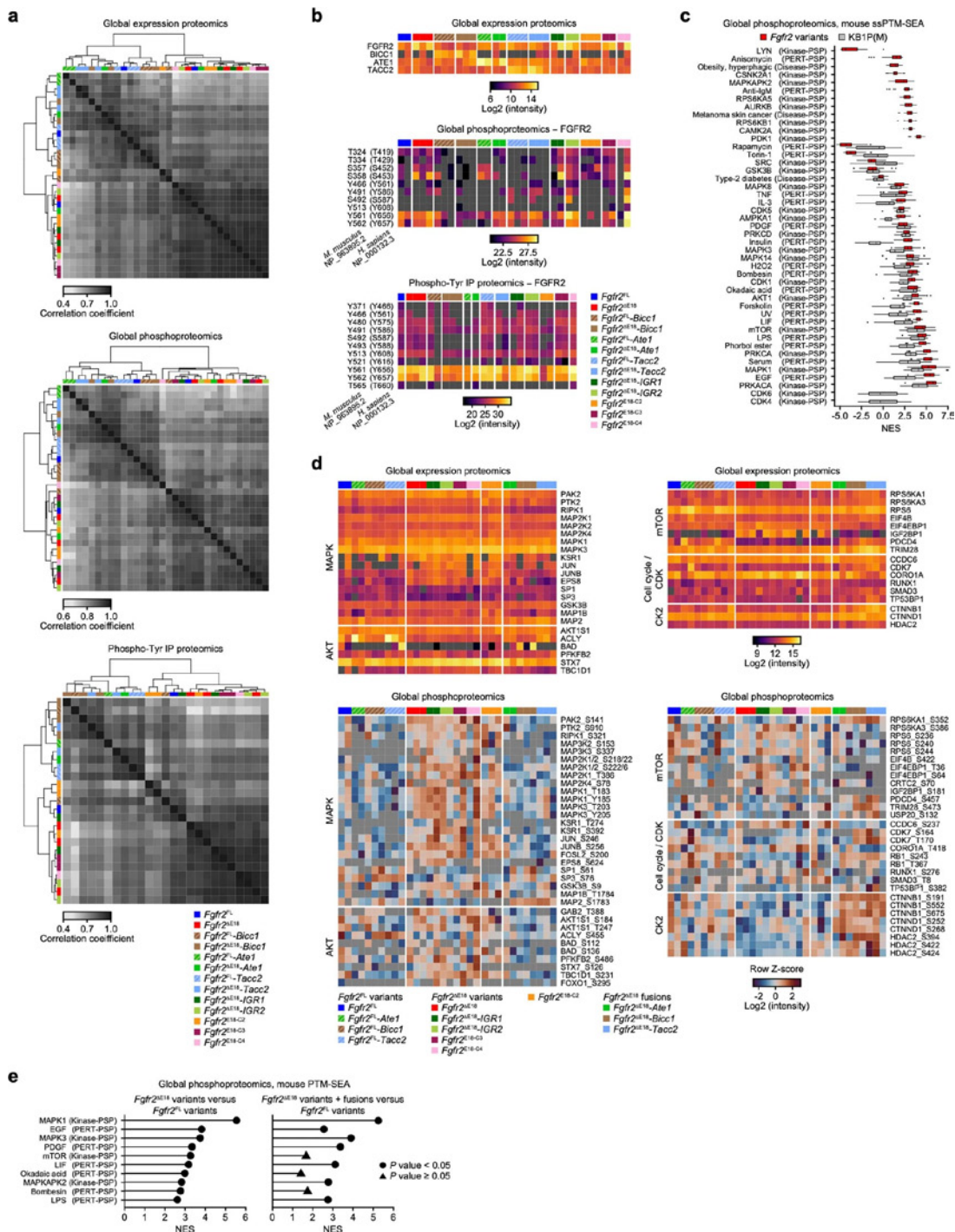
Extended Data Fig. 7: In vivo oncogenic capacities of *Fgfr2* variants in somatic models and GEMMs.

a, b, Kaplan-Meier curves showing mammary tumour-free (**a**) and -specific (**b**) survival of female *Wap-Cre;Cdh1^{fl/fl}* mice intraductally injected with lentiviruses encoding indicated *Fgfr2* variants. *Fgfr2^{16/78*}*, *n* = 13 of 7; *Fgfr2^{2686*}*, *n* = 13 of 4; *Fgfr2^{2694*}*, *n* = 12 of 4; *Fgfr2^{1702*}*, *n* = 12 of 5; *Fgfr2^{2717*}*, *n* = 11 injected MGs of 4 mice. *Fgfr2^{FL}* and *Fgfr2^{ΔE18}* curves in **a, b** are duplicates from Extended Data Fig. 6f, g. **c, d**, Kaplan-Meier curves showing mammary tumour-free (**c**) and -specific (**d**) survival of female WT mice intraductally injected with lentiviruses encoding indicated *Fgfr2* variants. *Fgfr2^{25156W}*, *Fgfr2^{C287R}*, *Fgfr2^{M454K}*, *Fgfr2^{K554E}*, *n* = 20 injected MGs of 10 mice. *Fgfr2^{FL}* and *Fgfr2^{ΔE18}* curves in **c, d** are duplicates from Fig. 2c and Extended Data Fig. 6d. **e**, Schematic representation of the engineered *Fgfr2^{FL}* and *Fgfr2^{ΔE18}* alleles. *Frt-invCAG-Fgfr2^{FL}-IRES-Luc* and *Frt-invCAG-Fgfr2^{ΔE18}-IRES-Luc* were inserted into the *Col1a1* locus using the genetically engineered mouse model – embryonic stem cell (GEMM-ESC) methodology [54]. Cre activity inverts the CAG promoter resulting in coherent FGFR2 and luciferase (Luc) expression. IRES, internal ribosome entry site. **f**, RT-qPCR quantification of *Fgfr2* and *Luc* expression in mouse mammary epithelial cells (MMECs) isolated from pooled MGs of 10-week-old WT control, *Fgfr2^{FL}-IRES-Luc*, and *Fgfr2^{ΔE18}-IRES-Luc* female mice and mock-treated or treated with adenoviral Ad5CMVCre (AdCre) to switch *Fgfr2* alleles in vitro. Data are represented as mean ± s.d. of WT, *n* = 1; *Fgfr2^{FL}-IRES-Luc*, *Fgfr2^{ΔE18}-IRES-Luc*, *n* = 4 MMEC cultures each from MG pools of individual mice. **g, h**, Western blot showing FGFR2 expression of mock- or AdCre-treated MMEC cultures (**g**) and quantification of relative FGFR2 intensities normalized to β-actin (**h**). β-Actin was run on a separate gel as sample processing control, and membranes were stained with Ponceau S to ensure equal loading of total protein. For gel source data, see Supplementary Fig. 1h, i. In **h**, data are represented as mean ± s.d. of WT, *n* = 1; *Fgfr2^{FL}-IRES-Luc*, *Fgfr2^{ΔE18}-IRES-Luc*, *n* = 3 MMEC cultures each from MG pools of individual mice. **i**, Luciferase activity measured using luciferin and bioluminescence imaging on mock- or AdCre-treated MMEC cultures. Data are represented as simple linear regressions across *Fgfr2^{FL}-IRES-Luc*, *n* = 4; *Fgfr2^{ΔE18}-IRES-Luc*, *n* = 3 MMEC cultures (each from MG pools of individual mice) at indicated cell densities. **j**, Representative in vivo bioluminescence images showing luciferase activity following luciferin administration measured as photon flux in 10-week-old *Wap-Cre;Cdh1^{fl/fl}*, *Wap-Cre;Cdh1^{fl/fl};Fgfr2^{FL}-IRES-Luc*, and *Wap-Cre;Cdh1^{fl/fl};Fgfr2^{ΔE18}-IRES-Luc* female mice. Scale bars, 1 cm. **k**, Quantification of luciferase activity using recurrent bioluminescence imaging in indicated GEMMs. *Wap-Cre;Cdh1^{fl/fl}* female mice show background luminescence. *Wap-Cre;Cdh1^{fl/fl}*, *n* = 3; *Wap-Cre;Cdh1^{fl/fl};Fgfr2^{FL}-IRES-Luc*, *n* = 6; *Wap-Cre;Cdh1^{fl/fl};Fgfr2^{ΔE18}-IRES-Luc*, *n* = 4 mice. **l, m**, Kaplan-Meier curves showing mammary tumour-free (**l**) and -specific (**m**) survival of indicated GEMMs. *Wap-Cre;Cdh1^{fl/fl}*, *n* = 12; *Wap-Cre;Cdh1^{fl/fl};Fgfr2^{FL}-IRES-Luc*, *n* = 5; *Wap-Cre;Cdh1^{fl/fl};Fgfr2^{ΔE18}-IRES-Luc*, *n* = 6; *Wap-Cre;Cdh1^{fl/fl}*, *n* = 16; *Wap-Cre;Cdh1^{fl/fl};Fgfr2^{FL}-IRES-Luc*, *Wap-Cre;Cdh1^{fl/fl};Fgfr2^{ΔE18}-IRES-Luc*, *n* = 19 mice. *P* values were calculated with log rank (Mantel-Cox) tests (**a–d, l, m**), one-tailed two-way ANOVA and FDR multiple-testing corrections using the two-stage step-up method from Benjamini, Krieger, and Yekutieli (**f**), a two-tailed unpaired Student's *t*-test (**h**), or one-way analysis of covariance (ANCOVA) to compare linear regression slopes (**i**). *****P* < 0.0001.



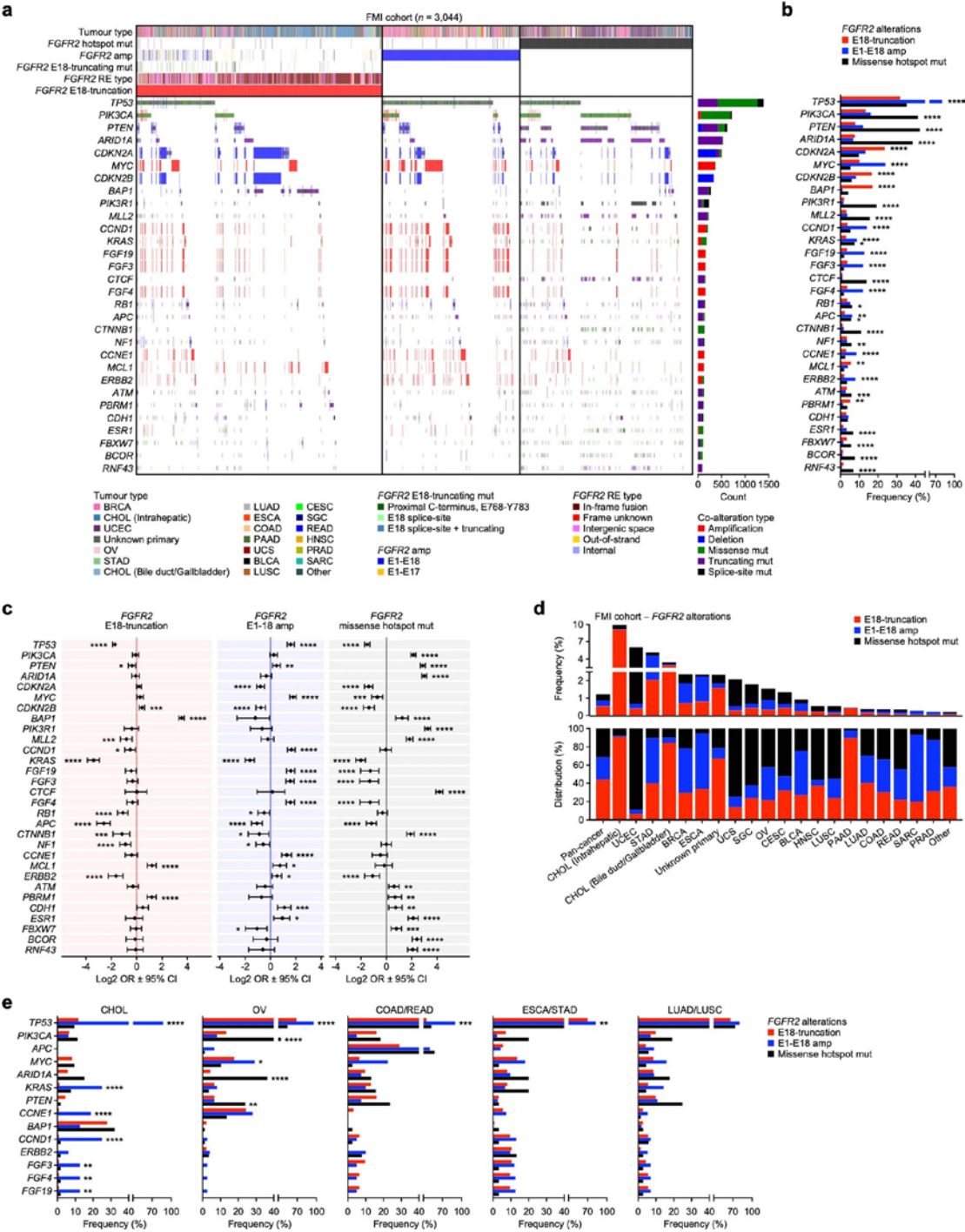
Extended Data Fig. 8: Mammary tumour types observed in *Fgfr2* mouse models.

a, Representative hematoxylin and eosin (H&E) histochemistry and FGFR2 and E-cadherin immunohistochemistry (IHC) stains on mammary tissue sections from indicated *Fgfr2* somatic mouse models. Per MG one tissue section was stained and quantified for each of the indicated stains acquired in multiple independent randomized batches across all *Fgfr2* variants. Numbers of stained and quantified MGs are in **(b, c)**. ILC, invasive lobular carcinoma. **b**, Mammary tumour type classifications of *Fgfr2* somatic mouse models based on H&Es and E-cadherin IHC stains. WT, *Fgfr2*^{FL}, *n* = 46 of 20; *Fgfr2*^{ΔE18}, *n* = 45 of 22; *Fgfr2*^{FL}-*Biccl1*, *n* = 23 of 10; *Fgfr2*^{ΔE18}-*Biccl1*, *n* = 26 of 11; *Fgfr2*^{ΔE18}-*Biccl1*^{ΔSAM}, *Fgfr2*^{ΔE18}-*Ate1*, *Fgfr2*^{FL}-*Tacc2*, *Fgfr2*^{E18-C3}, *n* = 20 of 10; *Fgfr2*^{FL}-*Ate1*, *Fgfr2*^{ΔE18}-*Tacc2*, *Fgfr2*^{E18-C4}, *n* = 19 of 10; *Fgfr2*^{ΔE18}-*IGR1*, *n* = 18 of 9; *Fgfr2*^{FL}-*IGR2*, *Fgfr2*^{E18-C2}, *n* = 17 of 10; *Fgfr2*^{1674*}, *Fgfr2*^{C287R}, *n* = 16 of 8; *Fgfr2*^{1678*}, *n* = 15 of 4; *Fgfr2*^{ΔE18}-*IGR1*, *n* = 14 of 7; *Fgfr2*^{2686*}, *n* = 16 of 5; *Fgfr2*^{5694*}, *n* = 13 of 5; *Fgfr2*^{5697/5*4}, *n* = 8 of 4; *Fgfr2*^{1702*}, *n* = 8 of 3; *Fgfr2*^{1717*}, *n* = 14 of 5; *Fgfr2*^{N454K}, *n* = 4 injected MGs of 2 mice. *Wap-Cre*; *Cdh1*^{FL/FL}, *Fgfr2*^{FL}, *n* = 34 of 15; *Fgfr2*^{ΔE18}, *n* = 39 of 15; *Fgfr2*^{FL}-*Biccl1*, *n* = 17 of 9; *Fgfr2*^{ΔE18}-*Biccl1*, *n* = 21 of 11; *Fgfr2*^{1678*}, *n* = 13 of 7; *Fgfr2*^{2686*}, *n* = 14 of 4; *Fgfr2*^{5694*}, *n* = 12 of 4; *Fgfr2*^{1702*}, *n* = 12 of 5; *Fgfr2*^{1717*}, *n* = 11 injected MGs of 4 mice. **c**, Histo-scoring of FGFR2 IHC stains on mammary tumours from *Fgfr2* somatic mouse models. WT, *Fgfr2*^{FL}, *n* = 6 of 5; *Fgfr2*^{ΔE18}, *n* = 39 of 22; *Fgfr2*^{FL}-*Biccl1*, *n* = 17 of 10; *Fgfr2*^{ΔE18}-*Biccl1*, *n* = 22 of 11; *Fgfr2*^{ΔE18}-*Biccl1*^{ΔSAM}, *n* = 17 of 9; *Fgfr2*^{FL}-*Ate1*, *n* = 5 of 5; *Fgfr2*^{ΔE18}-*Ate1*, *Fgfr2*^{E18-C4}, *n* = 16 of 9; *Fgfr2*^{FL}-*Tacc2*, *n* = 12 of 8; *Fgfr2*^{ΔE18}-*Tacc2*, *n* = 11 of 8; *Fgfr2*^{ΔE18}-*IGR1*, *Fgfr2*^{E18-C3}, *n* = 14 of 9; *Fgfr2*^{ΔE18}-*IGR2*, *n* = 15 of 10; *Fgfr2*^{E18-C2}, *n* = 12 of 9; *Fgfr2*^{1674*}, *n* = 14 of 8; *Fgfr2*^{1678*}, *n* = 11 of 4; *Fgfr2*^{ΔE18}-*IGR1*, *n* = 13 of 8; *Fgfr2*^{2686*}, *Fgfr2*^{5694*}, *n* = 10 of 5; *Fgfr2*^{5697/5*4}, *n* = 7 of 4; *Fgfr2*^{1702*}, *n* = 6 of 3; *Fgfr2*^{1717*}, *n* = 10 of 4; *Fgfr2*^{C287R}, *n* = 15 of 8; *Fgfr2*^{N454K}, *n* = 3 tumours of 2 mice. *Wap-Cre*; *Cdh1*^{FL/FL}, *Fgfr2*^{FL}, *n* = 12 of 8; *Fgfr2*^{ΔE18}, *n* = 23 of 9; *Fgfr2*^{FL}-*Biccl1*, *n* = 14 of 7; *Fgfr2*^{ΔE18}-*Biccl1*, *n* = 18 of 10; *Fgfr2*^{1678*}, *Fgfr2*^{1702*}, *n* = 10 of 5; *Fgfr2*^{2686*}, *n* = 9 of 4; *Fgfr2*^{5694*}, *n* = 10 of 4; *Fgfr2*^{1717*}, *n* = 6 tumours of 3 mice. **d**, Representative H&E histochemistry and E-cadherin and FGFR2 IHC stains on mammary tissue sections from indicated GEMMs. Per MG one tissue section was stained and quantified for each of the indicated stains acquired in two independent randomized batches across all genotypes. Numbers of stained and quantified MGs are in **(e, f)**. **e**, Mammary tumour type classifications of GEMMs based on H&Es and E-cadherin IHC stains. *Wap-Cre*; *Cdh1*^{FL/FL}, *n* = 45 of 12; *Wap-Cre*; *Cdh1*^{FL/FL}; *Fgfr2*^{FL}, *n* = 20 of 5; *Wap-Cre*; *Cdh1*^{FL/FL}; *Fgfr2*^{ΔE18}, *n* = 29 of 6; *Wap-Cre*; *Cdh1*^{FL/FL}, *n* = 60 of 16; *Wap-Cre*; *Cdh1*^{FL/FL}; *Fgfr2*^{FL}, *n* = 73 of 19; *Wap-Cre*; *Cdh1*^{FL/FL}; *Fgfr2*^{ΔE18}, *n* = 79 MGs of 19 mice. **f**, Histo (H)-score quantifications of FGFR2 IHC stains on mammary tumours from GEMMs. *Wap-Cre*; *Cdh1*^{FL/FL}, *n* = 2 of 2; *Wap-Cre*; *Cdh1*^{FL/FL}; *Fgfr2*^{FL}, *n* = 1 of 1; *Wap-Cre*; *Cdh1*^{FL/FL}; *Fgfr2*^{ΔE18}, *n* = 8 of 5; *Wap-Cre*; *Cdh1*^{FL/FL}, *n* = 9 of 8; *Wap-Cre*; *Cdh1*^{FL/FL}; *Fgfr2*^{FL}, *n* = 8 of 6; *Wap-Cre*; *Cdh1*^{FL/FL}; *Fgfr2*^{ΔE18}, *n* = 24 tumours of 19 mice. Data are represented as median (centre line) ± IQR (25th to 75th percentile, box) and ± full range (minimum to maximum, whiskers) and *P* values were calculated with one-tailed Kruskal-Wallis tests and Dunn's multiple-testing corrections. Scale bars, overview, 500 μm; inset, 50 μm.



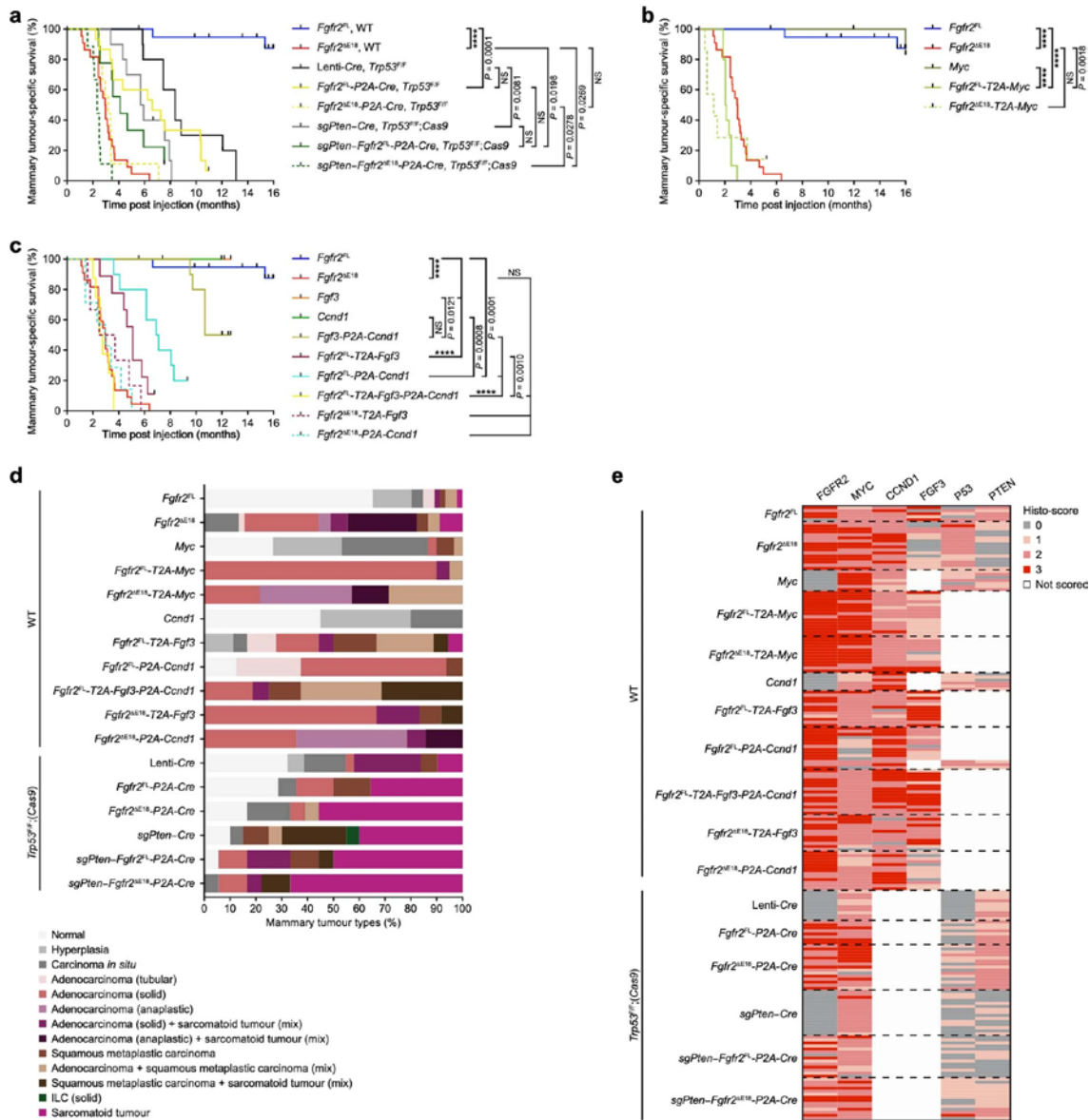
Extended Data Fig. 9: (Phospho)-proteomic analyses of tumours from *Fgfr2* somatic mouse models.

a, Mass spectrometry-based proteomic data showing correlation of indicated *Fgfr2* somatic mouse models for global protein expression, global phosphoproteomic analysis after enrichment with IMAC, and phospho-Tyr IP-enriched mammary tumours. Pearson's *R* correlation coefficients are depicted and heatmaps were clustered unsupervised. **b**, Relative protein expression of FGFR2 and its fusion partners BICC1, ATE1, and TACC2 next to FGFR2 phosphosites identified in datasets from **a**. Heatmaps colour-code relative intensities of protein expression and phosphorylation. **c**, ssPTM-SEA based on murine kinase/pathway definitions of PTMsigDB and the global phosphoproteomic dataset in **a** as well as phosphoproteomic data generated from mammary tumours from *K14-Cre;Brca^{fl/fl};Trp53^{fl/fl}* (KB1P) and *KB1P;Mdr1a/b^{-/-}* (KB1PM) GEMMs. Each boxplot represents one ssPTM-SEA term and shows NES of individual *Fgfr2* variant tumours or KB1P(M) tumours. ssPTM-SEA terms enriched in the *Fgfr2* variant and/or the KB1P(M) tumour cohorts are shown. Significant single-sample NES were calculated using gene permutation ($n=1,000$) and permutation-derived *P* values by applying PTM-SEA standard settings [82]. No further statistical selections were applied. Boxplots are represented as median (centre line) $\pm$ IQR (25th to 75th percentile, box) and IQR $\pm$ 1.5 $\times$ IQR (whiskers). *Fgfr2* variants, $n=32$; KB1P, $n=14$; KB1PM, $n=10$ tumours. **d**, Relative candidate protein expression (top panels) and phosphorylation (bottom panels) levels of MAPK, AKT, mTOR, cell cycle/ CDK, and CK2 substrates identified in **a**. For the phosphoproteomic analysis, samples were grouped into *Fgfr2^{FL}* variants, *Fgfr2^{ΔE18}* variants, and *Fgfr2^{ΔE18}* fusion variants and compared pairwise using the robust kinase activity inference (RoKAI) tool at default settings [83] including two-tailed hypothesis testing on Z-scores and FDR multiple-testing correction using the Benjamini-Hochberg method. Group comparison fold change (FC) values of $-1.5 \geq FC \geq 1.5$ and $P < 0.05$ were considered. The RoKAI output was used to manually curate phosphosites of interest, and phosphosites were manually grouped into indicated signalling pathways guided by RoKAI. The heatmaps depict relative expression intensities (top panels) and Z-scores of phosphosite intensities calculated per row from log₂-transformed intensity values (bottom panels). **e**, PTM-SEA based on murine kinase/pathway definitions of PTMsigDB and performed with global phosphoproteomic data and limma-based two-group comparisons of *Fgfr2^{ΔE18}* variants versus *Fgfr2^{FL}* variants groups (left panel) and *Fgfr2^{ΔE18}* variants including fusions versus *Fgfr2^{FL}* variants groups (right panel). Significant NES were calculated by using gene permutation ($n=1,000$) and one-tailed permutation testing without multiple-testing correction by applying PTM-SEA standard settings [82]. Lollipop plots show NES of terms significantly enriched in either of the two comparisons ($P < 0.05$).



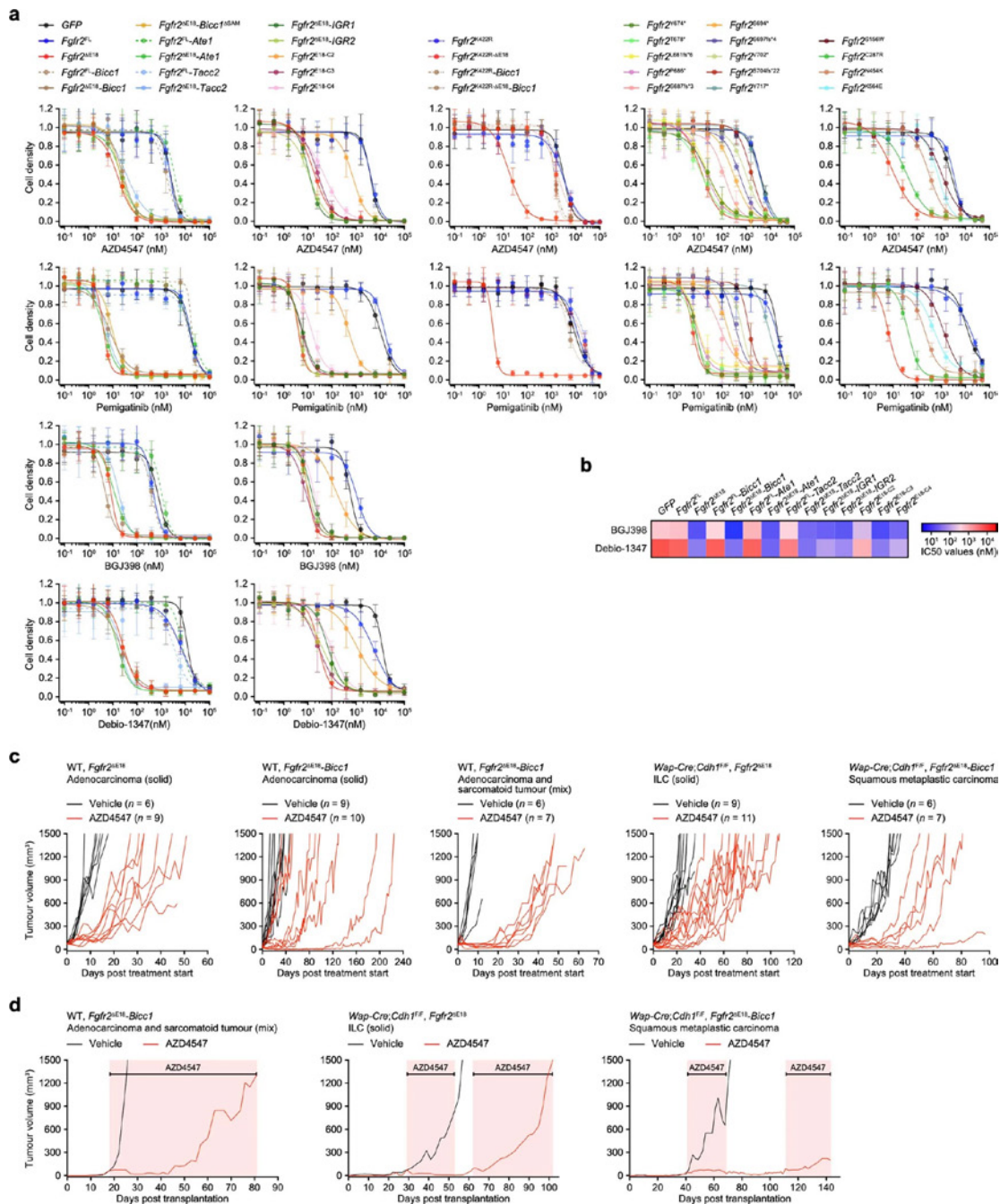
Extended Data Fig. 10: Top driver genes co-occurring in samples with *FGFR2* alterations.

a, 3,044 samples classified as either *FGFR2*-E18-truncated ($n = 1,344$, 44.2% of total, 0.54% incidence), amplified ($n = 757$, 24.8% of total, 0.30% incidence), or missense hotspot mutant ($n = 943$, 31.0% of total, 0.38% incidence) and top-30 co-enriched tumour driver alterations found in the FMI pan-cancer cohort ($n = 249,570$). **b**, Enrichments of top-30 tumour driver co-alterations in the indicated *FGFR2* alteration categories in the FMI pan-cancer cohort. **c**, Odds ratios (OR) of top-30 tumour driver co-alterations in the indicated *FGFR2* alteration categories (E18-truncation, $n = 1,344$; E1-E18 amp, $n = 757$; missense hotspot mut, $n = 943$) versus *FGFR2* WT samples ($n = 224,711$) of the FMI pan-cancer cohort. Data are represented as $\log_2$ -transformed OR $\pm$ 95% confidence interval (CI). Co-occurrence, OR > 1 ; mutual exclusivity, OR < 1 . **d**, Frequencies (top panel) and distributions (bottom panel) per tumour type of the indicated *FGFR2* alteration categories in the FMI pan-cancer cohort. **e**, Enrichment of top tumour driver co-alterations in the indicated *FGFR2* alteration categories in the FMI-CHOL, OV, COAD/READ, ESCA/STAD, and LUAD/LUSC cohorts. *P* values were calculated with one-tailed proportion z-tests (**b**, **e**) or two-tailed Fisher's exact tests (**c**) and FDR multiple-testing corrections using the Benjamini-Hochberg method (**b**, **c**, **e**). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$. Sample sizes and statistical details for **b**, **c**, **e** are in Supplementary Table 2.



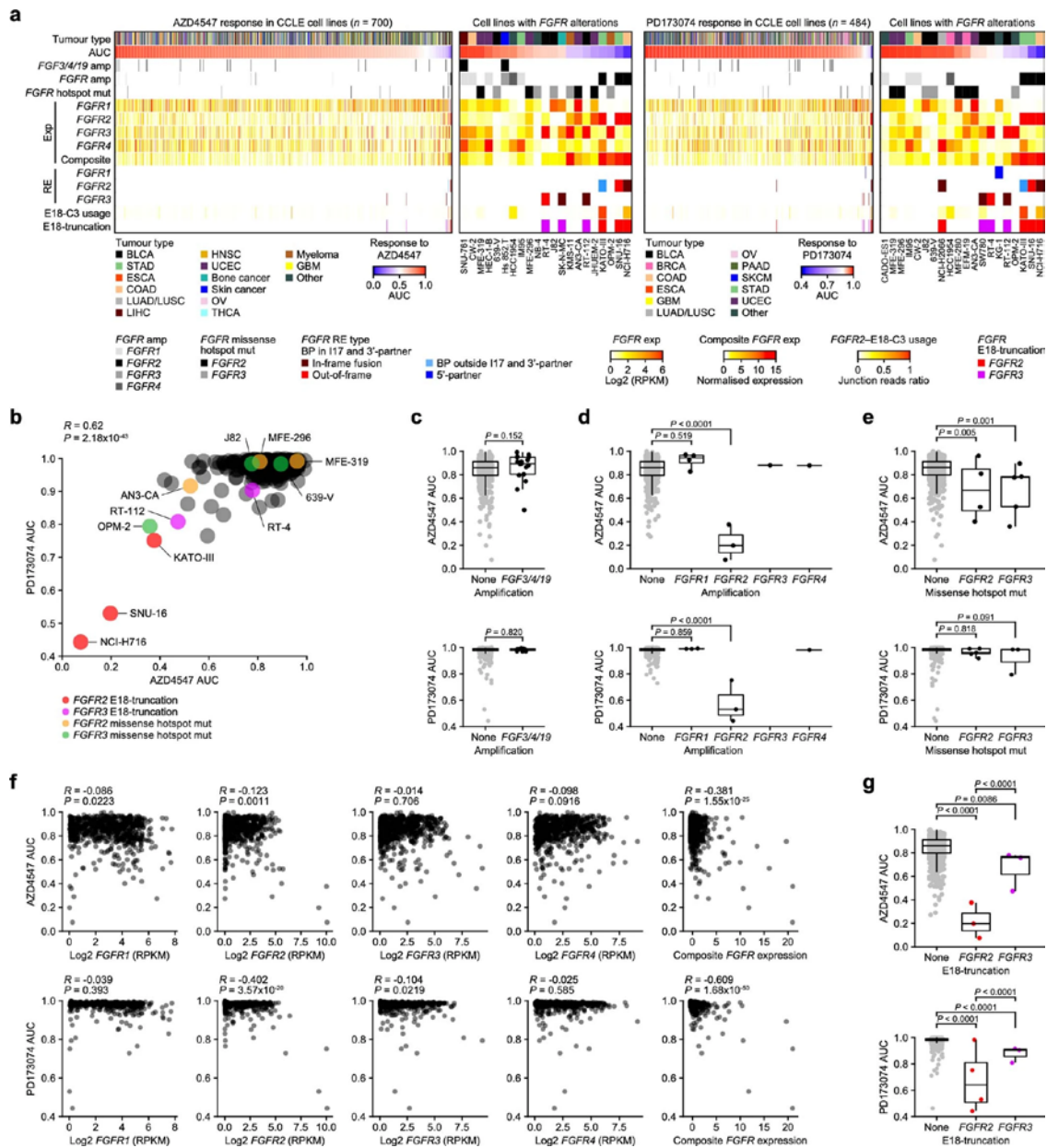
Extended Data Fig. 11: Somatic modelling of *Fgfr2* variants and co-occurring driver alterations.

a-c, Kaplan-Meier curves showing mammary tumour-specific survival of *Trp53^{fl/fl}* and *Trp53^{fl/fl};Rosa26-Cas9* (**a**) and WT (**b, c**) female mice intraductally injected with lentiviruses encoding indicated variants. *Trp53^{fl/fl};Rosa26-Cas9*, Lenti-Cre, *sgPten-Cre*, *n* = 10; *Fgfr2^{FL}-P2A-Cre*, *n* = 15; *Fgfr2^{ΔE18}-P2A-Cre*, *sgPten-Fgfr2^{FL}-P2A-Cre*, *sgPten-Fgfr2^{ΔE18}-P2A-Cre*, *n* = 9 mice. WT, *Myc*, *Fgfr2^{FL}-T2A-Fgf3*, *n* = 9; *Fgfr2^{FL}-T2A-Myc*, *Fgf3*, *Ccnd1*, *Fgf3-P2A-Ccnd1*, *Fgfr2^{FL}-P2A-Ccnd1*, *n* = 10; *Fgfr2^{ΔE18}-T2A-Myc*, *Fgfr2^{ΔE18}-P2A-Ccnd1*, *n* = 7; *Fgfr2^{FL}-T2A-Fgf3-P2A-Ccnd1*, *n* = 8; *Fgfr2^{ΔE18}-T2A-Fgf3*, *n* = 6 mice. *Fgfr2^{FL}* and *Fgfr2^{ΔE18}* curves are duplicates from Extended Data Fig. 6d. *P* values were calculated with log rank (Mantel-Cox) tests. *****P* < 0.0001. **d**, Mammary tumour type classifications of somatic mouse models based on H&Es. WT, *Myc*, *n* = 30 of 9; *Fgfr2^{FL}-T2A-Myc*, *n* = 20 of 10; *Fgfr2^{ΔE18}-T2A-Myc*, *Fgfr2^{ΔE18}-P2A-Ccnd1*, *n* = 14 of 7; *Ccnd1*, *n* = 40 of 10; *Fgfr2^{FL}-T2A-Fgf3*, *n* = 18 of 9; *Fgfr2^{FL}-P2A-Ccnd1*, *Fgfr2^{FL}-T2A-Fgf3-P2A-Ccnd1*, *n* = 16 of 8; *Fgfr2^{ΔE18}-T2A-Fgf3*, *n* = 12 injected MGs of 6 mice. *Trp53^{fl/fl};Rosa26-Cas9*, Lenti-Cre, *n* = 31 of 10; *Fgfr2^{FL}-P2A-Cre*, *n* = 14 of 7; *Fgfr2^{ΔE18}-P2A-Cre*, *sgPten-Fgfr2^{FL}-P2A-Cre*, *sgPten-Fgfr2^{ΔE18}-P2A-Cre*, *n* = 18 of 9; *sgPten-Cre*, *n* = 20 injected MGs of 10 mice. WT *Fgfr2^{FL}* and *Fgfr2^{ΔE18}* classifications are duplicates from Extended Data Fig. 8b. **e**, Histo-scoring of indicated IHC stains on mammary tumours from somatic mouse models. WT, *Fgfr2^{FL}*, *n* = 5 of 5; *Fgfr2^{ΔE18}*, *n* = 16 of 11; *Myc*, *n* = 7 of 3; *Fgfr2^{FL}-T2A-Myc*, *n* = 15 of 10; *Fgfr2^{ΔE18}-T2A-Myc*, *n* = 12 of 6; *Ccnd1*, *n* = 6 of 2; *Fgfr2^{FL}-T2A-Fgf3*, *n* = 12 of 8; *Fgfr2^{FL}-P2A-Ccnd1*, *n* = 14 of 9; *Fgfr2^{FL}-T2A-Fgf3-P2A-Ccnd1*, *n* = 15 of 8; *Fgfr2^{ΔE18}-T2A-Fgf3*, *n* = 12 of 7; *Fgfr2^{ΔE18}-P2A-Ccnd1*, *n* = 13 tumours of 7 mice. *Trp53^{fl/fl};Rosa26-Cas9*, Lenti-Cre, *n* = 10 of 8; *Fgfr2^{FL}-P2A-Cre*, *n* = 8 of 5; *Fgfr2^{ΔE18}-P2A-Cre*, *n* = 15 of 10; *sgPten-Cre*, *n* = 15 of 9; *sgPten-Fgfr2^{FL}-P2A-Cre*, *n* = 14 of 7; *sgPten-Fgfr2^{ΔE18}-P2A-Cre*, *n* = 14 tumours of 9 mice. In **d, e**, one tissue section per MG was stained and quantified for each of the indicated stains acquired in 4 independent randomized batches across all *Fgfr2* variants and genotypes.



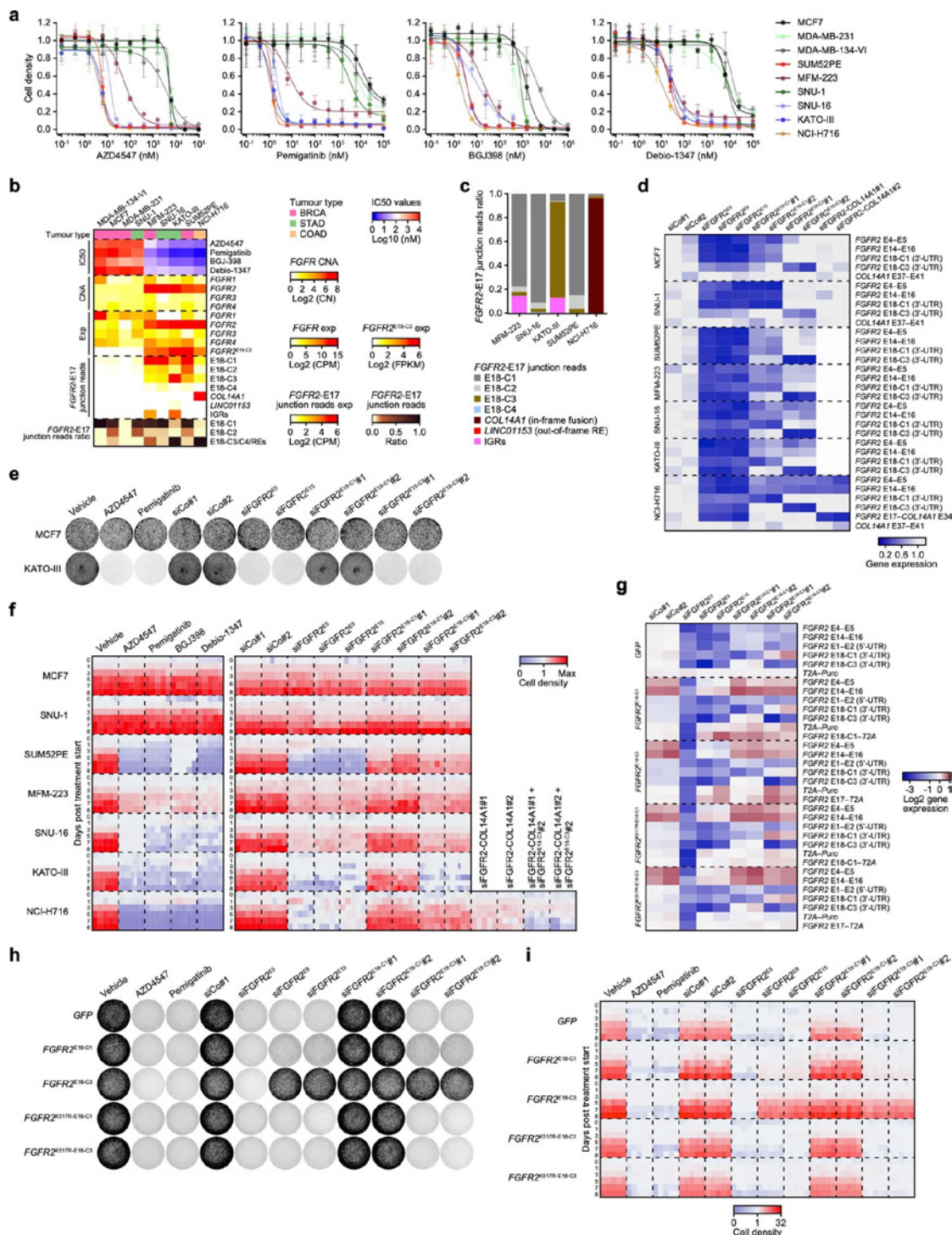
Extended Data Fig. 12: In vitro and in vivo sensitivity of *Fgfr2* variants to FGFRi.

a, Dose-response curves of 2D-grown NMuMG cells expressing *GFP* or indicated *Fgfr2* variants and treated with AZD4547, pemigatinib, BGJ398, or debio-1347 for 4 days. Data are represented as mean $\pm$ s.d. of $n=5$ replica per group collected across 5 independent experiments. **b**, Half-maximum inhibitory concentration (IC_{50}) value quantifications of BGJ398 and debio-1347 dose-response curves in **a**. Data are represented as mean of 3 independent experiments (*GFP*, *Fgfr2^{GL}*, *Fgfr2^{ΔE18}*) or 1 experiment (other *Fgfr2* variants). IC_{50} values for AZD4547 and pemigatinib are displayed in Fig. 4a. **c**, Individual growth curves of indicated tumour donors transplanted into the mammary fat pad of female syngeneic WT mice and treated daily orally with vehicle or 12.5 mg/kg AZD4547 using a previously established [55] intermittent dosing regimen. **d**, Selected tumour transplant growth curves of mice in **c**. Durations of AZD4547 treatments according to intermittent dosing regimen are indicated.



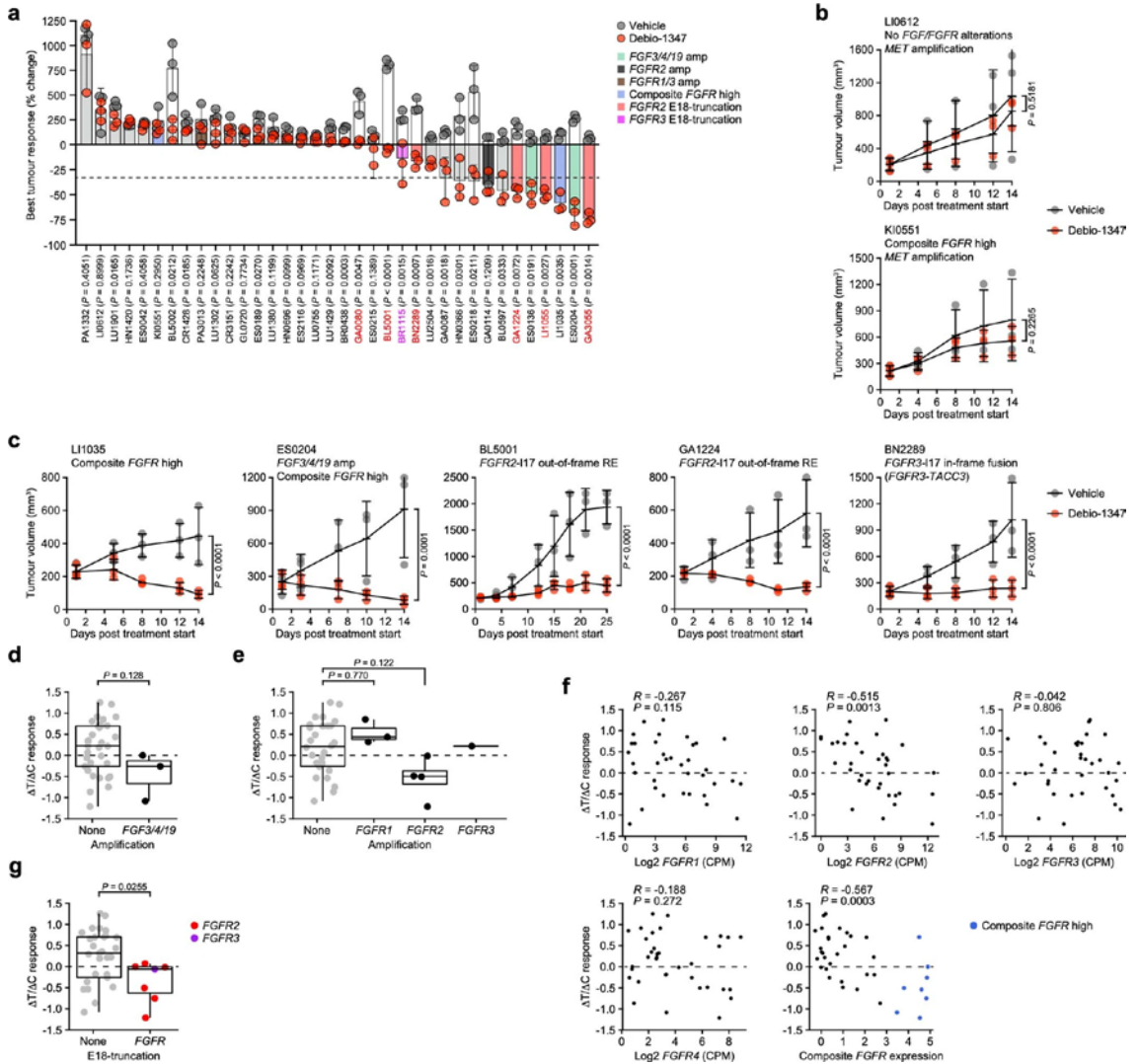
Extended Data Fig. 13: Correlation of *FGFR2* alterations to FGFRi sensitivity across CCLE.

a. The Broad Institute Cancer Cell Line Encyclopedia (CCLE) cell lines rank-ordered according to their AZD4547 or PD173074 FGFRi area under the sigmoid-fit concentration-response curve (AUC) values derived from the Cancer Therapeutics Response Portal (CTRP) v2 deposited in the PharmacDB database ($n=700$) and the Genomics of Drug Sensitivity in Cancer (GDSC) database ($n=484$), respectively. Data on *FGF/FGFR* mutations, CN status, REs, and expression and *FGFR2*-E18-C3 usage was obtained from CCLE. RPKM, reads per kilobase of transcript per million mapped reads. **b.** Correlation of AZD4547 versus PD173074 AUC values across shared CCLE cell lines ($n=384$). **c.** AZD4547 and PD173074 AUC values in CCLE cell lines with *FGF3/4/19* amp (AZD4547, $n=17$; PD173074, $n=16$) versus unaltered cell lines (AZD4547, $n=658$; PD173074, $n=455$). **d.** AZD4547 and PD173074 AUC values in CCLE cell lines with *FGFR1* amp (AZD4547, $n=4$; PD173074, $n=3$), *FGFR2* amp (AZD4547, $n=3$; PD173074, $n=3$), *FGFR3* amp (AZD4547, $n=1$; PD173074, $n=0$), or *FGFR4* amp (AZD4547, $n=1$; PD173074, $n=1$) versus unaltered cell lines (AZD4547, $n=666$; PD173074, $n=464$). **e.** AZD4547 and PD173074 AUC values in CCLE cell lines with *FGFR2* missense hotspot mut (AZD4547, $n=4$; PD173074, $n=5$) or *FGFR3* missense hotspot mut (AZD4547, $n=5$; PD173074, $n=3$) versus unaltered cell lines (AZD4547, $n=691$; PD173074, $n=476$). **f.** Correlations of *FGFR1*, *FGFR2*, *FGFR3*, *FGFR4*, or composite *FGFR* expression versus AZD4547 or PD173074 AUC values across CCLE cell lines. **g.** AZD4547 or PD173074 AUC values in CCLE cell lines expressing E18-truncated *FGFR2* (AZD4547, $n=3$; PD173074, $n=4$) or *FGFR3* (AZD4547, $n=3$; PD173074, $n=3$) versus no truncation (AZD4547, $n=694$; PD173074, $n=477$). Data in **c-e**, **g** are represented as median (centre line) $\pm$ IQR (25th to 75th percentile, box) and IQR $\pm$ 1.5 $\times$ IQR (whiskers). *P* values were calculated with two-tailed *t*-transformations of Pearson's *R* correlation coefficients (**b**, **f**), two-tailed Wilcoxon rank-sum tests (**c**), or one-tailed one-way ANOVA and Tukey's multiple-testing corrections (**d**, **e**, **g**).



Extended Data Fig. 14 Human cancer cell lines depend on *FGFR2*^{ΔE18} variants.

a, Dose-response curves of indicated human cancer cell lines treated with AZD4547, pemigatinib, BGJ398, or debio-1347 for 4 days. Data are represented as mean $\pm$ s.d. of $n=5$ replica per group collected across 2 independent experiments. **b**, Human cancer cell lines rank-ordered according to IC_{50} values for indicated FGFRi in **(a)**. *FGFR1-4* CNA and expression is based on low-coverage WGS and RNA-seq profiles. *FGFR2*^{E18-C3} isoform expression, *FGFR2*-E17 junction reads expression, and *FGFR2* RE types were identified in RNA-seq profiles. FPKM, fragments per kilobase of transcript per million mapped reads. **c**, Distribution of *FGFR2*-E17 junction reads to canonical full-length E18-C1 versus noncanonical E18-C2/C3/C4, RE partners, and IGRs in indicated human cancer cell lines. **d**, Heatmap showing silencing of *FGFR2* variants in indicated human cancer cell lines using small interfering (si) RNAs. Cells were transfected with the following siRNAs: non-targeting siRNAs (siCo), siRNAs targeting shared exons among *FGFR2* isoforms (siFGFR2^{E5,E9,E15}), siRNAs specifically targeting canonical E18-C1 of *FGFR2*^{FL} (siFGFR2^{E18-C1}) or E18-C3 of truncated *FGFR2*^{E18-C3} (siFGFR2^{E18-C3}), or siRNAs specifically targeting the *FGFR2*-COL14A1 fusion (siFGFR2-COL14A1). Silencing of specific *FGFR2* variants was detected with RT-qPCR using primers spanning indicated cDNA segments. Expression of each cDNA segment is normalized to *USF1* expression and cDNA segment expression in siCo condition of each cell line (average of siCo#1 and siCo#2). Data are represented as mean of $n=3$ technical replica per group. Data represent 1 replica of 2 independent experiments. **e, f**, Representative images of 6-well plate wells at 8 days post treatment start **(e)** and cell density quantifications over 8 days **(f)** of 2D-grown indicated cell lines treated with vehicle or 100 nM AZD4547, pemigatinib, BGJ398, or debio-1347 or (co)-transfected with siCo, siFGFR2^{E5}, siFGFR2^{E9}, siFGFR2^{E15}, siFGFR2^{E18-C1}, siFGFR2^{E18-C3}, and/or siFGFR2-COL14A1. Data in **f** represent $n=6$ independent replica collected across 1 experiment (MCF7, siRNA treatments), $n=10$ independent replica collected across 2 independent experiments (KATO-III vehicle, AZD4547, and siRNA treatments), or $n=5$ independent replica collected across 2 independent experiments (other cell lines and/or treatment conditions). **g**, Heatmap showing silencing of *FGFR2* isoforms using indicated siRNAs in KATO-III cells expressing *GFP* or indicated *FGFR2* variants. Validation of overexpression of *FGFR2* variants using RT-qPCR is in Supplementary Table 4. *FGFR2*^{K517R} variants encode KD-dead *FGFR2* variants. siFGFR2^{E5} targets endogenous *FGFR2* transcripts and *FGFR2* transcripts derived from lentiviral constructs. Other siRNAs specifically target endogenous *FGFR2* transcripts. Silencing of specific *FGFR2* variants was detected with RT-qPCR using primers spanning indicated cDNA segments. E4–E5 and E14–E16 primers detect endogenous *FGFR2* transcripts and *FGFR2* transcripts derived from lentiviral constructs. E1–E2 (5'-UTR), E18-C1 (3'-UTR), and E18-C3 (3'-UTR) primers specifically detect endogenous *FGFR2* transcripts. E18-C1–T2A, E17–T2A, and T2A–Puro primers specifically detect *FGFR2* transcripts derived from lentiviral constructs. Log₂-transformed expression of each cDNA segment is normalized to *USF1* expression and cDNA segment expression in siCo-treated *GFP*-expressing cells (average of siCo#1 and siCo#2). Data are represented as mean of $n=3$ technical replica per group of 1 experiment. **h, i**, Representative images of 6-well plate wells at 8 days post treatment start **(h)** and cell density quantifications over 8 days **(i)** of 2D-grown KATO-III cells expressing *GFP* or indicated *FGFR2* variants and treated with vehicle, 100 nM AZD4547, or 100 nM pemigatinib or transfected with siCo, siFGFR2^{E5}, siFGFR2^{E9}, siFGFR2^{E15}, siFGFR2^{E18-C1}, or siFGFR2^{E18-C3}. Data in **i** represent $n=6$ independent replica per group collected across 1 experiment.



Extended Data Fig. 15 PDXs expressing *FGFR2*^{ΔE18} variants are sensitive to debio-1347.

a, Best percentage change from baseline tumour volume in patient-derived xenograft (PDX) models ($n=36$) engrafted in female NOD-SCID (BL5001, BL5002) or BALB/c Nude (all other PDX models) mice and treated daily orally with vehicle or debio-1347 ($n=3$ mice per PDX model and treatment group). Coloured bars indicate identified *FGF/FGFR2* alterations detailed in Fig. 4c. **b, c**, Growth curves of indicated PDXs models engrafted in NOD-SCID or BALB/c Nude mice and treated daily orally with vehicle or debio-1347 (ES0204, Li0612, Li1035, BN2289, 60 mg/kg; GA1224, KI0551, 80 mg/kg; BL5001, day 1-14, 40 mg/kg; day 15-25, 60 mg/kg; $n=3$ mice per PDX model and treatment group). **d**, Debio-1347 $\Delta T/\Delta C$ response ratios in PDXs with *FGF3/4/19* amp ($n=3$) versus normal CN ($n=33$). **e**, Debio-1347 $\Delta T/\Delta C$ response ratios in PDXs with *FGFR1* amp ($n=3$), *FGFR2* amp ($n=4$), or *FGFR3* amp ($n=1$) versus normal CN ($n=28$). **f**, Correlations of *FGFR1*, *FGFR2*, *FGFR3*, *FGFR4*, or composite *FGFR* expression versus debio-1347 $\Delta T/\Delta C$ response ratios across PDXs. Composite *FGFR* expression was defined as high, if normalized expression >3 . **g**, Debio-1347 $\Delta T/\Delta C$ response ratios in PDXs expressing E18-truncated *FGFR2* ($n=6$) or *FGFR3* ($n=1$) versus no truncation ($n=29$). Data are represented as mean $\pm$ s.d. (**a, b**) or as median (centre line) $\pm$ IQR (25th to 75th percentile, box) and IQR $\pm$ 1.5 $\times$ IQR (whiskers) (**d, f**). *P* values were calculated with two-tailed unpaired Student's *t*-tests (**a**), one-tailed two-way ANOVA and FDR multiple-testing corrections using the two-stage step-up method from Benjamini, Krieger, and Yekutieli (**b, c**), two-tailed Wilcoxon rank-sum tests (**d, g**), one-tailed one-way ANOVA and Tukey's multiple-testing corrections (**e**), or two-tailed *t*-transformations of Pearson's *R* correlation coefficients (**f**).



CHAPTER 4

THE C-TERMINAL KINASE DOMAIN-BINDING AND SUPPRESSION MOTIF PREVENTS CONSTITUTIVE ACTIVATION OF FGFR2

Daniel Zingg*#, Chi-Chuan Lin*, **Julia Yemelyanenko**, Lukasz Wieteska, Sjors M. Kas, Onno B. Bleijerveld, Xue Chao, Jinhyuk Bhin, Catrin Lutz, Ellen Wientjens, Sjoerd Klarenbeek, Giulia Zanetti, Stefano Annunziato, Bjørn Siteur, Eline van der Burg, Anne Paulien Drenth, Marieke van de Ven, Lodewyk F.A. Wessels, Maarten Altelaar, John E. Ladbury# & Jos Jonkers#

*equal contribution

#co-corresponding author

Cancer Research, 2025, Sep 2;85(17):3234-3257.
<https://doi.org/10.1158/0008-5472.CAN-24-3349>

ABSTRACT

Genetic alterations in receptor tyrosine kinase genes can generate potent oncogenic drivers. Truncation of the *FGFR2* gene by its last exon 18 (E18) is caused by structural alterations, such as focal amplifications and gene fusions/rearrangements, as well as by mutations. All the E18-truncating *FGFR2* variants (*FGFR2*^{ΔE18}) act as strong driver alterations in cancer, and they commonly encode a receptor lacking the carboxy (C) terminal tail. In this study, we analyzed a compendium of *Fgfr2-E18* variants to uncover the mechanism by which loss of the C-tail renders FGFR2 oncogenic. Although permutation of previously annotated C-terminal FGFR motifs did not recapitulate the tumorigenicity of *FGFR2*^{ΔE18}, the functional annotation efforts led to the discovery of a C-terminal phenylalanine–serine motif that mediates binding of the C-tail to the kinase domain and thereby suppresses FGFR2 kinase activity. The permutation of this kinase domain–binding and suppression motif in conjunction with other FGFR2-regulatory C-terminal sites fully phenocopied the oncogenic competence of *FGFR2*^{ΔE18}. Together, these findings delineate how the C-terminal tail prevents FGFR2 from aberrant oncogenic activation.

Significance

A C-terminal phenylalanine–serine motif suppresses FGFR2 kinase activity by mediating binding of the C-terminal tail to the kinase domain, explaining how C-terminal truncation activates FGFR2 to promote tumorigenesis.

INTRODUCTION

Fibroblast growth factor (FGF) signaling controls key processes during embryogenesis and adult tissue homeostasis. FGF ligands exert their action through adhesion to FGF receptors (FGFR) that belong to the family of receptor tyrosine kinases (RTK; refs. [1–3]). FGFRs consist of two to three extracellular immunoglobulin-like domains that bind the FGF ligands, a transmembrane domain, the intracellular juxtamembrane and split tyrosine kinase domains that mediate RTK signaling, and the carboxy (C) terminal tail [2–4]. Upon FGF stimulation, FGFRs undergo conformational changes resulting in receptor dimerization and trans-autophosphorylation of the tyrosine kinase domains, which in turn induces downstream signaling transduction pathways. These are the MAPK, the PI3K–AKT–mTOR, the STAT3, and the PLCγ–PKC pathways [3,4]. FGFRs mediate these signaling cascades via FRS2 and CRKL adapter proteins [5–8] and by direct binding and activation of PIK3R1, STAT3, and PLCγ [9–14]. Accurate FGFR signaling kinetics are tightly controlled by protein tyrosine phosphatases, such as SHP2, which directly dephosphorylate FGFRs [14–16], and through the recruitment of inhibitory adapter proteins, such as GRB2 and GRB14 [16–20]. Notably, the FGFR C-terminal tail is crucial to fine-tune FGFR signaling [7,20]. Its most distal part is bound by GRB2 [17,18], and it contains the YXXΦ motif that is required for clathrin-mediated endocytosis [6,20]. Moreover, several C-terminal tyrosine and serine residues are phosphorylated by kinases that are part of the downstream signaling pathways. These feedback loops can elevate receptor endocytosis and attenuate FGFR signaling [20–28].

Dysregulated FGF/FGFR signaling leads to a variety of pathologies. These include cancer as well as congenital syndromes that affect skeleton development, impair the response to injury, and/or result in metabolic disorders [2–4]. In human cancers, the *FGFR* genes can be affected by hotspot missense mutations or structural alterations, such as amplifications and fusions/rearrangements [4,29–31]. Notably, focal *FGFR2* amplifications can produce C-terminally truncated isoforms owing to genomic breakpoints that perturb intron 17 or exon 18 (E18), which encodes the FGFR2 C-terminal tail [31–33]. *FGFR2* and *FGFR3* fusion/ rearrangement breakpoints typically occur in the intron 17/E18 hotspot, thus also producing C-terminally truncated receptors [4,30,31]. Moreover, E18 splice site, nonsense, and frameshift mutations can likewise generate C-tail–truncated FGFR2. These types of E18-truncating *FGFR2* alterations (*FGFR2^{ΔE18}*) are recurrently found in cholangiocarcinoma (intrahepatic, 8.7%; extrahepatic, 2.7%) as well as in gastric (2.0%), esophageal (0.8%), and breast (0.7%) cancers [4,29–31] and act as potent tumor driver genes [6,21,31]. Hence, loss of the C-tail is vital to render FGFR2 and potentially other FGFRs oncogenic.

However, the motifs or amino acid residues within the C-terminal tail that are responsible for suppressing oncogenic FGFR2 signaling remain unclear. Here, we analyzed a compendium of C-terminal *Fgfr2-E18* variants using *in vitro* signaling assays and somatic *in vivo* tumor

modeling to functionally dissect the tumor-suppressive motifs within the FGFR2 C-terminal region and to uncover the mechanism by which loss of the C-tail renders FGFR2 oncogenic.

RESULTS

Constitutive FGFR2^{ΔE18} activity promotes PI3K–AKT–mTOR and STAT3 signaling

FGF-10/FGFR2 signaling plays pivotal roles during mammary gland development and homeostasis in both the mammary epithelium and the mesenchyme [2], and human *FGFR2* oncogenes including *FGFR2*^{ΔE18} act as potent tumor drivers in the mouse mammary gland [31]. To study the oncogenic signaling cascades induced by FGFR2^{ΔE18} versus wild-type full-length FGFR2 (FGFR2^{FL}), we transduced NMuMG mouse mammary epithelial cells with lentiviruses containing *GFP*, *Fgfr2*^{FL}-*GFP*, or *Fgfr2*^{ΔE18}-*GFP* vectors. We isolated GFP^{hi} cells via FACS and verified the expression of signaling-active FGFR2^{FL}-GFP and FGFR2^{ΔE18}-GFP (Supplementary Fig. S1A and S1B). We subjected serum-starved GFP^{hi} cells to GFP coimmunoprecipitation (co-IP) coupled with mass spectrometry–based proteomics. Examination of the proteomic data identified factors that were significantly bound by GFP-tagged FGFR2^{FL} and/or FGFR2^{ΔE18} versus the GFP control (**Fig. 1A**; Supplementary Fig. S1C and S1D; Supplementary Table S3). GO analysis of FGFR2-bound proteins revealed GO terms related to RTK signaling and especially the FGF/FGFR signaling pathway (**Fig. 1B**). The FGFR2-bound proteins comprised several other RTKs including FGFR1 and FGFR3 (Supplementary Fig. S1E) as well as key factors of canonical FGFR signaling including the adapter proteins FRS2 and CRKL; the effector protein STAT3; and phosphatases, such as SHP2 (**Fig. 1A and C**). However, the proteomic data revealed enrichments for neither the adapter protein GRB14 nor the phosphatases PTPN1 and PTPN2 (**Fig. 1C**) that regulate FGFR1 and FGFR3 activity [15,19].

We next treated serum-starved GFP^{hi} cells with the FGFR2 ligand FGF-10 [2], which promoted FGFR2 signaling in both *Fgfr2*^{FL}-*GFP*–expressing and *Fgfr2*^{ΔE18}-*GFP*–expressing cells versus *GFP* control cells, as evidenced by the phosphorylation of ERK1/2 and RPS6 (**Fig. 1D and E**; Supplementary Fig. S1F and S1G). Prolonged FGFR2 stimulation with FGF-10 fostered receptor downregulation and signaling loss in *Fgfr2*^{FL}-*GFP* cells, whereas *Fgfr2*^{ΔE18}-*GFP* cells maintained receptor levels and signaling (Supplementary Fig. S1F and S1G).

Although expression of *Fgfr2*^{ΔE18} resulted in baseline ERK1/2 phosphorylation, the MAPK signaling pathway components ERK1/2 and p90 RSK were predominantly induced through FGF-10 in both *Fgfr2*^{FL}-*GFP*–expressing and *Fgfr2*^{ΔE18}-*GFP*–expressing cells (**Fig. 1D and E**; Supplementary Fig. S1F and S1G). In contrast, marked phosphorylation of the PI3K–AKT–mTOR signaling pathway components AKT, MTOR, RPS6, EIF4E, EIF4EBP1, p70 S6K, and

EIF4B and the resulting stabilization of MYC prevailed in *Fgfr2^{ΔE18}-GFP* cells independently of FGF-10 stimulation (**Fig. 1D** and **F**; Supplementary Fig. S1F–S1I). Phosphorylation of STAT3 was also independent of FGF-10 stimulation and occurred exclusively in cells expressing *Fgfr2^{ΔE18}-GFP* (**Fig. 1D** and **G**). Conversely, PLCγ–PKC signaling was mediated by FGF-10 solely in *Fgfr2^{FL}-GFP*-expressing cells (**Fig. 1D** and **G**), consistent with the requirement of a FGFR2 C-terminal tyrosine residue for PLCγ activation [12–14] that is absent in C-terminally truncated FGFR2^{ΔE18}. The ATP-competitive inhibitor fexagratinib that blocks the catalytic function of the FGFR2 kinase domain [34] completely abrogated FGFR2^{FL} and FGFR2^{ΔE18} signaling (**Fig. 1D–G**; Supplementary Fig. S1H and S1I). Thus, the ligand-inducible activity of FGFR2^{FL} and the constitutive activity of FGFR2^{ΔE18} both relied on canonical kinase domain function to stimulate MAPK and PLCγ–PKC versus PI3K–AKT–mTOR and STAT3 signaling transduction pathways.

Disruption of FGFR2 C-tail regulatory motifs does not phenocopy FGFR2^{ΔE18}

Loss of the complete C-tail in FGFR2 promoted PI3K–AKT–mTOR and STAT3 signaling (**Fig. 1D**), and C-terminal truncation converts the FGFR2 proto-oncogene into an effective oncogenic driver [31]. To dissect the signaling and tumor-suppressive function of the FGFR2 C-tail, we asked whether previously annotated C-terminal single amino acid residues and motifs were required to suppress FGFR2^{FL} oncogenicity. These conserved residues/motifs have been studied mostly in FGFR1 and to a lesser extent in FGFR2 and are relevant for proper control of receptor activity. As displayed in **Fig. 2A**, these are a tyrosine residue (human, Y769; mouse, Y674) that is critical for PLCγ binding and activation [12–14] and that is also part of the YXXΦ consensus motif that mediates receptor endocytosis [6,20]; a tyrosine residue (human, Y779; mouse, Y684) that is connected to suppression of PI3K signaling and oncogenic transformation [21,22]; a serine residue (human, S780; mouse, S685) that can bind and is a phosphotarget of p38 MAPK and ERK1/2, thus mitigating FGFR signaling [23–25]; a serine residue (human, S782; mouse, S687) that can bind and is a phosphotarget of 14-3-3 and PKC, thus enhancing FGFR signaling [35,36]; two serine residues (human, S791/2; mouse, S696/7) that can bind and are phosphotargets of p90 RSK, thus attenuating MAPK signaling and augmenting receptor endocytosis [26–28]; and a proline-rich motif and the GRB2-binding site, which are both implicated in signaling attenuation [7,16–18].

We first assessed whether in human cancers these residues/motifs are affected by *FGFR2* missense mutations. We analyzed oncogenomic data from Foundation Medicine that were derived from 249,570 targeted tumor sequencing assays [31] and from cBioPortal that comprehends 226 non-redundant tumor-sequencing studies of 105,424 samples [29]. We identified 5,447 (2.18%) and 1,192 (1.13%) samples with *FGFR2* missense mutations in the respective Foundation Medicine and cBioPortal datasets (**Fig. 2B**). A fraction of these mutations located to E18 of *FGFR2*, thus affecting amino acid residues of the FGFR2 C-tail (**Fig. 2C**). Missense mutation counts were recurrent at previously annotated FGFR2

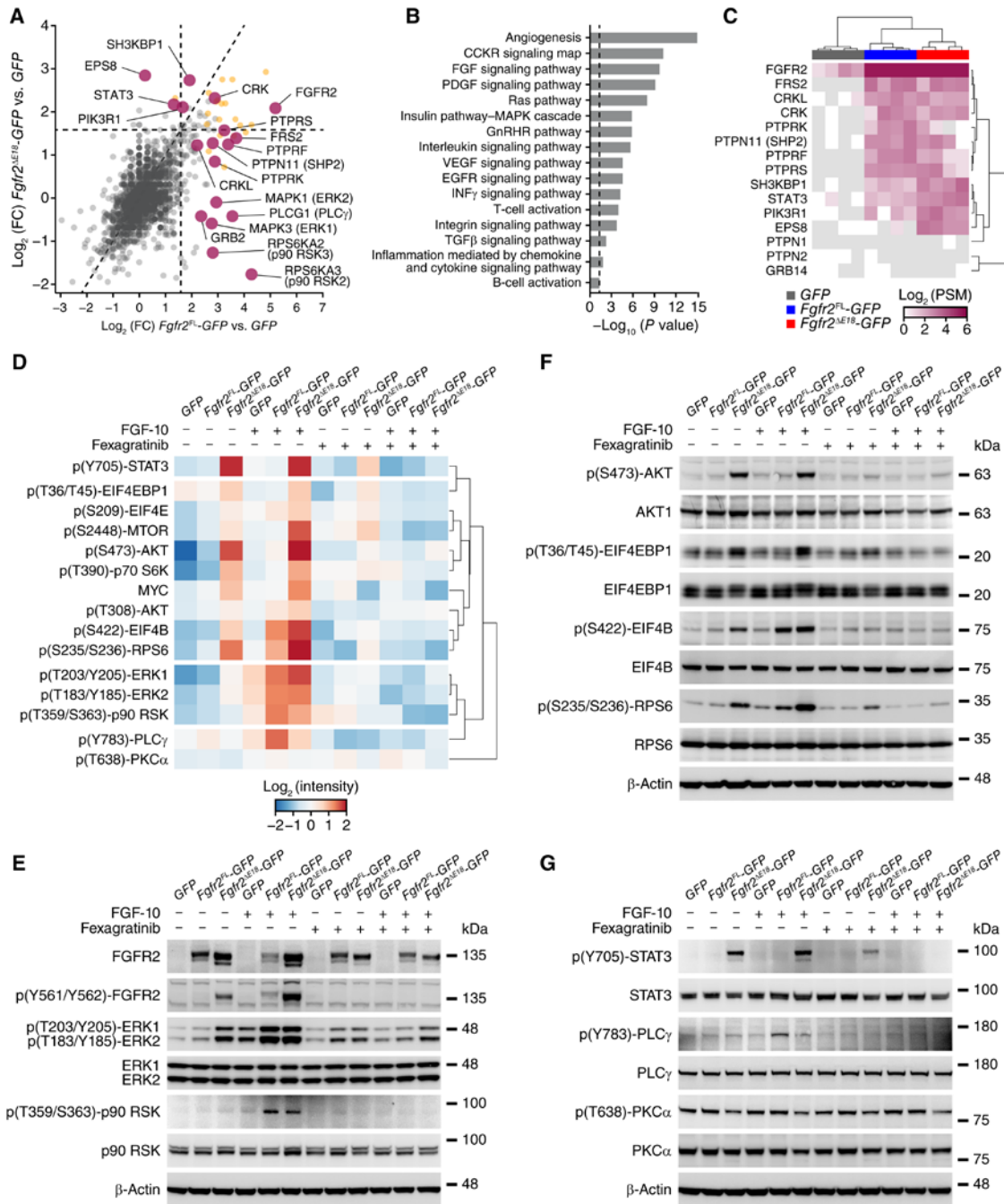


Figure 1. C-terminally truncated FGFR2 signals through PI3K–AKT–mTOR and STAT3.

A, Proteomic data showing differentially enriched proteins after GFP IP on serum-starved GFP^{hi} NMuMG cells expressing GFP, *Fgfr2^{FL}-GFP*, or *Fgfr2^{ΔE18}-GFP*. Proteins with FC ≥3 and FDR-adjusted *P* values < 0.05 for *Fgfr2^{FL}-GFP* and/or *Fgfr2^{ΔE18}-GFP* PSMs vs. GFP PSMs were considered enriched and are highlighted in yellow. FGFR2 signaling components are highlighted in purple. Log₂ (FC) and *P* values were calculated using PSM +1 data transformation and one-tailed two-way ANOVA combined with FDR multiple testing corrections. **B**, GO pathway enrichment analysis for proteins enriched in *Fgfr2^{FL}-GFP* and/or *Fgfr2^{ΔE18}-GFP* vs. GFP samples using two-tailed Fisher exact tests with Bonferroni multiple testing corrections. Significantly enriched pathways (*P* < 0.05) are indicated. CCKR, cholecystokinin receptor; GnRHR, gonadotropin-releasing hormone receptor; PDGF, platelet-derived growth factor. **C**, Heatmap depicting log₂ (PSM) values for indicated FGFR2 signaling components in indicated GFP-IP samples. Gray squares indicate undetected proteins (PSM = 0). Unsupervised hierarchical clustering was performed on row mean-centered log₂ (PSM +1) values using the Pearson distance measure and Ward minimum variance clustering. Data in **A** to **C** are represented as *n* = 4 independent replicates per group of two independent experiments. **D–G**, Quantifications of Western blots showing expression and phosphorylation of indicated proteins in serum-starved GFP^{hi} NMuMG cells expressing GFP, *Fgfr2^{FL}-GFP*, or *Fgfr2^{ΔE18}-GFP* and treated for 15 minutes with 20 ng/mL FGF-10 and/or for 3 hours with 100 nmol/L fexagratinib. Quantifications in **(D)** are based on Western blots in **E** to **G** and Supplementary Fig. S1H and S1I. Band intensities were normalized to their individual lane backgrounds and to the corresponding total protein and β-actin bands. Unsupervised hierarchical clustering was performed on row mean-centered log₂ (intensity) values using the Pearson distance measure and Ward minimum variance clustering. Data in **D** to **G** represent one replica of two independent experiments.

4

hotspots [4,29,31] locating to the extracellular, transmembrane, and kinase domains (**Fig. 2B**). The only C-tail amino acid residues that showed significant enrichment of missense mutations were lysine 770 (L770) that is part of the endocytosis consensus motif, glutamic acid 806 (E806) that is part of the proline-rich motif, and the previously unreported hotspot site E778 (**Fig. 2C**).

Next, we examined whether the FGFR2 C-terminal tail is indeed required for the recruitment of factors that were previously reported to interact with the C-tails of FGFRs. Therefore, we analyzed the proteomic data to identify differentially enriched proteins in FGFR2^{ΔE18}-GFP versus FGFR2^{FL}-GFP samples. The capability of FGFR2^{ΔE18}-GFP to bind PLCγ, ERK1/2, p90 RSK2/3, and GRB2 was markedly compromised as compared with FGFR2^{FL}-GFP (**Figs. 1A, 3A** and **B**). The expression of FLAG-*Fgfr2* variants followed by FLAG co-IP verified that FLAG-FGFR2^{FL} and FLAG-FGFR2^{ΔE18} were both bound by FRS2, whereas FLAG-FGFR2^{ΔE18} had lost binding of ERK1/2 and p90 RSK (Supplementary Fig. S2A). Thus, binding of these factors to FGFR2 depended on its C-terminal tail. In contradiction to previous studies on the recruitment of 14-3-3 proteins, PKC, and p38 MAPK to FGFRs [22,35,36], we found no evidence for these factors to specifically bind to FGFR2 (**Fig. 3B**; Supplementary Fig. S2A and S2B).

We next asked whether disruptions of the previously annotated C-tail residues/motifs in FGFR2^{FL} were sufficient to recapitulate FGFR2^{ΔE18} oncogenicity. We generated a series of *Fgfr2-E18* variants in which we either deleted or permuted the reported C-terminal regulatory sites (**Fig. 3C**). The expression of *Fgfr2* variants that were mutant for the PLCγ-binding or PKC-binding sites (*Fgfr2^{Y674F}* and *Fgfr2^{S687G}*) lowered phosphorylation of PLCγ and PKCα as compared with cells expressing *Fgfr2^{FL}* (**Fig. 3D**; Supplementary Fig.

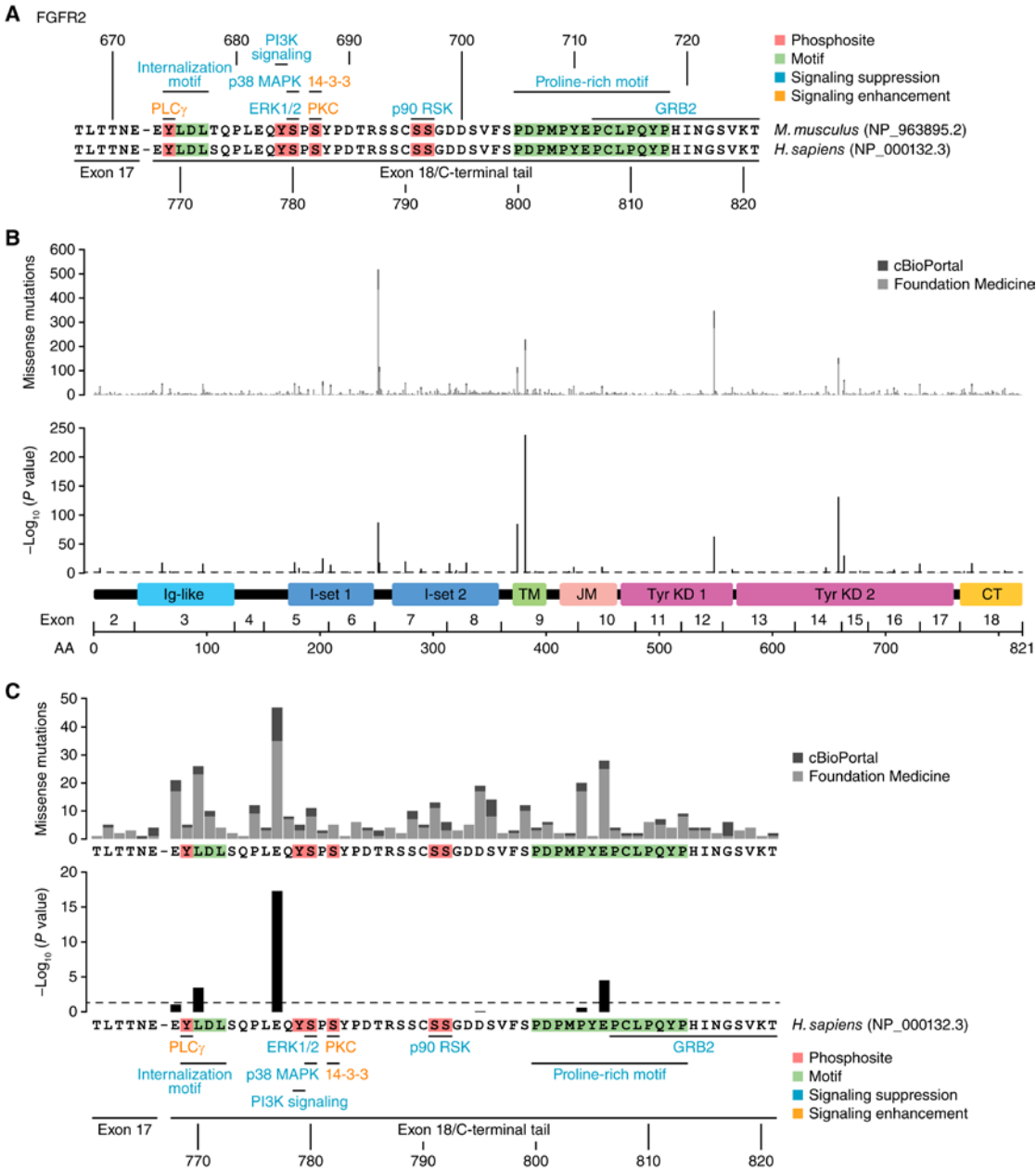


Figure 2. Occurrence of *FGFR2* missense mutations in human cancer.

A, Graphical representation of the *FGFR2-E18*-encoded amino acid sequence of the C-terminal tail and its previously annotated binding sites and motifs. **B** and **C**, Analysis of 249,570 (Foundation Medicine) and 105,424 (cBioPortal) pan-cancer samples for the occurrence of *FGFR2* missense mutations (Foundation Medicine: $n = 5,447$; incidence = 2.18%. cBioPortal: $n = 1,192$; incidence = 1.13%). Mutation counts (top) and significant enrichments over background mutational rates (bottom) across the full *FGFR2* amino acid sequence (**B**) and the C-tail (**C**). P values were calculated using one-tailed binomial tests with Bonferroni multiple testing corrections. AA, amino acid; CT, C-terminal tail; JM, juxtamembrane; KD, kinase domain; TM, transmembrane.

S2C). The expression of *Fgfr2* containing a mutant endocytosis consensus motif (*Fgfr2*^{YLDL674-7FANA}) enhanced MAPK signaling nearly to the extent of *Fgfr2*^{ΔE18} but only mildly affected RPS6 and STAT3 phosphorylation (**Fig. 3D** and **E**). FLAG-FGFR2 binding to ERK1/2 and p90 RSK was prevented when the respective C-terminal binding sites were targeted in *FLAG-Fgfr2*^{Y684F/S685A} and *FLAG-Fgfr2*^{S696A/S697A} (Supplementary Fig. S2A). However, the expression of *Fgfr2*^{Y684F/S685A} or *Fgfr2*^{S696A/S697A} only modestly elevated phosphorylation of ERK1/2 and RPS6 and had virtually no effect on p38 MAPK, p90 RSK, and STAT3 phosphorylation as compared with the expression of *Fgfr2*^{FL} (**Fig. 3D** and **E**; Supplementary Fig. S2C). Of note, none of the *Fgfr2*-E18 mutants analyzed were able to phenocopy the strong inductions of the PI3K-AKT-mTOR and STAT3 signaling pathways observed in mammary epithelial cells expressing *Fgfr2*^{ΔE18} (**Fig. 3D** and **E**; Supplementary Fig. S2C).

Finally, we evaluated the oncogenicity of the C-terminal *Fgfr2*-E18 mutants in mouse mammary tumor models representative of different breast cancer subtypes, including invasive lobular carcinoma hallmarked by the loss of E-cadherin [37]. To this end, using intraductal injection [31,38], we somatically delivered lentiviruses encoding individual *Fgfr2* variants to the mammary glands of wild-type or *Wap-cre;Cdh1*^{F/F} female mice, the latter of which express Cre recombinase in the mammary epithelium, resulting in conditional ablation of loxP site-flanked E-cadherin (*Cdh1*) alleles [37]. In wild-type mice, the majority of the *Fgfr2*-E18 variants tested did not affect mammary tumor latencies as compared with mice injected with *Fgfr2*^{FL} (**Fig. 3F-H**). The remaining *Fgfr2* variants mutant for the PLCγ-binding site (*Fgfr2*^{Y674F}), the endocytosis consensus motif (*Fgfr2*^{YLDL674-7FANA}), the PI3K/ERK negative feedback loop regulatory sites (*Fgfr2*^{Y684F/S685A}), and the p90 RSK binding site (*Fgfr2*^{ΔS696/S697} and *Fgfr2*^{S696A/S697A}) modestly decreased tumor latencies (**Fig. 3F-H**). In *Wap-cre;Cdh1*^{F/F} mice, the endocytosis consensus motif (*Fgfr2*^{ΔYLDL674-7} and *Fgfr2*^{YLDL674-7FANA}), the PI3K/ERK negative feedback loop regulatory site (*Fgfr2*^{Y684F/S685A}), and the p90 RSK binding site (*Fgfr2*^{ΔS696/S697}) mutants similarly accelerated mammary tumor onsets as compared with *Fgfr2*^{FL} (**Fig. 3I-K**). However, regardless of *Cdh1* mutation status, none of the *Fgfr2*-E18 variants recapitulated the rapid tumor onset and short tumor-specific survival observed for *Fgfr2*^{ΔE18} in wild-type and *Wap-cre;Cdh1*^{F/F} animals (**Fig. 3F-K**; Supplementary Fig. S2D-S2I). Histopathologic evaluation of the mammary glands of mice injected with *Fgfr2*-E18 variants revealed mostly healthy tissue or low-grade lesions comparable with mammary glands of mice injected with *Fgfr2*^{FL} (Supplementary Fig. S3A and S3B). The remaining glands contained high-grade tumors, among which a substantial fraction showed no or low FGFR2 expression (Supplementary Fig. S3A and S3C). The FGFR2-negative lesions likely represented spontaneous tumor events owing to the mammary tumor-prone genetic backgrounds of FVB wild-type and especially *Wap-cre;Cdh1*^{F/F} female mice [31,37,39]. In sum, few of the previously annotated FGFR2 C-terminal regulatory sites harbored recurrent missense mutations in human cancers, but none of these sites were required to suppress a tumor driver phenotype reminiscent of FGFR2^{ΔE18} in our experimental settings.

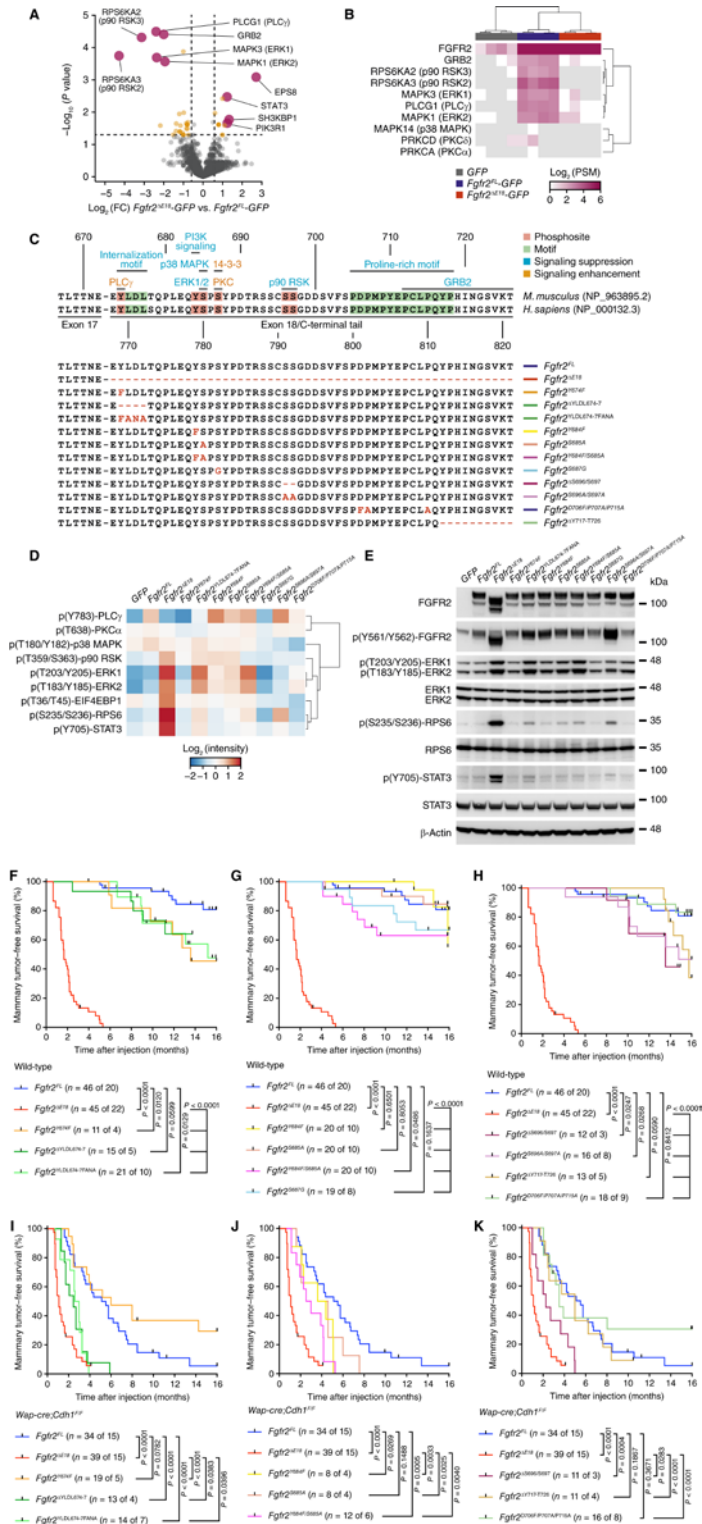


Figure 3. Ablation of known C-terminal FGFR2 motifs does not phenocopy FGFR2^{ΔE18} oncogenicity.

A, Proteomic data showing differentially enriched proteins after GFP IP on serum-starved GFP^{hi} NMuMG cells expressing GFP, *Fgfr2*^{FL}-GFP, or *Fgfr2*^{ΔE18}-GFP. Volcano plot shows differentially enriched proteins in *Fgfr2*^{ΔE18}-GFP vs. *Fgfr2*^{FL}-GFP samples. Significantly changed proteins were defined with $-0.58 \geq \log_2(\text{FC}) \geq 0.58$ and FDR-adjusted *P* values < 0.05 and are highlighted in yellow. FGFR2 signaling components enriched over GFP controls (*Fgfr2*^{FL}-GFP PSMs and/or *Fgfr2*^{ΔE18}-GFP PSMs vs. GFP PSMs, $\text{FC} \geq 3$) are highlighted in purple. $\log_2(\text{FC})$ and *P* values were calculated using PSM +1 data transformation and two-tailed unpaired Student *t* tests combined with FDR multiple testing corrections. **B**, Heatmap depicting $\log_2(\text{PSM})$ values for indicated FGFR2-CT-binding factors in indicated GFP-IP samples. Gray squares indicate undetected proteins (PSM = 0). Unsupervised hierarchical clustering was performed on row mean-centered $\log_2(\text{PSM} + 1)$ values using the Pearson distance measure and Ward minimum variance clustering. Data in **A** and **B** are represented as *n* = 4 independent replicates per group of two independent experiments. **C**, Graphical representation of the FGFR2-CT amino acid sequence and its previously annotated binding sites and motifs. *Fgfr2*-E18 variants used to target the C-terminal binding sites/motifs are indicated. **D** and **E**, Quantifications of Western blots showing expression and phosphorylation of indicated proteins in serum-starved NMuMG cells expressing indicated *Fgfr2* variants. Quantifications in **D** are based on Western blots in **E** and Supplementary Fig. S2C. Band intensities were normalized to their individual lane backgrounds and to the corresponding total protein and β -actin bands. Unsupervised hierarchical clustering was performed on row mean-centered $\log_2(\text{intensity})$ values using the Pearson distance measure and Ward minimum variance clustering. Data in **D** and **E** represent one replica of three independent experiments. **F–K**, Kaplan–Meier curves showing mammary tumor-free survival of wild-type (**F–H**) or *Wap-cre;Cdh1*^{T9F} (**I–K**) mice intraductally injected with lentiviruses encoding indicated *Fgfr2* variants. Cohort counts (*n*) were injected mammary glands per number of mice. The *Fgfr2*^{FL} and *Fgfr2*^{ΔE18} curves in **F** and **I** are duplicated in **G** and **H** and in **J** and **K**, respectively. The *Fgfr2*^{FL}, *Fgfr2*^{ΔE18}, and *Fgfr2*^{ΔY717-T726} curves were previously published [31]. *P* values were calculated using log-rank (Mantel–Cox) tests.

4

The C-terminal F703–S704 motif is critical for FGFR2 signaling suppression

The benign nature of previously described FGFR2 C-terminal motifs prompted us to perform an unbiased screen of the FGFR2 C-terminal region for amino acid residues with tumor-suppressive attributes. Phosphorylation of RPS6 represented a robust marker to discriminate ligand-independent signaling induction by FGFR2^{ΔE18} versus FGFR2^{FL} (**Fig. 1D** and **F**). Therefore, we developed a FACS assay to systematically measure FGFR2 activities via phospho-RPS6 intensities in cells expressing different *Fgfr2*-GFP variants (Supplementary Fig. S4A). In contrast to cells expressing *Fgfr2*^{FL}-GFP or the GFP control, the expression of *Fgfr2*^{ΔE18}-GFP enhanced the mean phospho-RPS6 intensity specifically in the GFP-positive cell population (**Fig. 4A** and **B**; Supplementary Fig. S4A). We also analyzed fluorescence intensities at an individual cell level. In *Fgfr2*^{FL}-GFP-expressing cells, GFP signal intensities, as a surrogate for FGFR2 expression levels, only correlated moderately with phospho-RPS6 intensities, but we observed a strong GFP-to-phospho-RPS6 correlation in *Fgfr2*^{ΔE18}-GFP cells (**Fig. 4C**). It is noteworthy that the expression of *Fgfr2*^{YLDL674-7FANA}-GFP (mutant endocytosis consensus motif), *Fgfr2*^{Y684F/S685A}-GFP (mutant PI3K/ERK negative feedback loop regulatory sites), and *Fgfr2*^{S696A/S697A}-GFP (mutant p90 RSK binding site) was hardly able to induce phospho-RPS6 (**Fig. 4B**).

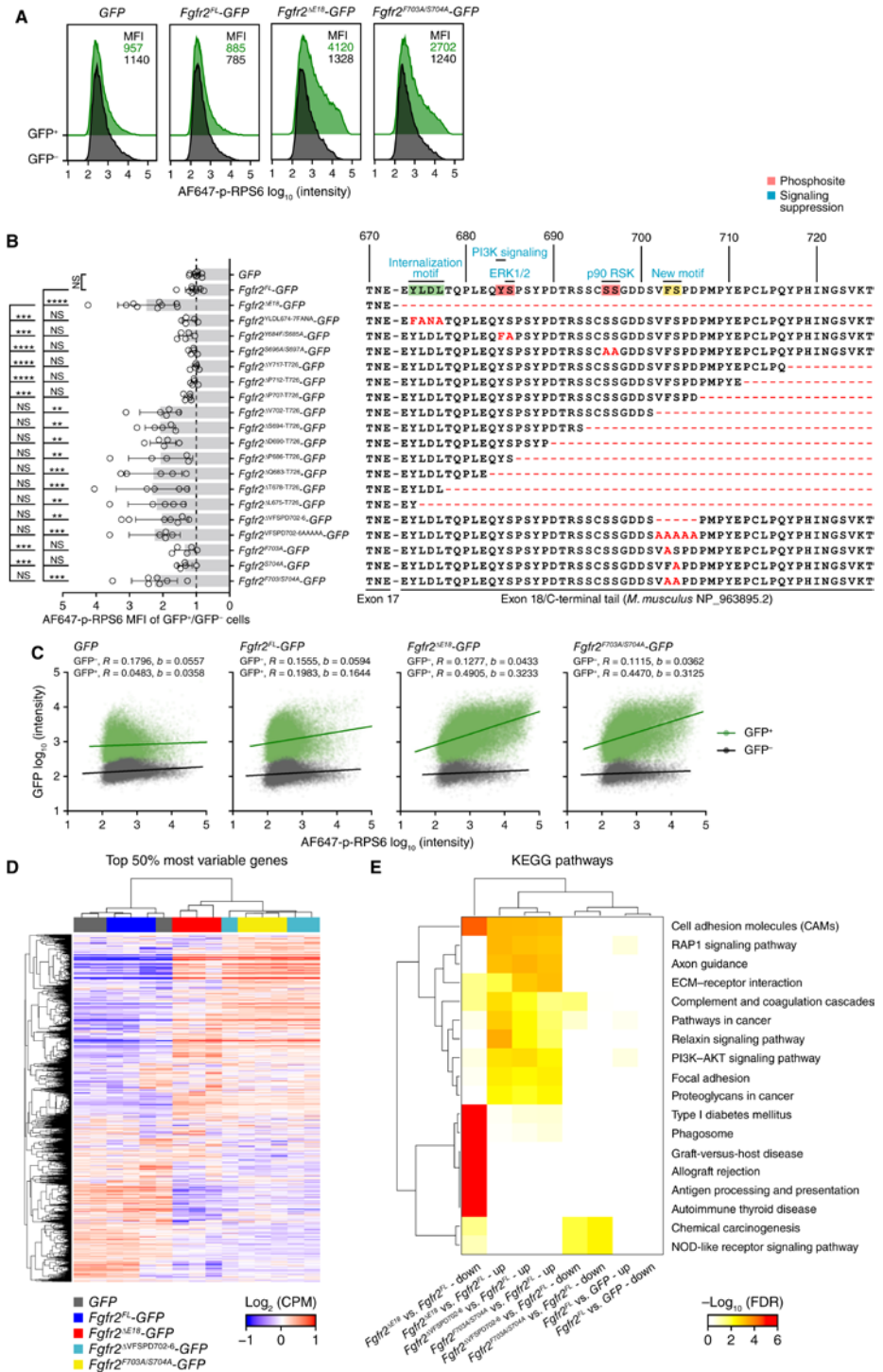
We next analyzed a series of C-terminal deletion derivatives of *Fgfr2*^{FL}-GFP. Progressive truncation of the FGFR2 C-terminal tail resulted in phospho-RPS6 induction following a bimodal pattern. Variants with truncations of the distal C-tail showed phospho-RPS6

intensities comparable with FGFR2^{FL}, whereas variants with truncations affecting the proximal C-tail beyond proline 707 (P707) significantly induced RPS6 phosphorylation to levels that approximated those observed with FGFR2^{ΔE18} (**Fig. 4B**). This was in line with the differential oncogenic potentials of proximal versus distal E18-truncating mutations found in human cancers [21,31] and suggested that the VFSPD702-6 sequence preceding P707 or parts thereof may be critical for FGFR2 signaling suppression. Deletion or substitution of the VFSPD residues (*Fgfr2*^{ΔVFSPD702-6}-GFP and *Fgfr2*^{VFSPD702-6AAAAA}-GFP) indeed recapitulated the phospho-RPS6 induction observed with the *Fgfr2*^{ΔV702-T726}-GFP distal C-tail ablation variant (**Fig. 4B**). To further map VFSPD702-6, we generated the *Fgfr2*^{F703A}-GFP, *Fgfr2*^{S704A}-GFP, and *Fgfr2*^{F703A/S704A}-GFP variants. Dual targeting of the phenylalanine and the serine in FGFR2^{F703A/S704A} drove robust RPS6 phosphorylation, whereas the single amino acid mutants only produced marginal phospho-RPS6 signals (**Fig. 4A–C**).

Next, we analyzed RNA-seq profiles derived from GFP^{hi} cells expressing GFP, *Fgfr2*^{FL}-GFP, *Fgfr2*^{ΔE18}-GFP, *Fgfr2*^{ΔVFSPD702-6}-GFP, or *Fgfr2*^{F703A/S704A}-GFP (Supplementary Fig. S4B; Supplementary Table S4). Unsupervised clustering of the top 50% variable genes revealed that FGFR2^{FL} had no effect on gene expression when compared with the GFP control (**Fig. 4D**; Supplementary Fig. S4B). In contrast, FGFR2^{ΔE18}, FGFR2^{ΔVFSPD702-6}, and FGFR2^{F703A/S704A} promoted a shared set of differentially expressed genes as compared with FGFR2^{FL} (**Fig. 4D**; Supplementary Fig. S4B). We explored the biological processes underlying these genes, which revealed terms related to MAPK and PI3K–AKT signaling, the cross-talk with the immune compartment, and overall protumorigenic processes (**Fig. 4E**; Supplementary Fig. S4C and S4D). Taken together, disruption of the C-terminal F703–S704 motif fostered *in vitro* oncogenic signaling that mimicked FGFR2^{ΔE18}, thus pinpointing F703 and S704 to be pivotal for C-terminal tail–mediated suppression of FGFR2 activity.

Figure 4. F703 and S704 in the C-tail are critical to suppress FGFR2 signaling activity.

A and **B**, FACS was used to measure phospho (p)-RPS6 induction in serum-starved NMuMG cells expressing indicated *Fgfr2*-GFP variants. p-RPS6 mean fluorescence intensities (MFI) were analyzed in GFP-positive vs. GFP-negative cells (**A**) to calculate p-RPS6 MFI ratios (**B**). Histograms in **A** show one representative replica of data quantified in **B**, which is represented as mean ± SD of GFP, *Fgfr2*^{FL}-GFP, and *Fgfr2*^{ΔE18}-GFP, *n* = 9; *Fgfr2*^{ΔVFSPD702-6}-GFP and *Fgfr2*^{F703A/S704A}-GFP, *n* = 7; and other variants, *n* = 5 independent replica. *P* values were calculated using a one-tailed one-way ANOVA and FDR multiple testing correction. NS, not significant (*P* ≥ 0.05); **, *P* < 0.01; ***, *P* < 0.001; ****, *P* < 0.0001. Exact *P* values are shown in Supplementary Table S2. **C**, Comparison of GFP intensity vs. p-RPS6 intensity using Pearson *R* correlation coefficient and the slope (*b*) calculated by simple linear regression across GFP-positive and GFP-negative cell populations. *X–Y* plots show one representative replica of the GFP and indicated *Fgfr2*-GFP variants from **B**. **D**, RNA-seq data showing differential gene expression in serum-starved GFP^{hi} NMuMG cells expressing indicated *Fgfr2*-GFP variants. Unsupervised hierarchical clustering was performed on row mean-centered log₂ (counts per million, CPM) values of the top 50% most variable genes defined by median absolute deviation using the Pearson distance measure with complete clustering across *n* = 3 independent replica per group. **E**, Kyoto Encyclopedia of Genes and Genomes pathway enrichment analysis for differentially expressed genes identified from indicated two-group comparisons of RNA-seq data (Supplementary Fig. S4B; Supplementary Table S4) using two-tailed Fisher exact tests with FDR multiple testing corrections using the Benjamini–Hochberg method. Pathways that were significantly enriched (FDR < 0.01) by differentially expressed genes in at least one of the comparisons are represented in the heatmap. Negative log₁₀ (FDR) values were clustered unsupervised using the Pearson distance measure with complete clustering.



The F703–S704 motif is dispensable for FGFR2 endocytosis

We next tested whether the C-terminal F703–S704 motif regulates FGFR2 endocytosis. We made use of the SPIDA methodology in which HALO-tagged cell surface proteins are sequentially labeled using first, a cell-impermeable fluorescent AF660-HTL, and second, a cell-permeable fluorescent TMR-HTL, followed by FACS to measure AF660-HTL and TMR-HTL fluorescent intensities [40]. We subjected cells expressing *HALO-Fgfr2* variants to SPIDA (Supplementary Fig. S5A and S5B) and computed relative cell surface versus internal FGFR2 abundances during FGF-10 stimulation by applying the SPIDA quantification method [40]. Cells expressing *HALO-Fgfr2^{FL}* lost cell surface FGFR2 upon FGF-10 stimulation, whereas intracellular FGFR2 levels only marginally increased (**Fig. 5A–C**; Supplementary Fig. S5B and S5C). This indicated that the internalized receptors were degraded [40] in line with the overall reduced FGFR2^{FL} levels observed during FGF-10 stimulation (**Fig. 1E**; Supplementary Fig. S1F and S1G). In contrast, HALO-FGFR2^{ΔE18}–mutant and endocytosis consensus motif–mutant HALO-FGFR2^{YLDL674-7FANA} were both inert to ligand-induced receptor internalization and degradation (**Fig. 5A–C**; Supplementary Fig. S5C). Recruitment of p90 RSK to the FGFR1 C-terminal region was previously shown to enhance receptor endocytosis [28], yet the HALO-FGFR2^{S696A/S697A} mutant preventing p90 RSK binding hardly showed impaired receptor internalization (**Fig. 5B and C**; Supplementary Fig. S5C). The HALO-FGFR2^{Y684F/S685A} derivative lacking the PI3K/ERK negative feedback loop regulatory sites was internalized as efficiently as HALO-FGFR2^{FL} (**Fig. 5B and C**; Supplementary Fig. S5C), which was in line with previous reports [24,25]. Importantly, endocytosis of HALO-FGFR2^{F703A/S704A} was unaffected during FGF-10 exposure (**Fig. 5A–C**; Supplementary Fig. S5C). Thus, the C-terminal F703–S704 motif suppressed constitutive FGFR2 activation through a mechanism other than receptor internalization.

The F703–S704 motif suppresses FGFR2 signaling through kinase domain binding

The FGFR2 C-tail can directly interfere with the FGFR2 kinase domain to moderate receptor signaling [7]. We therefore explored the possibility that the F703 and/or S704 residues might be relevant for interference of the C-tail with the kinase domain within FGFR2. We made use of a HSQC-2D-NMR spectroscopy dataset [7] to assess CSPs of a recombinant ¹⁵N-labeled polypeptide that included the last 58 amino acids of the human FGFR2 C-tail (residues 764–821) upon titration of the recombinant unlabeled human FGFR2 kinase domain (residues 466–763). A monophosphorylated FGFR2 kinase domain (FGFR2-KD^{pY1}) variant was used, in which, except for the activation loop tyrosine 656 (Y656), all phosphorylatable tyrosine residues were substituted for phenylalanine residues, thus representing the phosphorylation state of the receptor under basal conditions [7,41]. Mapping of the chemical shifts while increasing the ratio of the concentration of FGFR2-KD^{pY1} to the C-tail polypeptide revealed several C-terminal residues with significant perturbations corresponding to binding with the kinase domain (**Fig. 5D**). Strikingly, F798 and S799 (mouse, F703 and S704) showed the most prominent CSPs in the HSQC-2D-NMR

spectra (**Fig. 5D** and **E**; Supplementary Fig. S5D), indicating that these two residues are critical for binding of the C-terminal tail to the kinase domain.

To address whether binding of the FGFR2 C-terminal tail to the kinase domain relied on F798 and S799, we produced human recombinant polypeptides of the wild-type C-tail (residues 764–821) and its F798A, S799A, and F798A/S799A derivatives (Supplementary Fig. S5E and S5F) and human recombinant 6xHis-tagged FGFR2-KD^{pY1} that retains the juxtamembrane domain (6xHis-FGFR2-JM-KD^{pY1}; residues 413–763). We measured the binding of the C-terminal peptides to 6xHis-FGFR2-JM-KD^{pY1} using microscale thermophoresis assays. Substitution of F798 or S799 to an alanine reduced the affinity of the C-tail to the kinase domain (**Fig. 5F**). The F798A/S799A double-mutant C-tail showed a strongly reduced binding affinity for the kinase domain, corresponding to an increased dissociation constant of an order of magnitude as compared with the wild-type C-tail (**Fig. 5F** and **G**). GST IP confirmed that GST-tagged wild-type C-tail strongly bound to 6xHis-FGFR2-JM-KD^{pY1}, whereas the F798A and S799A single mutants and especially the F798A/S799A double mutant showed impaired binding to the kinase domain (**Fig. 5H**; Supplementary Fig. S5G and S5H). We modeled the structure of the FGFR2-KD^{pY1} and a C-terminal polypeptide sequence centered around F798–S799 using AlphaFold [42] and the ATP-bound FGFR2-KD^{pY1} crystal structure [7]. This confirmed that the amino acid sequence proximal to and including the F798–S799 motif was able to interact with the hydrophobic residues proximal to the ATP-bound catalytic site, whereas the F798A/S799A permutation resulted in the inability of the peptide to dock into this site (**Fig. 5I**).

Next, we examined the activity of the FGFR2 kinase domain in the presence of the C-terminal tail. Coincubation of 6xHis-FGFR2-JM-KD^{pY1} (active kinase) with human unphosphorylated kinase-dead FGFR2-KD (6xHis-FGFR2-KD^{K517L}; residues 466–763) that functions as the kinase substrate resulted in robust trans-phosphorylation of 6xHis-FGFR2-KD^{K517L} in the presence of ATP and Mg²⁺ (**Fig. 5J** and **K**). However, coincubation together with the untagged wild-type C-tail strongly hindered phosphorylation of FGFR2 kinase substrate (**Fig. 5J** and **K**). In contrast, the F798A-, the S799A-, and particularly the F798A/S799A-mutant C-tails showed reduced suppression of FGFR2 kinase domain activity (**Fig. 5J** and **K**). Hence, the FGFR2 C-tail depended on the F798–S799 motif to efficiently bind to the kinase domain and suppress its activity. We therefore denoted the F798–S799 (mouse, F703–S704) residues of FGFR2 the kinase domain-binding and suppression (KDBS) motif.

Finally, we assessed whether the F798–S799 KDBS motif is conserved across C-terminal regions of the FGFR homologs. Amino acid sequence alignments showed that F798 and S799 are indeed conserved across vertebrate FGFR2 orthologs (Supplementary Fig. S6A) and between the FGFR1, FGFR2, and FGFR4 paralogs (Supplementary Fig. S6B). Thus, other FGFRs might depend on the C-terminal phenylalanine-serine motif to regulate kinase domain activity.

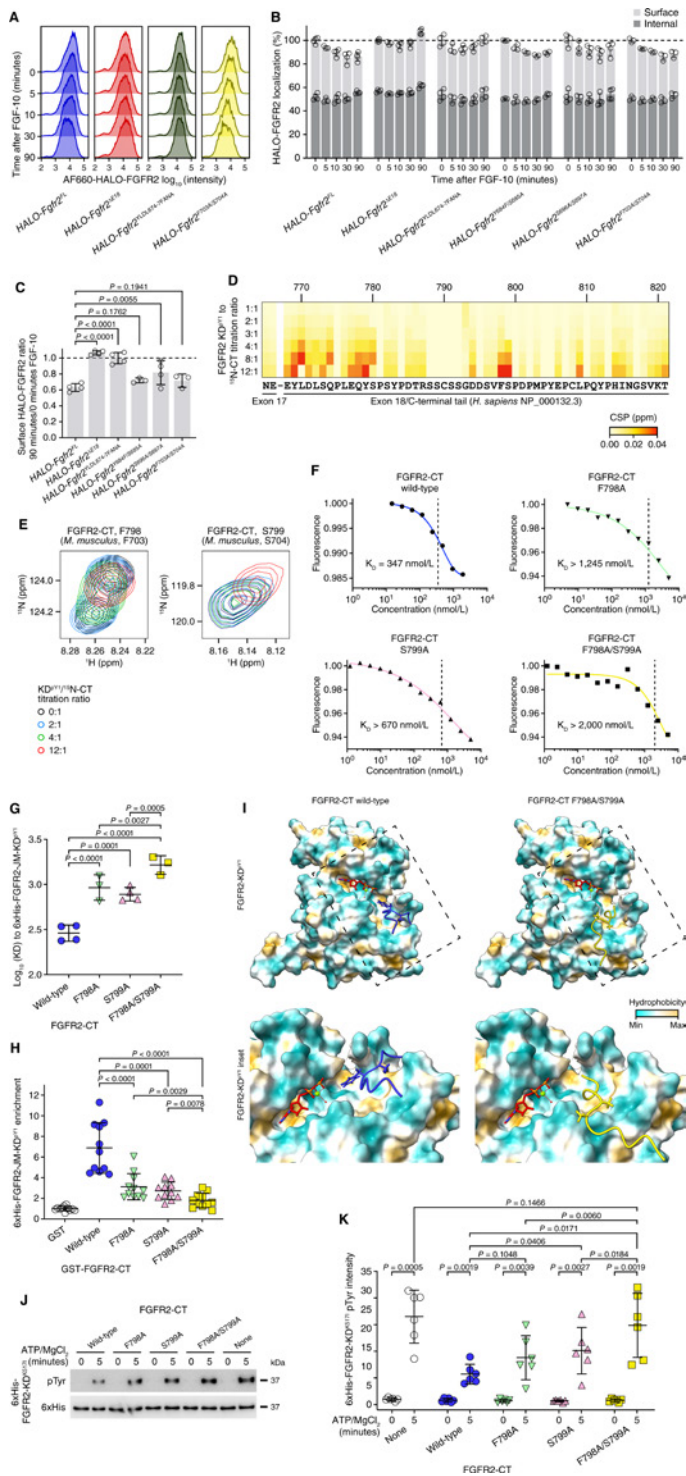


Figure 5. Binding of the FGFR2 C-tail to the kinase domain depends on F703–S704.

A–C, FACS to quantify HALO-FGFR2 internalization in serum-starved NMuMG cells expressing indicated HALO-*Fgfr2* variants using SPIDA [40]. Cells were treated with 20 ng/mL FGF-10 and subsequently labeled first with cell membrane impermeable HTL-AF660 and second with permeable HTL-TMR. AF660-HALO-FGFR2 mean fluorescence intensity (**A**) and TMR-HALO-FGFR2 mean fluorescence intensity were used to calculate surface and internal FGFR2 percentages (**B**) and ratios of surface FGFR2 levels in cells stimulated with FGF-10 for 90 minutes vs. unstimulated cells (**C**). Histograms in **A** show one representative replica of data quantified in **B** and **C**, which are represented as mean $\pm$ SD of HALO-*Fgfr2*^{FL}, HALO-*Fgfr2* ^{Δ E18}, HALO-*Fgfr2*^{YLDL674-TFANA}, and HALO-*Fgfr2*^{Y684F/S685A} at FGF-10 0, 10, 30, and 90 minutes, with $n = 4$; other variants and time points, $n = 3$ independent replicas. **D**, ¹H-¹⁵N HSQC-2D-NMR spectroscopy used to quantify CSPs (parts per million, ppm) of recombinant human ¹⁵N-labeled FGFR2-CT titrated with monophosphorylated FGFR2-KD (KD^{PT}). The heatmap visualizes color-coded CSP values of indicated KD^{PT} to ¹⁵N-CT ratios for each FGFR2-CT amino acid. **E**, Spectral chemical shift maps of FGFR2-CT F798 and S799 from indicated KD^{PT} to ¹⁵N-CT ratios. **F** and **G**, Microscale thermophoresis assays to measure binding capacity of indicated FGFR2-CT to Atto488-labeled 6xHis-FGFR2-JM-KD^{PT}. K_D values were calculated using Richard five parameter dose–response curves and least squares regressions. Dose–response curves in **F** show one representative replica to quantify K_D values in **G**, which are represented as mean $\pm$ SD of FGFR2-CT wild-type and S799A, $n = 4$; F798A, F798A/S799A, $n = 3$ independent replica. **H**, 6xHis-FGFR2-JM-KD^{PT} enrichments after coinubation of 6xHis-FGFR2-JM-KD^{PT} with indicated GST-FGFR2-CTs, followed by GST IP. Quantifications are based on Western blots exemplified in Supplementary Fig. S5G and S5H. Data are represented as mean $\pm$ SD of $n = 11$ independent replicas per group. **I**, Modeling of FGFR2-KD interactions with wild-type (blue) and F798A/S799A-mutant (yellow) FGFR2-CT using AlphaFold2 Multimer and the FGFR2-KD^{PT} crystal structure 6V6Q. The FGFR2-KD surface is color-coded according to a hydrophobicity scale. The side chains of F798/S799 (blue) and A798/A799 (yellow) of the C-tails and the ATP molecule (red) are shown as stick models. **J** and **K**, Western blots (**J**) showing Tyr phosphorylation of kinase-dead 6xHis-FGFR2-KD^{S17} (substrate) coinubated with 6xHis-FGFR2-JM-KD^{PT} (active kinase) and indicated FGFR2-CT variants together with ATP/MgCl₂. 6xHis-FGFR2-KD^{S17} phosphorylation levels were quantified (**K**) based on pTyr band intensities, which were normalized to their lane backgrounds and to the corresponding 6xHis bands. Data are represented as mean $\pm$ SD of $n = 6$ independent replicas per group. P values in **C**, **G**, **H**, and **K** were calculated using a one-tailed one-way ANOVA (**C** and **G**) or a one-tailed one-way repeated measures ANOVA with Geisser–Greenhouse corrections (**H** and **K**) and FDR multiple testing corrections.

The F703–S704 KDBS motif inhibits FGFR2 oncogenicity

To address whether the F703–S704 (human, F798–S799) KDBS motif was required to suppress oncogenicity of FGFR2, we tested *Fgfr2* KDBS motif variants in our *in vitro* signaling pathway transduction and *in vivo* tumor modeling assays (**Fig. 6A**). Among the *Fgfr2* variants tested, the expression of FLAG-*Fgfr2* ^{Δ VFSPD702–6} (lacking the five residues flanking the KDBS motif) and FLAG-*Fgfr2*^{F703A/S704A} in mammary epithelial cells followed by FLAG co-IP showed that both FLAG-FGFR2 ^{Δ VFSPD702–6} and FLAG-FGFR2^{F703A/S704A} were bound by the FGFR adapter proteins EPS8 [43,44], PIK3R1, and STAT3 (**Fig. 6B**). Binding of these effector proteins was considerably increased as compared with FLAG-FGFR2^{FL} and comparable with the high binding observed for FLAG-FGFR2 ^{Δ E18} (**Fig. 6B**), in accordance with the proteomic data (**Figs. 1A, C** and **2B**). Noteworthy, FLAG-FGFR2 ^{Δ VFSPD702–6} and FLAG-FGFR2^{F703A/S704A} were not impaired in binding ERK1/2 and p90 RSK, in contrast to FLAG-FGFR2 ^{Δ E18}, FLAG-FGFR2^{Y684F/S685A}, and FLAG-FGFR2^{S696A/S697A} (**Fig. 6B**; Supplementary Fig. S2A). Accordingly, the expression of *Fgfr2* ^{Δ VFSPD702–6} or *Fgfr2*^{F703A/S704A} activated the MAPK and particularly the PI3K–AKT–mTOR and STAT3 signaling pathways, whereas the expression of the *Fgfr2*^{F703A} and *Fgfr2*^{S704A} single mutants hardly activated downstream signaling pathways except inducing the MAPK pathway (**Fig. 6C** and **D**; Supplementary Fig. S7A and S7B). This was in line with the phospho-RPS6 inductions quantified by FACS (**Fig. 4A–C**). However, none of the *Fgfr2* KDBS motif variants fully recapitulated FGFR2 ^{Δ E18}-induced PI3K–AKT–mTOR signaling and especially STAT3 phosphorylation (**Fig. 6C** and **D**; Supplementary Fig. S7A and S7B).

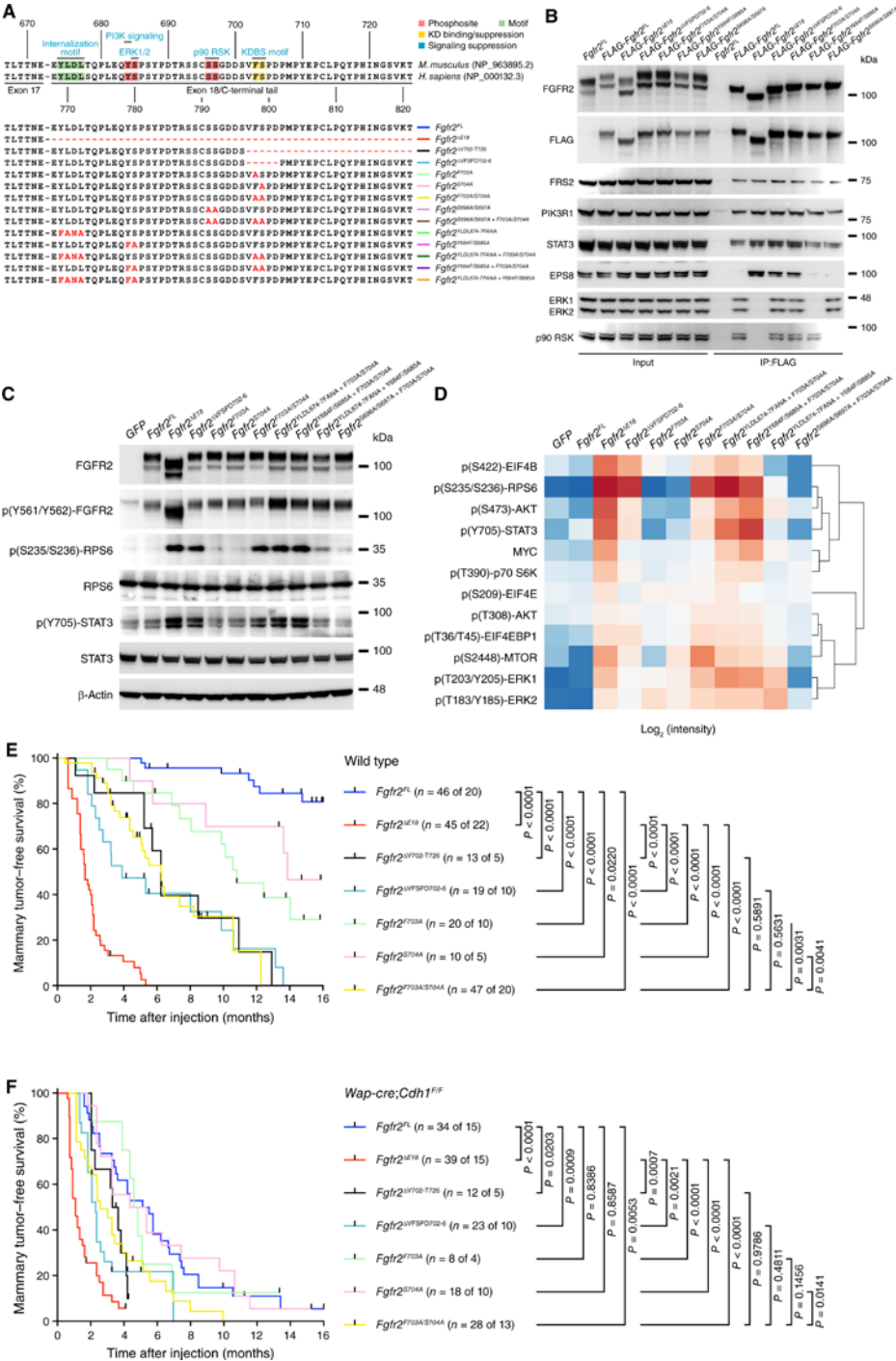


Figure 6. Loss of F703–S704 in the C-tail enables FGFR2 oncogenicity.

A, Graphical representation of the FGFR2-CT amino acid sequence and its previously annotated binding sites and motifs and the KDBS motif identified here. *Fgfr2-E18* variants used to target the C-terminal motifs are indicated. *Fgfr2-E18* variants used to target the C-terminal motifs or combinations thereof are indicated. **B**, Western blots showing enrichment of indicated proteins in protein lysate input and FLAG-IP samples of serum-starved NMuMG cells expressing indicated FLAG-*Fgfr2* variants. Data represent one replica of two independent experiments. **C**, Western blots showing expression and phosphorylation of indicated proteins in serum-starved NMuMG cells expressing indicated *Fgfr2* variants. **D**, Quantifications of Western blots in **C** and Supplementary Fig. S7A and S7B. Band intensities were normalized to their individual lane backgrounds and to the corresponding total protein and β -actin bands. Unsupervised hierarchical clustering was performed on row mean-centered $\log_2$ (intensity) values using the Pearson distance measure and Ward minimum variance clustering. Data in **C** and **D** represent one replica of three independent experiments. **E** and **F**, Kaplan–Meier curves showing mammary tumor-free survival of wild-type (**E**) or *Wap-cre;Cdh1^{F/F}* (**F**) mice intraductally injected with lentiviruses encoding indicated *Fgfr2* variants. Cohort counts (*n*) are injected mammary glands per number of mice. The *Fgfr2^{FL}* and *Fgfr2^{ΔE18}* curves are duplicates from Fig. 2F and I. The *Fgfr2^{ΔV702-726}* curves were previously published [31]. *P* values were calculated using log-rank (Mantel–Cox) tests.

Injection of lentiviral *Fgfr2^{F703A}* or *Fgfr2^{S704A}* into the mammary glands of wild-type or *Wap-cre;Cdh1^{F/F}* mice accelerated tumorigenesis in comparison with glands transduced with *Fgfr2^{FL}* (**Figs. 6E and F**; Supplementary Fig. S7C and S7D). Histopathologic evaluation of the *Fgfr2^{F703A}*- or *Fgfr2^{S704A}*-injected mammary glands revealed mostly FGFR2-positive low-grade lesions (Supplementary Fig. S8A–S8C). This was in contradiction to the inability of these FGFR2 variants to induce downstream signaling *in vitro* (**Figs. 4B and 6C and D**; Supplementary Fig. S7A and S7B). Possibly, the presence of FGF ligands in the mouse mammary gland [2] sufficed to promote oncogenic FGFR2^{F703A} and FGFR2^{S704A} signaling *in vivo*, whereas these variants were incapable of signaling in the starvation conditions used *in vitro*. The injection of *Fgfr2^{F703A/S704A}* markedly shortened mammary tumor onset resulting in reduced tumor-specific survival versus the injection of *Fgfr2^{F703A}*, *Fgfr2^{S704A}*, or *Fgfr2^{FL}* (**Fig. 6E and F**; Supplementary Fig. S7C and S7D). The tumor latencies of wild-type and *Wap-cre;Cdh1^{F/F}* mice injected with *Fgfr2^{F703A/S704A}* were equivalent to the tumor latencies observed for the respective *Fgfr2^{ΔVFS702-6}* and *Fgfr2^{ΔV702-T726}* (distal C-tail ablation) cohorts (**Fig. 6E and F**), and glands transduced with *Fgfr2^{F703A/S704A}*, *Fgfr2^{ΔVFS702-6}*, or *Fgfr2^{ΔV702-T726}* prevalently contained FGFR2-positive high-grade lesions (Supplementary Fig. S8A–S8C). Yet, tumorigenesis did not occur as rapidly as within *Fgfr2^{ΔE18}*-injected mice (**Fig. 6E and F**; Supplementary Fig. S7C and S7D), suggesting that the C-terminal tail exerts its FGFR2-inhibitory function through the F703–S704 KDBS motif in conjunction with other regulatory sites.

Loss of KDBS and proximal C-tail motifs drives FGFR2 oncogenicity

To address whether co-permutation of other C-terminal motifs located before F703 would enhance oncogenicity of the *Fgfr2^{F703A/S704A}* KDBS motif variant, we generated double mutants for the KDBS motif and the p90 RSK binding site (*Fgfr2^{S696A/S697A+F703A/S704A}*), the PI3K/ERK negative feedback loop regulatory site (*Fgfr2^{Y684F/S685A+F703A/S704A}*), or the endocytosis consensus motif (*Fgfr2^{YLD674-7FANA+F703A/S704A}*; **Fig. 7A**). The use of HALO-tagged derivatives of the double mutants revealed impaired FGFR2 internalization upon endocytosis motif targeting, whereas internalization of the other FGFR2 variants was unaffected (Supplementary

Fig. S9A–S9C). Interestingly, loss of the p90 RSK binding site in FGFR2^{S696A/S697A+F703A/S704A} reduced signaling output as compared with cells expressing *Fgfr2*^{F703A/S704A} (Fig. 6C and D; Supplementary Fig. S7A and S7B). Consistent with the impaired signaling activity *in vitro*, mice that were intraductally injected with FGFR2^{S696A/S697A+F703A/S704A} displayed mammary tumor onsets similar to *Fgfr2*^{F703A/S704A}-injected mice (Supplementary Fig. S9D and S9E), and *Fgfr2*^{S696A/S697A+F703A/S704A}-injected animals even showed prolonged survival versus *Fgfr2*^{F703A/S704A} mice (Supplementary Fig. S9F and S9G).

In contrast, preventing endocytosis or ablating the PI3K/ERK negative feedback loop in FGFR2^{YLDL674-7FANA+F703A/S704A} and FGFR2^{Y684F/S685A+F703A/S704A}, respectively, further enhanced the signaling capacity as compared with FGFR2^{F703A/S704A} and resulted in the induction of PI3K–AKT–mTOR and STAT3 pathways to levels observed with FGFR2^{ΔE18} (Fig. 6C and D; Supplementary Fig. S7A and S7B). Wild-type and *Wap-cre;Cdh1^{F/F}* mice that were intraductally injected with lentiviral *Fgfr2*^{YLDL674-7FANA+F703A/S704A} or *Fgfr2*^{Y684F/S685A+F703A/S704A} showed rapid mammary tumor onsets matching those of the respective *Fgfr2*^{ΔE18} cohorts (Fig. 7A–C), and the mice reached tumor-related endpoint equally fast (Supplementary Fig. S9H and S9I). Histopathologic evaluation of the *Fgfr2*^{YLDL674-7FANA+F703A/S704A} and *Fgfr2*^{Y684F/S685A+F703A/S704A} mammary glands revealed FGFR2-positive high-grade tumors representing the spectrum of mammary tumor subtypes driven by *Fgfr2*^{ΔE18}, i.e., adenocarcinomas and metaplastic carcinomas in wild-type mice, invasive lobular carcinomas in *Wap-cre;Cdh1^{F/F}* mice, and sarcomatoid tumors in animals of both genotypes (Supplementary Fig. S8A–S8C). Importantly, co-perturbation of the endocytosis and PI3K/ERK negative feedback loop motifs in *Fgfr2* that retain the F703–S704 KDBS motif (*Fgfr2*^{YLDL674-7FANA+Y684F/S685A}) enhanced MAPK signaling but had only minor effects on the PI3K–AKT–mTOR and STAT3 signaling pathways (Fig. 6C and D; Supplementary Fig. S7A and S7B). In line with this, the *Fgfr2*^{YLDL674-7FANA+Y684F/S685A} double mutant did not display enhanced *in vivo* oncogenicity as compared with the *Fgfr2*^{YLDL674-7FANA} and *Fgfr2*^{Y684F/S685A} single mutants (Fig. 7B and C; Supplementary Fig. S9H and S9I). Hence, the C-terminal tail restrained oncogenicity of FGFR2 through the F703–S704 KDBS motif and the concurrent regulation of endocytosis and the PI3K/ERK negative feedback loop. Consequently, targeting of these regulatory sites revealed FGFR2 oncogenes reminiscent of FGFR2^{ΔE18}.

Figure 7. Loss of F703–S704 KDBS and proximal C-tail motifs drives FGFR2 oncogenicity.

A, Graphical representation of the FGFR2-CT amino acid sequence and its binding sites and motifs. *Fgfr2*-E18 variants used to target the C-terminal motifs or combinations thereof are indicated. **B** and **C**, Kaplan–Meier curves showing mammary tumor-free survival of wild-type (**B**) or *Wap-cre;Cdh1^{F/F}* (**C**) mice intraductally injected with lentiviruses encoding indicated *Fgfr2* variants. Cohort counts (*n*) are injected mammary glands per number of mice. The *Fgfr2*^{F^L}, *Fgfr2*^{ΔE18}, *Fgfr2*^{YLDL674-7FANA}, *Fgfr2*^{Y684F/S685A}, and *Fgfr2*^{F703A/S704A} curves are duplicates from Figs. 2F, G, I, and J, and 5E and F. *P* values were calculated using log-rank (Mantel–Cox) tests. **D**, GR-corrected dose–response curves of NMuMG cells expressing indicated *Fgfr2* variants and treated for 4 days with pemigatinib. Data are represented as mean ± SEM of *n* = 5 replicates per group. **E**, GR₅₀ values for pemigatinib based on dose–response experiments in **D**. Data are represented as mean ± 95% confidence interval. GR dose–response curves (**D**) and GR₅₀ values (**E**) were modeled using [inhibitor] vs. response with variable slope (four parameters), least squares regressions, and asymmetrical confidence intervals (CI). Data in **D** and **E** represent one replica of two independent experiments.



KDBS motif–mutant FGFR2 is sensitive to FGFR inhibitors

Finally, we tested GR-corrected sensitivities of mammary epithelial cells expressing different *Fgfr2-E18* variants to the ATP-competitive FGFR inhibitors pemigatinib and fexagratinib [34,45] by using culture conditions that are permissive to FGFR2-induced cell growth. The expression of *Fgfr2^{ΔE18}* rendered cells highly sensitive to these FGFR inhibitors, whereas *Fgfr2^{FL}*-expressing cells were less sensitive to FGFR2 blockade (**Fig. 7D** and **E**; Supplementary Fig. S9J and S9K). Pemigatinib and fexagratinib GR₅₀ values of *Fgfr2^{FL}*-expressing cells were comparable with those measured in cells expressing *Fgfr2^{YLDL674-7FANA+Y684F/S685A}* that retain the KDBS motif (**Fig. 7E**; Supplementary Fig. S9K). In contrast, the expression of KDBS motif–mutant *Fgfr2^{ΔVFSPD702-6}* and *Fgfr2^{F703A/S704A}* modestly increased GR-corrected FGFR inhibitor sensitivities as compared with *Fgfr2^{FL}* (**Fig. 7D**; Supplementary Fig. S9J). Notably, the expression of *Fgfr2^{YLDL674-7FANA+F703A/S704A}* and *Fgfr2^{Y684F/S685A+F703A/S704A}* that are double mutant for the KDBS motif and the endocytosis consensus motif or the PI3K/ERK negative feedback loop regulatory site further increased sensitivities to pemigatinib and fexagratinib, with GR₅₀ values matching those measured in *Fgfr2^{ΔE18}*-expressing cells (**Fig. 7D** and **E**; Supplementary Fig. S9J and S9K). Together, these studies show that loss of the KDBS motif and further C-terminal regulatory sites generates FGFR2 oncogenes that are sensitive to FGFR inhibitors, thus providing a mechanistic rationale for the actionability of *FGFR2^{ΔE18}*-altered cancers by FGFR-targeted therapies.

DISCUSSION

Here, we reveal the previously unannotated phenylalanine–serine KDBS motif to be central for efficient FGFR2 kinase domain inhibition by its C-terminal tail. The C-tail exerts this function through direct KDBS-mediated binding to the kinase domain (**Fig. 8A**). Conversely, loss of the entire C-tail or the KDBS motif enables FGFR2 oncogenes that activate the PI3K–AKT–mTOR and STAT3 signal transduction pathways (**Fig. 8B**). The signaling activities of RTKs are controlled by autoinhibition [1]. Within the RTK homodimer, the kinase domain activation loop itself (e.g., insulin receptor), the juxtamembrane domain (e.g., KIT), and the C-terminal tail (e.g., EGFR, FGFR2, and MET) can suppress kinase domain activity, typically *in cis* [1,7,46–49]. The C-terminal tail of EGFR stabilizes an unphosphorylated kinase domain dimer and the MET C-tail competitively hinders the access of ATP to the kinase domain, thus both suppressing receptor activity [48,49]. Notably, truncation of EGFR and MET by their C-tails drives pro-tumorigenic activities by releasing autoinhibition of these RTKs [50,51]. However, it remains to be investigated whether EGFR, MET, and other RTKs beyond the FGFR paralogs contain phenylalanine–serine KDBS motifs and whether those are functionally involved in the autoinhibitory regulation of kinase activity.

C-terminal tyrosine and serine residues in the FGFRs are subject to phosphorylation, which in turn fine-tunes receptor activity [12–14,21–28,35,36]. In the FGFR2 C-tail, the

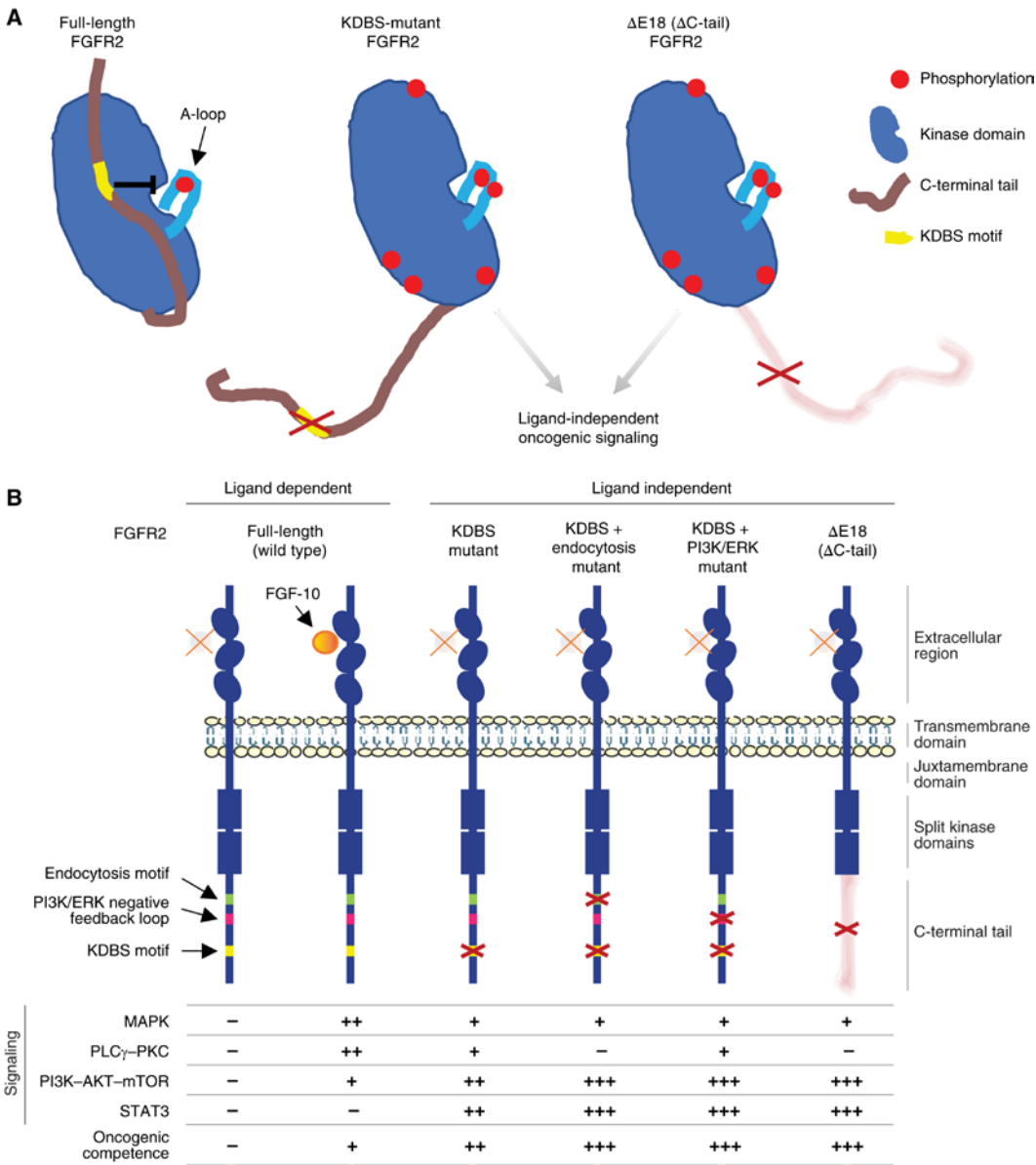


Figure 8. Graphical summary.

A, Graphical representation of the C-terminal tail and its interaction with the kinase domain of FGFR2. In the wild-type FGFR2^{FL} receptor, the C-tail directly binds to the kinase domain via the KDBS motif and suppresses its activity. Upon mutation of the KDBS motif or complete loss of E18, this binding and suppressing capacity is abrogated. **B**, Scoring of the experimentally assessed oncogenic signaling and tumor-forming capacities of FGFR2^{FL}, KDBS-mutant FGFR2 (FGFR2^{F703A/S704A}), FGFR2 mutant for the KDBS and the endocytosis motifs (FGFR2^{YLDL674-7FANA+F703A/S704A}), FGFR2 mutant for the KDBS and the PI3K/ERK negative feedback loop motifs (FGFR2^{Y684F/S685A+F703A/S704A}), and C-terminally truncated FGFR2 (FGFR2 ^{Δ E18}).

serine of the F703–S704 KDBS motif identified here is followed by a proline (P705), which constitutes a serine–proline consensus phosphosite for ERK1/2, CDK1, and additional kinases [52]. Conceivably, the kinase-suppressive function of the F703–S704 KDBS motif might depend on the phosphorylation of S704, possibly mediated by downstream FGFR2 signaling components, such as ERK1/2. However, only combined loss of F703 and S704 was detrimental for suppression of FGFR2 kinase activity, whereas the S704 single mutant displayed mild phenotypes (**Fig. 6**; Supplementary Figs. S7 and S8). Hence, phosphorylated S704 would only exert its FGFR2 inhibitory function in conjunction with F703.

The conformational changes of FGFR2 required for receptor dimerization and trans-autophosphorylation are thought to be directly regulated by the C-terminal tail [7,20], in line with the autoinhibition of other RTKs by their C-terminal regions [1,48,49]. As supported by our data, the direct steric hindrance of FGFR2 by its C-tail may indeed critically rely on the F703–S704 KDBS motif (**Fig. 8A**). Alternatively, the KDBS motif could also indirectly restrain FGFR2 activity through the recruitment of phosphatases or by preventing access of effector proteins. In comparison with FGFR2^{FL}, FGFR2^{ΔE18} indeed displayed reduced binding of the phosphatases PTPRF, PTPRK, PTPRS, and SHP2 (**Fig. 1A and C**). Conversely, FGFR2^{ΔE18} as well as KDBS motif–mutant FGFR2^{ΔVFSPD702-6} and FGFR2^{F703A/S704A} versus FGFR2^{FL} showed elevated binding of PIK3R1 and STAT3 (**Figs. 2B and 7B**), both of which can directly bind to FGFR1 and FGFR3 to mediate PI3K–AKT–mTOR and STAT3 signaling [9–11,22,53]. Interestingly, the direct engagement of PIK3R1 by constitutively active FGFR3 represses MAPK signaling in favor of the PI3K–AKT–mTOR pathway [9], and breast cancer cell lines expressing FGFR2^{ΔE18} variants [31–33] predominantly stimulate and depend on PI3K–AKT–mTOR and STAT3 signaling [11,32,33,54]. FGFR2^{ΔE18}, FGFR2^{ΔVFSPD702-6}, and FGFR2^{F703A/S704A} also showed markedly increased binding to the adapter protein EPS8 as compared with FGFR2^{FL} (**Figs. 2B and 7B**). EPS8 is a direct EGFR phosphotarget that potentiates the signaling output of EGFR and facilitates EGFR-driven cancers [43,55,56]. Notably, EPS8 activity primarily augments the PI3K–AKT–mTOR and STAT3 signaling pathways [56,57], and it is also a direct phosphotarget of FGFRs [43,44]. In line with these findings, we observed that FGFR2^{ΔE18} as well as KDBS motif–mutant FGFR2 prominently signaled through the PI3K–AKT–mTOR and STAT3 pathways (**Fig. 8B**). Taken together, the FGFR2 C-terminal tail in conjunction with the phenylalanine–serine KDBS motif might prevent the recruitment of the PIK3R1, STAT3, and/or EPS8 effector proteins, whereas the oncogenic PI3K–AKT–mTOR and STAT3 signal transductions of FGFR2^{ΔE18} might rely on the aberrant engagement of these adapters.

Patients with cancers driven by oncogenic FGFR2^{ΔE18} variants display high objective response rates to FGFR-targeted therapies [4,30,31]. However, primary and acquired resistance to FGFR inhibitors inevitably occurs, frequently through the acquisition of FGFR2 on-target mutations [31,58,59]. Notably, the dependence of FGFR2^{ΔE18} on the PI3K–AKT–mTOR and STAT3 signaling pathways could be exploited therapeutically. Combination

therapies targeting FGFR2 together with PI3K–AKT–mTOR or STAT3 might circumvent FGFR2 on-target resistance and thereby prolong the clinical responses in patients with FGFR2^{ΔE18}-positive cancer [59,60].

MATERIALS AND METHODS

Recombinant protein expression and purification

The pET-28b(+) and pGEX-4T-1 vectors encoding a 6xHis-tagged human FGFR2 (NP_075259.4) kinase-dead kinase domain (6xHis-FGFR2-KD^{K517I}, amino acid residues 466–763) or a monophosphorylated juxtamembrane-kinase domain (6xHis-FGFR2-JM-KD^{pY1}, amino acid residues 413–763) and a GST-tagged C-terminal tail polypeptide (GST-FGFR2-CT, amino acid residues 764–821), respectively, and site-directed mutagenesis thereof, were previously described [7,14,17]. FGFR2-JM-KD^{pY1} contained phenylalanine conversions for phospho-relevant tyrosine residues, except for the activation loop Y656, and upon *in vitro* autophosphorylation, represented a JM-KD activation state corresponding to basal conditions among the sequential phosphorylation steps [7,41]. 6xHis-FGFR2-KD^{K517I}, 6xHis-FGFR2-JM-KD^{pY1}, and GST-FGFR2-CT variants were expressed in BL21(DE3)-competent *Escherichia coli* (C2527H, New England Biolabs), and proteins were purified as previously described [7,14,16]. Briefly, LB cultures were grown at 37°C to an optical density at 600 nm of 0.8 and, following induction of protein expression with 0.5 mmol/L isopropyl β-D-1-thiogalactopyranoside, further grown for 12 hours at 20°C. *E. coli* were harvested by centrifugation and lysed by sonication [7,14], insoluble material was removed by centrifugation, and the soluble fractions were applied to HiTrap TALON crude columns (28953809, Cytiva) or GSTrap HP Columns (17528205, Cytiva) for purification. Columns were subjected to washing buffer [14] and proteins were eluted with 6xHis tag-specific or GST tag-specific elution buffers [14].

To completely dephosphorylate purified 6xHis-FGFR2-KD^{K517I}, calf intestinal alkaline phosphatase (CIP; M0290, New England Biolabs) was conjugated to UltraLink Biosupport beads (53110, Thermo Fisher Scientific), and CIP beads were incubated with 6xHis-FGFR2-KD^{K517I} for 4 hours at room temperature as previously described [7,14]. After dephosphorylation, protein solution and CIP beads were separated by centrifugation. Partially purified proteins were concentrated and purified with a gel filtration CD10/40 Column (19-5037-01, Cytiva) that was self-packed with Superdex 75 Prep Grade size exclusion chromatography resin (17104402, Cytiva). To generate monophosphorylated 6xHis-FGFR2-JM-KD^{pY1}, protein was autophosphorylated by incubating with 5 mmol/L ATP (A1852, Sigma-Aldrich) and 10 mmol/L MgCl₂ for 2 hours at room temperature, reactions were quenched with 100 mmol/L EDTA, and 6xHis-FGFR2-JM-KD^{pY1} was purified using a gel filtration column containing Superdex 75 Prep Grade resin as previously described [7,14]. All recombinant proteins were prepared in a gel filtration buffer (20 mmol/L HEPES,

150 mmol/L NaCl, and 1 mmol/L Tris (2-carboxy-ethyl)-phosphine-HCl, pH 7.5) for further application.

SDS-PAGE was used to verify protein purities >98%. GST-tagged FGFR2-CT variants were re-captured to GST beads for GST pulldown experiments. Where required, GST was cleaved with 1 unit thrombin (T6884, Sigma-Aldrich) per 1 mg recombinant protein overnight at 4°C. Benzamidine Sepharose 4 Fast Flow (high sub) beads (17512302, Cytiva) were used to remove thrombin, and untagged FGFR2-CT variants were purified using GStap HP Columns.

For ^{15}N single labeling or ^{13}C - ^{15}N double labeling of the FGFR2-CT, *E. coli* were grown in minimal medium supplemented with $^{15}\text{NH}_4\text{Cl}$ (99%, NLM-467-PK, Cambridge Isotope Laboratories) and ^{13}C -labeled D-glucose (U- $^{13}\text{C}6$, 99%, CLM-1396-PK, Cambridge Isotope Laboratories) as previously described [7,14].

***In vitro* FGFR2 kinase assay**

For this assay, 0.1 $\mu\text{mol/L}$ active kinase (6xHis-FGFR2-JM-KD $^{\text{PY1}}$) and 4 $\mu\text{mol/L}$ dephosphorylated kinase-dead substrate (6xHis-FGFR2-KD $^{\text{K517I}}$) were mixed (1:40 ratio). Then 1 $\mu\text{mol/L}$ FGFR2-CT was added, and the mixture was incubated with 5 mmol/L ATP and 10 mmol/L MgCl_2 for 5 minutes at room temperature and quenched with 100 mmol/L EDTA. The phosphorylation level of the 6xHis-FGFR2-KD $^{\text{K517I}}$ substrate was analyzed using immunoblotting.

Microscale thermophoresis assay

Human recombinant 6xHis-FGFR2-JM-KD $^{\text{PY1}}$ was fluorescently labeled with Atto 488 NHS ester (41698, Sigma-Aldrich) according to the manufacturer's protocol using a previously optimized 1:1 protein to Atto 488 ratio [7,14]. Labeled protein was buffer-exchanged into the gel filtration buffer with 0.01% Tween 20. Dilution series of unlabeled human recombinant FGFR2-CT variants were prepared in a gel filtration buffer and mixed with Atto 488-labeled 6xHis-FGFR2-JM-KD $^{\text{PY1}}$ at a 1:1 ratio [7,14]. Samples were loaded into Monolith Premium Capillaries (MO-K025, NanoTemper Technologies) and binding affinities were measured at 25°C using a Monolith NT.115 (NanoTemper Technologies). Data were analyzed with MO.Affinity Analysis Software (v.1.2.101, NanoTemper Technologies) and equilibrium dissociation constants (K_D) were calculated with Richard five parameter dose-response curves and least squares regressions using Prism (v.9.4.1, GraphPad Software).

Nuclear magnetic resonance spectroscopy

The acquisition of the ^1H - ^{15}N heteronuclear single quantum coherence two-dimensional nuclear magnetic resonance (HSQC-2D-NMR) spectra was previously described [7]. Briefly, HSQC-2D-NMR spectroscopy was performed using a 750 MHz/89-mm Bruker Avance III HD NMR spectrometer equipped with a ^1H -optimized triple-resonance NMR 5-mm TCI

CryoProbe (Bruker) and TopSpin software (v.4.0). Human recombinant ^1H - ^{13}C - ^{15}N -FGFR2-CT was concentrated to 300 $\mu\text{mol/L}$ in HEPES buffer (20 mmol/L HEPES, 150 mmol/L NaCl, and 1 mmol/L Tris (2-carboxy-ethyl)-phosphine-HCl, pH 7.5) and backbone resonance assignment spectra were recorded at 25°C. To collect high-resolution spectra, three-dimensional backbone resonance assignment experiments using standard Bruker library pulse sequences or BEST versions of amide TROSY pulse sequences were conducted by applying nonuniform sampling (20%–30%). To quantify ^1H - ^{15}N -FGFR2-CT chemical shift perturbations (CSP) upon interference with unlabeled FGFR2-KD^{py1}, 300 $\mu\text{mol/L}$ ^1H - ^{15}N -FGFR2-CT in HEPES buffer was titrated with FGFR2-KD^{py1} using a molar ratio range of 0.5:1 to 12:1. Amide spectra were recorded at 25°C using an hsqcetf3gpsi pulse sequence from the Bruker library. NMR data were processed using NMRPipe (v.10.8; <https://www.ibbr.umd.edu/nmrpipe/>) and further analyzed with the CcpNmr Analysis software package (v.2.4; <https://ccpn.ac.uk/software/version-2/>). CSPs were calculated as previously described [7].

Lentiviral vectors and virus production

The SIN.LV.SF-Fgfr2^{FL}, SIN.LV.SF-Fgfr2^{ΔE18}, SIN.LV.SF-Fgfr2^{FL}-T2A-Puro, SIN.LV.SF-Fgfr2^{ΔE18}-T2A-Puro, SIN.LV.SF-GFP, and SIN.LV.SF-GFP-T2A-Puro lentiviral vectors were previously described [31,38]. The 3xFLAG and HALO tag sequences including murine Igk signal and linker sequences are disclosed in Supplementary Materials and Methods and were purchased as custom-synthesized gBlocks Gene Fragments (Integrated DNA Technologies). The GFP and (parts of) the full-length *Fgfr2* (*Fgfr2*^{FL}) or *Fgfr2*^{ΔE18} sequences, the 3xFLAG and HALO gBlocks Gene Fragments, and the lentivector backbones were assembled using PCR amplification of backbone and insert and the In-Fusion HD Cloning Plus kit (638911, Takara Bio) according to the manufacturer's recommendations. Point mutations and small deletions were introduced to *Fgfr2*^{FL} or *Fgfr2*^{ΔE18} using the QuikChange Lightning Site-Directed Mutagenesis Kit (210519, Agilent). In-Fusion and site-directed mutagenesis primers were designed with SnapGene (v.6.1.2, Dotmatics) and QuikChange Primer Design (Agilent), respectively. All lentivectors were verified using Sanger sequencing. Concentrated lentiviral stocks were produced in HEK293T cells as previously described [31,38], and viral titers were determined using the qPCR Lentivirus Titer Kit (LV900, Applied Biological Materials).

Cell culture

HEK293T cells (CRL-3216, ATCC, obtained in Q1 2015) were cultured in Iscove's modified Dulbecco's medium (31980, Thermo Fisher Scientific) containing 10% FBS (S-FBS-EU-015, Serana) and 100 IU/mL penicillin/streptomycin (15070, Thermo Fisher Scientific). NMuMG cells (CRL-1636, ATCC, obtained in Q3 2011) were cultured in DMEM/F-12 (31331, Thermo Fisher Scientific) containing 10% FBS and 100 IU/mL penicillin/streptomycin. The cell lines used were previously authenticated by ATCC, but no re-authentication was carried out for this study. To stably express lentiviral GFP or Fgfr2-GFP constructs, NMuMG cells (passage 7–12 after purchase) were transduced with lentiviral supernatants at equal TU/

mL in the presence of 8 µg/mL Polybrene (H9268, Sigma-Aldrich) for 72 hours. To stably express lentiviral GFP-T2A-Puro, Fgfr2-T2A-Puro, 3xFLAG-Fgfr2-T2A-Puro, or HALO-Fgfr2-T2A-Puro constructs, NMuMG cells (passage 15–20 after purchase) were transduced with lentiviral supernatants at equal TU/mL in the presence of 8 µg/mL Polybrene for 24 hours, selected with 2 µg/mL puromycin (A11138, Thermo Fisher Scientific) for 5 days, and subsequently grown in DMEM/F-12 containing 10% FBS, 100 IU/mL penicillin/streptomycin, 1 µg/mL puromycin, and 10 µmol/L Y-27632 (M1817, AbMole BioScience). All cell lines were cultured in standard incubators at 37°C with 5% CO₂ and routinely tested for *Mycoplasma* contamination using the MycoAlert Mycoplasma Detection Kit (LT07-218, Lonza). Experiments were carried out with cells cultured for a maximum of 10 passages after lentiviral transduction.

Dose–response curves

NMuMG cells expressing *Fgfr2-T2A-Puro* variants were seeded in 96-well plates (2,000 cells per well) using DMEM/F-12 containing 1% FBS and 100 IU/mL penicillin/streptomycin. After 24 hours, the cells were treated with pemigatinib (HY-109099, MedChemExpress) or fexagratinib (HY-13330, MedChemExpress) with a concentration range of 0.1 nmol/L to 50 µmol/L for 4 days. Cell viability was measured 4 days after treatment start using the CellTiter-Blue Cell Viability Assay (G808A, Promega) for 4 hours and an Infinite M Plex plate reader (Tecan) operated with Tecan i-control software (v.3.9.1). Growth rate (GR) inhibition values were calculated as described [61]. GR dose–response curves and half maximum GR inhibitory concentration (GR₅₀) values were modeled using inhibitor versus response with variable slope (four parameters) and least squares regressions in Prism.

Preparative FACS of GFP^{hi} NMuMG cells

NMuMG cells expressing *Fgfr2-GFP* variants were cultured in DMEM/F-12 medium containing 10% FBS and 100 IU/mL penicillin/streptomycin for 6 days. The cells were collected with 2 mmol/L EDTA, resuspended in DMEM/F-12 medium containing 1% FBS and 2 mmol/L EDTA, and passed through a 70-µm cell strainer prewetted with DMEM/F-12 medium containing 1% FBS and 2 mmol/L EDTA. The cells were counterstained with propidium iodide ReadyProbes Reagent (R37108, Thermo Fisher Scientific) before FACS. The top 50% of all GFP-positive cells (GFP^{hi} cells) were isolated by FACS at 45 psi sheath pressure using a BD FACSAria II Cell Sorter (BD Biosciences) equipped with BD FACSDiva Software (v.8.0.1), an 85-µm nozzle, and 488 nm (530/30 pass filters) and 561 nm (610/20 pass filters) lasers to measure GFP and propidium iodide. Purities of post-sorted GFP^{hi} cell populations were verified by FACS and were typically >97% GFP positive. GFP^{hi} cells were collected in DMEM/F-12 containing 10% FBS and 100 IU/mL penicillin/streptomycin and allowed to recover for 6 days before subsection to follow-up experiments.

Protein isolation from cells

GFP^{hi} NMuMG cells expressing *Fgfr2-GFP* variants or NMuMG cells expressing *Fgfr2-T2A-Puro* variants were cultured in DMEM/F-12 starvation medium (0% FBS) for 48 hours and, where

indicated, treated with recombinant murine FGF-10 (450-61, PeproTech) with indicated doses and times and/or 100 nmol/L fexagratinib for 3 hours. The cells were lysed in RIPA Lysis and Extraction Buffer (89900, Thermo Fisher Scientific) containing Halt Protease and Phosphatase Inhibitor Cocktail (78440, Thermo Fisher Scientific). Protein concentrations were determined with the BCA Protein Assay Kit (23227, Thermo Fisher Scientific) and an Infinite M Plex plate reader.

Immunoprecipitation of recombinant protein and cell lysates

For GST immunoprecipitation (IP), 1 μ mol/L 6xHis-tagged FGFR2-JM-KD^{py1} and 1 μ mol/L GST-tagged FGFR2-CT variants were prepared in 500 mL HEPES buffer. The protein solutions were mixed with 50 mL of Glutathione Sepharose 4B GST-tagged protein purification resin (17075604, Cytiva) and incubated overnight at 4°C following the manufacturer's recommendations. The beads were washed five times using lysis buffer [14], and the samples were eluted from the beads using 2 $\times$ Laemmli Sample Buffer (S3401, Sigma-Aldrich) for 5 minutes at 95°C.

For FLAG IP, NMuMG cells expressing 3xFLAG-Fgfr2-T2A-Puro variants were cultured in DMEM/F-12 starvation medium for 48 hours. The cells were lysed in IP Lysis Buffer (87788, Thermo Fisher Scientific) containing Halt Protease and Phosphatase Inhibitor Cocktail, and protein concentrations were determined with the BCA Protein Assay Kit and an Infinite M Plex plate reader. Equal amounts of protein per sample were incubated with Anti-DYKDDDDK Magnetic Agarose beads (A36797, Thermo Fisher Scientific) overnight at 4°C following the manufacturer's recommendations. The beads were washed using IP Lysis Buffer and a magnetic stand, and the samples were eluted from the beads using NuPAGE Sample Reducing Agent (NP0009, Thermo Fisher Scientific) and NuPAGE LDS Sample Buffer (NP0008, Thermo Fisher Scientific) for 5 minutes at 90°C.

For GFP IP, GFP^{hi} NMuMG cells expressing Fgfr2-GFP variants were cultured in 15-cm dishes in DMEM/F-12 starvation medium for 48 hours, reaching 80% confluency. Cells were lysed in NP-40 lysis buffer (50 mmol/L Tris-HCl pH 8.0, 150 mmol/L NaCl, 10% glycerol, and 1% NP-40). Protein lysates were incubated with ChromoTek GFP-Trap Agarose beads (gta-20, Proteintech) overnight at 4°C following the manufacturer's recommendations. The beads were washed five times using NP-40 lysis buffer, and the samples were eluted from the beads using NuPAGE LDS Sample Buffer for 7 minutes at 95°C.

Western blotting

For human recombinant proteins, equal volumes of kinase assay or IP samples and the Precision Plus Protein All Blue Prestained Protein Standards (1610373, Bio-Rad) were separated on 4% to 20% Mini-PROTEAN TGX Precast Protein Gels (4561094 and 4561096, Bio-Rad) using Tris-Glycine-SDS running buffer [14] and transferred onto iBlot polyvinylidene difluoride membranes (IB24002, Thermo Fisher Scientific) using the iBlot 2 Gel Transfer Device (IB21001, Thermo Fisher Scientific). The membranes were blocked in 5%

BSA (A8022, Sigma-Aldrich) in TBS with 0.05% Tween 20 (TBS-T) and incubated with primary antibodies in 5% BSA in TBST overnight at 4°C. The membranes were washed with TBST and incubated with secondary antibodies in 5% BSA in TBST for 1 hour at room temperature. The membranes were washed in TBS-T and developed using ECL Western Blotting Substrate (32106, Thermo Fisher Scientific). The membranes were imaged with the G:Box Chemi XX6 gel doc system (Syngene). For sequential immunoblotting, the membranes were stripped using Restore Western Blot Stripping Buffer (21059, Thermo Fisher Scientific).

For whole-cell lysates and IP input samples, equal amounts of protein were adjusted to equal volumes using RIPA Lysis and Extraction Buffer and IP Lysis Buffer, respectively, and boiled with NuPAGE Sample Reducing Agent and NuPAGE LDS Sample Buffer for 5 minutes at 90°C. Equal volumes of whole-cell lysates, IP input samples, and IP samples and the BlueEye Prestained Protein Marker (PS-104, Jena Bioscience) were separated on NuPAGE 4% to 12% Bis-Tris Mini Protein Gels (NP0323 and NP0329, Thermo Fisher Scientific) using NuPAGE MOPS SDS Running Buffer (NP0001, Thermo Fisher Scientific) and transferred overnight at 4°C onto nitrocellulose membranes (88018, Thermo Fisher Scientific) in transfer buffer [31]. The membranes were stained with Ponceau S Solution (ab270042, Abcam) to verify equal loading of total protein. The membranes were blocked in 5% BSA in PBS with 0.05% Tween 20 (PBST) and incubated with primary antibodies in 5% BSA in PBST overnight at 4°C, except for the EPS8 and phospho(T308)-AKT antibodies, for which 5% non-fat milk powder in TBST was used. Sister membranes were stained with β -actin antibody as a sample processing control. The membranes were washed with PBST or TBST and incubated with secondary antibodies in 5% BSA in PBST or 5% milk in TBS-T for 1 hour at room temperature. The membranes were washed and developed using SuperSignal West Pico PLUS Chemiluminescent Substrate or Femto Maximum Sensitivity Substrate (34580 and 34095, Thermo Fisher Scientific). The membranes were imaged with Fusion FX (Vilber) operated with the Fusion FX7 Edge imaging system (v.18.05).

All antibodies used are listed in Supplementary Table S1. Images were processed with Photoshop 2023 (v.24.0.0, Adobe) using “input levels” and “output levels” to adjust brightness, and band intensities were measured with “mean gray value” in Fiji (v.2.9.0; <https://imagej.net/software/fiji/>). Band intensities were normalized to their individual lane backgrounds and to the corresponding β -actin bands, where applicable. Phosphoprotein band intensities were further normalized to the corresponding total protein band intensities. Unsupervised hierarchical clustering of relative (phospho)-protein intensities was performed using ClustVis (v.2.0; <https://biit.cs.ut.ee/clustvis/>).

Mass spectrometry-based proteomics

GFP IP eluates were run into the stacking of NuPAGE 4% to 12% Bis-Tris Mini Protein Gels. The gels were stained with Pierce Coomassie Brilliant Blue G-250 Dye (20278, Thermo Fisher Scientific), and Coomassie-stained bands were excised. Proteins were reduced with

6.5 mmol/L dithiothreitol, alkylated with 54 mmol/L iodoacetamide, and digested in-gel with 3 ng/ μ L Mass Spectrometry–Grade Trypsin Gold (V5280, Promega) overnight at 37°C. Extracted peptides were vacuum dried, reconstituted in 10% formic acid, and analyzed by LC/MS-MS using a Proxeon/Thermo EASY-nLC 1000 system coupled to an Orbitrap Fusion Tribrid Mass Spectrometer (Thermo Fisher Scientific). Peptides were loaded directly onto a Poroshell 120 EC-C18 Threaded Column (2.7 μ m, 50 μ m $\times$ 500 mm, Agilent) packed in-house and separated in a 90-minute gradient containing a 74-minute linear increase from 6% to 30% solvent B (0.1% formic acid and 80% acetonitrile) with solvent A (0.1% formic acid in water). Raw data files were processed with Proteome Discoverer (v.1.4.1.14 or v.2.2.0.388, Thermo Fisher Scientific) using default settings. Tandem mass spectrometry spectra were searched against the *Mus musculus* Swiss-Prot database (16,714 entries, release 2015/04) using the Mascot search engine (v.2.5 or v.2.6.1, Matrix Science), with carbamidomethylation and oxidation set as fixed and variable modifications, respectively, and with a trypsin cleavage specificity of maximum two missed cleavages. A decoy database defined in the Percolator node of Proteome Discoverer was used to validate and filter peptide–spectrum matches (PSM) using an FDR < 0.01 at peptide and protein levels. The PSM data from independent experiments were combined by Proteome Discoverer into multi-consensus reports, which were exported for determination of potential FGFR2 protein interactors. Proteins with fold change (FC) ≥ 3 and FDR-adjusted *P* values < 0.05 for *Fgfr2^{FL}-GFP* PSMs versus *GFP* PSMs and/or *Fgfr2^{ΔE18}-GFP* PSMs versus *GFP* PSMs were considered enriched. Proteins differentially enriched in *Fgfr2^{ΔE18}-GFP* versus *Fgfr2^{FL}-GFP* were defined as $|\log_2(\text{FC})| \geq 0.58$ and FDR-adjusted *P* values < 0.05. Principal component analysis and unsupervised hierarchical clustering of enriched proteins were performed using $\log_2(\text{PSM} + 1)$ data transformation and ClustVis. Gene set overrepresentation analysis for differentially enriched proteins was performed using gene ontology (GO) biological processes [62] and the PANTHER Classification System (v.17.0; <https://pantherdb.org/>) with Fisher exact testing and Bonferroni multiple testing corrections.

Analytic FACS

To measure RPS6 phosphorylation intensity, NMuMG cells expressing *Fgfr2-GFP* variants were cultured in DMEM/F-12 starvation medium for 24 hours, collected with 2 mmol/L EDTA, and passed through a 70- μ m cell strainer prewetted with FACS buffer (PBS with 10% FBS and 2 mmol/L EDTA). The cells were fixed with BD Phosflow Fix Buffer I (557870, BD Biosciences) and permeabilized with BD Phosflow Perm Buffer III (558050, BD Biosciences) each for 30 minutes at 4°C. The cells were incubated with phospho(S235/S236)-RPS6 primary antibody and subsequently with Alexa Fluor 647 (AF647) secondary antibody each in FACS buffer for 1 hour at 4°C. Details on antibodies used are in Supplementary Table S1.

To measure FGFR2 internalization, NMuMG cells expressing *HALO-Fgfr2-T2A-Puro* variants were cultured in DMEM/F-12 starvation medium for 48 hours and treated with 20 ng/mL recombinant murine FGF-10 for indicated times. Subsequently, the cells were sequentially

labeled with 350 nmol/L cell membrane-impermeable HaloTag ligand (HTL)-AF660 (G8471, Promega) and 250 nmol/L cell membrane-permeable HTL-TMR (G8251, Promega), each for 15 minutes at 37°C, following the surface protein internalization and degradation assay (SPIDA; ref. [40]). The cells were collected with 2 mmol/L EDTA and passed through a 70- μ m cell strainer prewetted with FACS buffer. Single cells were labeled with the LIVE/DEAD Fixable Violet Dead Cell Stain Kit (405 nm excitation, L34964, Thermo Fisher Scientific) and fixed with BD Phosflow Fix Buffer I each for 30 minutes at 4°C.

FACS was performed with a BD LSRFortessa Cell Analyzer (BD Biosciences) equipped with BD FACSDiva Software (v.8.0.2) and with 405 nm (450/50 pass filters), 488 nm (530/30 pass filters), 561 nm (586/15 pass filters), and 638 nm (670/30 pass filters) lasers to measure either GFP and AF647-phospho(S235/S236)-RPS6 or 405-Live/Dead, TMR-HALO-FGFR2, and AF660-HALO-FGFR2. Data were analyzed with FlowJo (v.10.7.1, BD Biosciences) and Prism, and cell surface and internalized FGFR2 percentages were calculated according to the SPIDA methodology [40].

RNA sequencing analysis

GFP^{hi} NMuMG cells expressing *Fgfr2-GFP* variants were cultured in DMEM/F-12 starvation medium for 24 hours and lysed in Buffer RLT (79216, Qiagen) containing 1% 2-mercaptoethanol. Total RNA extraction was performed using the RNeasy Mini Kit (74104, Qiagen). RNA quantity and quality were assessed using the 2100 Bioanalyzer system and a Nano chip (Agilent). RNA samples with RNA integrity number >8 were subjected to polyA-stranded library preparation using the TruSeq RNA Library Prep Kit v2 (RS-122-2001/2, Illumina), quality-checked with the 2100 Bioanalyzer system using a 7500 chip, and pooled equimolar into a 10 nmol/L sequencing stock solution. Libraries were sequenced with 65 bp single reads using the HiSeq 2500 System (Illumina) with V4 chemistry and operated by HiSeq Control Software (v.2.2.68). RNA sequencing (RNA-seq) profiles were analyzed as previously described [31] using R packages from Bioconductor (<https://bioconductor.org/packages/3.20/bioc/>). Briefly, sequencing reads were mapped to the mouse reference genome GRCm38 (mm10) using STAR (v.2.6.0) and gene and expression read counts were quantified by featureCounts (v.1.5.0) based on gene structures defined in GRCm38. Genes with counts per million values higher than one in at least 20% of the total number of samples were considered expressed and used for downstream analysis. Read counts for expressed genes were normalized by the trimmed mean of M-value method using edgeR (v.3.26.6). Differentially expressed genes were identified using Limma-voom (<https://ucdavis-bioinformatics-training.github.io/2018-June-RNA-Seq-Workshop/thursday/DE.html>) with cutoff criteria defined as $|\log_2(\text{FC})| > 0.58$ and Benjamini-Hochberg-adjusted *P* values < 0.05. Gene set overrepresentation analyses for differentially expressed genes were performed based on Kyoto Encyclopedia of Genes and Genomes pathways [63], Hallmark gene sets [64], and GO biological processes [62] using Fisher exact testing with FDR multiple testing corrections. The code used to analyze the RNA-seq data is accessible via Code Ocean (<https://codeocean.com/capsule/5041649/tree/v1>).

Mouse models

The FVB/NRj *Wap-cre;Cdh1^{+/F}* mouse strain (MGI:6296605) was maintained at the Netherlands Cancer Institute animal facility and PCR-genotyped as previously described [31,37]. To somatically model *Fgfr2* variants in the mouse mammary gland, 6-week-old FVB/NRj wild-type (IMSR_RJ:FVB, Janvier Labs) or *Wap-cre;Cdh1^{+/F}* female mice were intraductally injected with lentiviruses encoding *Fgfr2* variants as previously described [31,38]. Briefly, 20 μ L of high-titer lentiviruses (10^8 – 10^9 TU/mL) was injected into the fourth and/or the third mammary glands using a 34G needle. The somatic model cohorts were monitored twice weekly and mammary tumor-free survival was scored (event) for each injected mammary gland individually when palpable tumors were detected. Tumor volume was measured in two dimensions using a caliper with the formula $\text{volume} = \text{length} \times \text{width}^2 \times 0.5$. Mammary tumor-specific survival was scored when a single mammary tumor burden reached a volume of 1,500 mm³, the total mammary tumor burden reached a volume of 2,000 mm³, or the mice suffered from tumor-related clinical signs of distress, such as respiratory distress, ascites, distended abdomen, rapid weight loss, and severe anemia. The maximal permitted disease endpoints were not exceeded in any of the cohorts. Mammary glands were collected and analyzed for histologic abnormalities. Cohort sizes were determined using a sample size calculator (<http://powerandsamplesize.com/Calculators>) and were large enough to measure the effect sizes. Tumor measurements were performed in a blinded manner. The mouse colony was housed in a certified animal facility with a 12-hour light/dark cycle in a temperature- and humidity-controlled room set to 21°C and 55% relative humidity. Mice were kept in individually ventilated cages, and food and water were provided *ad libitum*. All animal experiments were approved by the Animal Ethics Committee of the Netherlands Cancer Institute and the Institutional Animal Care and Use Committee (DEC) of the Dutch government under the Central Animal Testing Committee license number 9 (CCD 9, AVD3010020172464, appendix 2). All animal experiments were performed in accordance with institutional and national guidelines for animal care and use and the European Animal Research Association regulations on animal research.

Histology and IHC

Formalin-fixed, paraffin-embedded mammary gland tissue was subjected to hematoxylin and eosin (H&E) histochemical and FGFR2 IHC stainings using previously described routine procedures [31]. For the FGFR2 staining, antigen retrieval was done with Citrate Buffer (CBB999, ScyTek) at pH 6. Sections were incubated with FGFR2 primary antibody (Supplementary Table S1) overnight at 4°C, labeled with the EnVision+ System HRP Labeled Polymer Anti-Rabbit (K4003, Dako), visualized with the Liquid DAB+ Substrate Chromogen System (K3468, Dako), and counterstained with hematoxylin. Slides were digitally processed using a Panoramic 1000 whole-slide scanner (3DHISTECH) and captured with CaseViewer software (v.2.2.1, 3DHISTECH). H&E slides were used to classify mammary tumor lesion types according to the international consensus of mammary pathology [65]. FGFR2 stainings were quantified by evaluating tumor cell-specific positivity using default histoscore. All slides were reviewed and quantified in a blinded manner.

Protein sequence alignments

Amino acid sequences of the C-terminal tails of FGFR2 orthologs and human FGFR paralogs were aligned using the NCBI Constraint-Based Multiple Alignment Tool (https://www.ncbi.nlm.nih.gov/tools/cobalt/re_cobalt.cgi) with default settings.

AlphaFold2 Multimer modeling

To model potential interactions between the FGFR2 kinase domain and C-terminal tail variants, the human FGFR2 (NP_000132.3) kinase domain amino acid sequence (T467-S792) and part of the human C-tail sequence (G793-T821) or its F798A-S799A double-mutant version (G793-F798A/S799A-T821) were subjected to computational simulations using AlphaFold2 Multimer [42]. The simulations were executed on a high-performance computing cluster with graphics processing unit acceleration utilizing AlphaFold (v.2.3.0) with CUDA (v.11.4.1, NVIDIA Developer). The “multimer” model preset was selected. Model assessment involved the computation of 25 models for each FGFR2-CT variant. The selection of the best-ranked models was determined based on the “iptm + ptm” scoring function. The ATP molecule was superimposed onto the FGFR2-KD-CT models based on the human FGFR2-KD^{py1} crystal structure (PDB 6V6Q; ref. [7]). The hydrophobicity scale of the FGFR2-KD surface was visualized using UCSF ChimeraX (v.1.6; <https://www.cgl.ucsf.edu/chimerax/>).

FGFR2 missense mutation analysis

The occurrence of *FGFR2* missense mutations in cancer was assessed using two independent oncogenomic datasets from Foundation Medicine [31] and cBioPortal [29]. The Foundation Medicine dataset is derived from 249,570 hybrid capture–based targeted tumor sequencing assays and cBioPortal comprehends a curated set of 226 non-redundant studies containing a total of 105,424 samples (queried 2024/11). Mutation counts per amino acid were normalized to the mutational rate across the full FGFR2 amino acid sequence, and significant mutation enrichments were calculated using binomial testing.

Statistical analysis

Data of *in vitro* and *in vivo* experiments were analyzed using Prism and ClustVis. Analyses of RNA-seq and *FGFR2* oncogenomic data were performed using R. Proteomic data were processed with Proteome Discoverer and gene set overrepresentation analysis for differentially enriched proteins was performed using the PANTHER Classification System. *In vitro* experiments were independently repeated at least twice, and all attempts at replication were successful. Across these, data on $n \geq 3$ replicas were collected. No sample size calculations were performed. Sample sizes of mouse cohorts and *ex vivo* H&E and IHC analyses thereof were based on previous calculations [31] or determined using a sample size calculator (<http://powerandsamplesize.com/Calculators>) and were large enough to measure the effect sizes. Data were reproducible across batches analyzed, and all attempts at replication were successful. The statistical tests used are described in the corresponding

figure legends. Unless specified, the FDR multiple testing correction model using the two-stage step-up method from Benjamini, Krieger, and Yekutieli was applied. $P < 0.05$ was considered statistically significant. Exact P values are disclosed in the corresponding figure panels or, where indicated, in Supplementary Table S2.

Key resources

Key resources (antibodies, commercial assays, experimental models, reagents, and software) are listed in Supplementary Table S1.

Data availability

The proteomic and RNA-seq data generated in this study are available in the Proteomics Identifications Database (PRIDE) at PXD041797 and in the European Nucleotide Archive (ENA) at PRJEB60417 as well as in Supplementary Tables S3 and S4. The mouse reference Swiss-Prot proteome and genome used for proteomic and RNA-seq data analyses are available at UniProt (16,714 entries, released 2015/04; <https://www.uniprot.org>) and NCBI (GRCm38 mm10; https://www.ncbi.nlm.nih.gov/assembly/GCF_000001635.20/). The human FGFR2-KD^{pY1} crystal structure is available in RCSB PDB at 6V6Q [7]. The two *FGFR2* missense mutation datasets were previously published [31] and are available at cBioPortal [29]. Raw data generated or analyzed during this study are either included in this article and its Supplementary Data or are available from the corresponding authors upon request.

ONLINE CONTENT

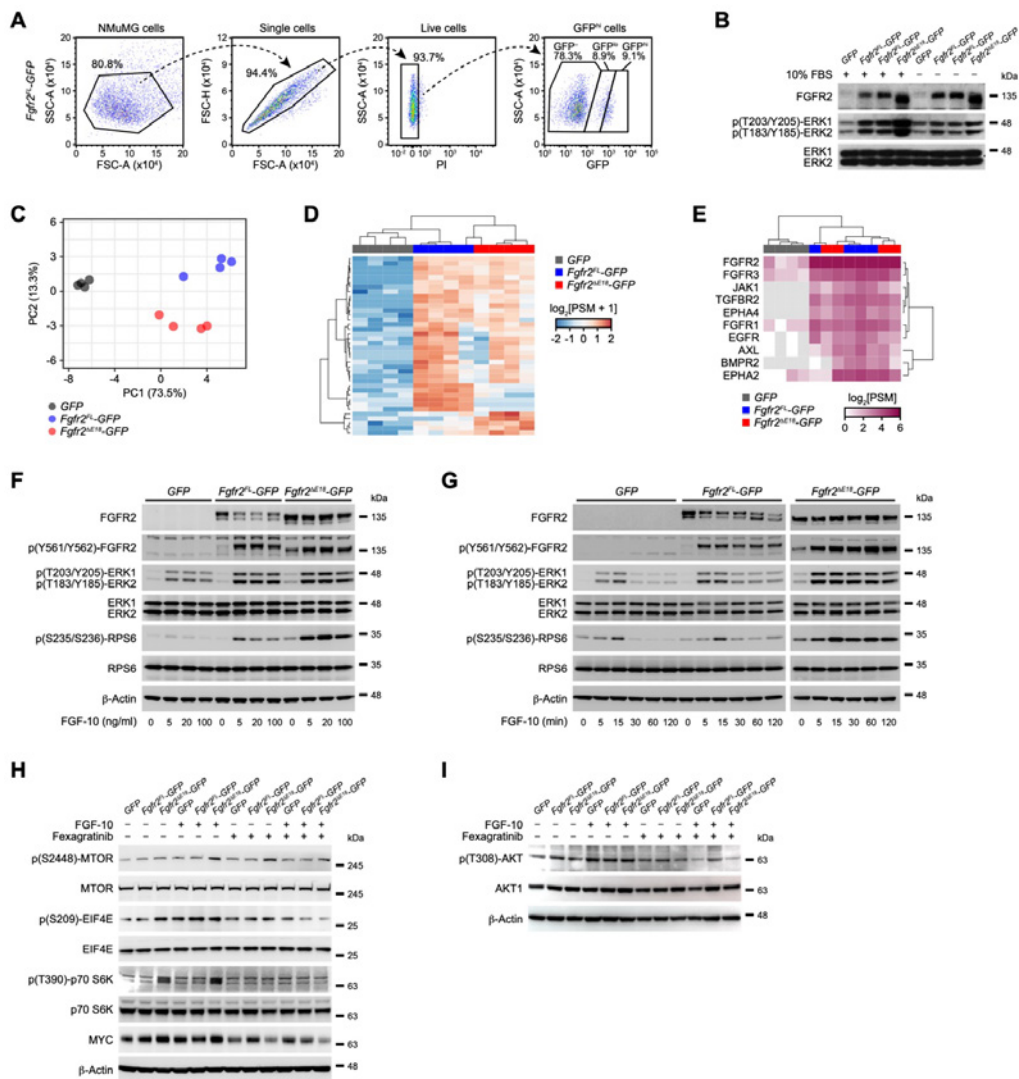
The supplementary data for this article, including Tables S1-S4 and Supplementary Methods are available at <https://doi.org/10.1158/0008-5472.CAN-24-3349>.

REFERENCES

1. Lemmon, M.A. and Schlessinger, J. (2010) Cell signaling by receptor tyrosine kinases. *Cell* 141, 1117–1134
2. Rivetti, S. *et al.* (2020) Fgf10/Fgfr2b Signaling in Mammary Gland Development, Homeostasis, and Cancer. *Front Cell Dev Biol* 8, 415
3. Xie, Y. *et al.* (2020) FGF/FGFR signaling in health and disease. *Signal Transduct Target Ther* 5, 181
4. Katoh, M. *et al.* (2024) FGFR-targeted therapeutics: clinical activity, mechanisms of resistance and new directions. *Nat Rev Clin Oncol* 21, 312–329
5. Ong, S.H. *et al.* (2000) FRS2 proteins recruit intracellular signaling pathways by binding to diverse targets on fibroblast growth factor and nerve growth factor receptors. *Mol Cell Biol* 20, 979–989
6. Cha, J.Y. *et al.* (2009) Aberrant receptor internalization and enhanced FRS2-dependent signaling contribute to the transforming activity of the fibroblast growth factor receptor 2 IIIb C3 isoform. *J Biol Chem* 284, 6227–6240
7. Lin, C.-C. *et al.* (2023) The combined action of the intracellular regions regulates FGFR2 kinase activity. *Commun Biol* 6, 728
8. Seo, J.-H. *et al.* (2009) Structural and functional basis of a role for CRKL in a fibroblast growth factor 8-induced feed-forward loop. *Mol Cell Biol* 29, 3076–3087
9. Salazar, L. *et al.* (2009) A novel interaction between fibroblast growth factor receptor 3 and the p85 subunit of phosphoinositide 3-kinase: activation-dependent regulation of ERK by p85 in multiple myeloma cells. *Hum Mol Genet* 18, 1951–1961
10. Francavilla, C. *et al.* (2013) Functional proteomics defines the molecular switch underlying FGF receptor trafficking and cellular outputs. *Mol Cell* 51, 707–722
11. Dudka, A.A. *et al.* (2010) Signal transducers and activators of transcription-3 binding to the fibroblast growth factor receptor is activated by receptor amplification. *Cancer Res* 70, 3391–3401
12. Mohammadi, M. *et al.* (1992) Point mutation in FGF receptor eliminates phosphatidylinositol hydrolysis without affecting mitogenesis. *Nature* 358, 681–684
13. Ceridono, M. *et al.* (2005) Tyrosine 769 of the keratinocyte growth factor receptor is required for receptor signaling but not endocytosis. *Biochem Biophys Res Commun* 327, 523–532
14. Lin, C.-C. *et al.* (2022) Receptor tyrosine kinases regulate signal transduction through a liquid-liquid phase separated state. *Mol Cell* 82, 1089–1106.e12
15. St-Germain, J.R. *et al.* (2015) Differential regulation of FGFR3 by PTPN1 and PTPN2. *Proteomics* 15, 419–433
16. Ahmed, Z. *et al.* (2013) Grb2 controls phosphorylation of FGFR2 by inhibiting receptor kinase and Shp2 phosphatase activity. *J Cell Biol* 200, 493–504
17. Ahmed, Z. *et al.* (2010) Direct binding of Grb2 SH3 domain to FGFR2 regulates SHP2 function. *Cell Signal* 22, 23–33
18. Lin, C.-C. *et al.* (2012) Inhibition of basal FGF receptor signaling by dimeric Grb2. *Cell* 149, 1514–1524
19. Reilly, J.F. *et al.* (2000) Association of fibroblast growth factor receptor 1 with the adaptor protein Grb14. Characterization of a new receptor binding partner. *J Biol Chem* 275, 7771–7778
20. Szybowska, P. *et al.* (2021) Negative Regulation of FGFR (Fibroblast Growth Factor Receptor) Signaling. *Cells* 10, 1342
21. Lorenzi, M.V. *et al.* (1997) Ligand-independent activation of fibroblast growth factor receptor-2 by carboxyl terminal alterations. *Oncogene* 15, 817–826

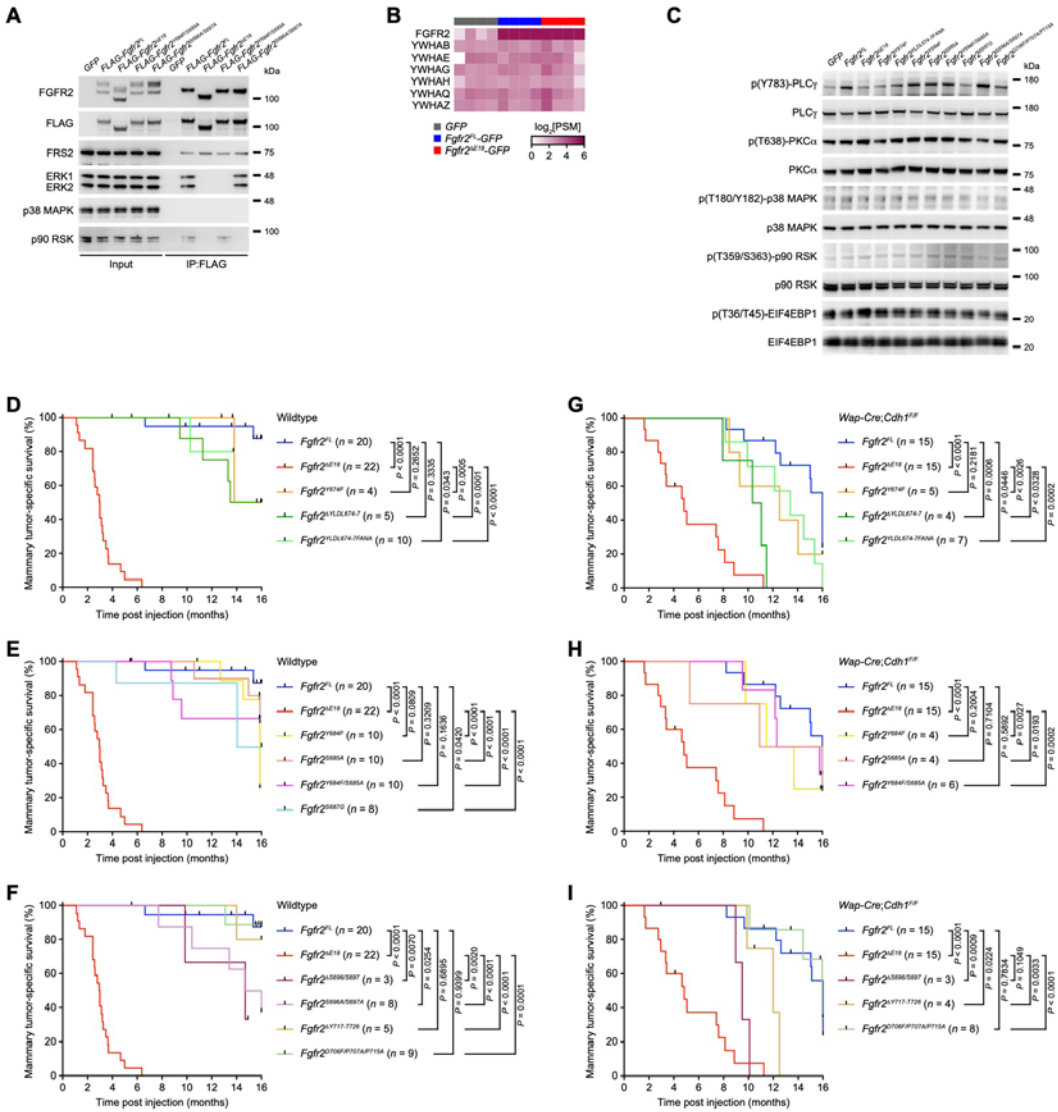
22. Hart, K.C. *et al.* (2001) Identification of tyrosine residues in constitutively activated fibroblast growth factor receptor 3 involved in mitogenesis, Stat activation, and phosphatidylinositol 3-kinase activation. *Mol Biol Cell* 12, 931–942
23. Sørensen, V. *et al.* (2008) Phosphorylation of fibroblast growth factor (FGF) receptor 1 at Ser777 by p38 mitogen-activated protein kinase regulates translocation of exogenous FGF1 to the cytosol and nucleus. *Mol Cell Biol* 28, 4129–4141
24. Zakrzewska, M. *et al.* (2013) ERK-mediated phosphorylation of fibroblast growth factor receptor 1 on Ser777 inhibits signaling. *Sci Signal* 6, ra11
25. Szybowska, P. *et al.* (2019) Cancer Mutations in FGFR2 Prevent a Negative Feedback Loop Mediated by the ERK1/2 Pathway. *Cells* 8, 518
26. Hu, Y. *et al.* (2004) 90-kDa ribosomal S6 kinase is a direct target for the nuclear fibroblast growth factor receptor 1 (FGFR1): role in FGFR1 signaling. *J Biol Chem* 279, 29325–29335
27. Fang, X. *et al.* (2005) Control of CREB-binding protein signaling by nuclear fibroblast growth factor receptor-1: a novel mechanism of gene regulation. *J Biol Chem* 280, 28451–28462
28. Nadratowska-Wesolowska, B. *et al.* (2014) RSK2 regulates endocytosis of FGF receptor 1 by phosphorylation on serine 789. *Oncogene* 33, 4823–4836
29. Cerami, E. *et al.* (2012) The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2, 401–404
30. Rodón, J. *et al.* (2024) Pemigatinib in previously treated solid tumors with activating FGFR1-FGFR3 alterations: phase 2 FIGHT-207 basket trial. *Nat Med* 30, 1645–1654
31. Zingg, D. *et al.* (2022) Truncated FGFR2 is a clinically actionable oncogene in multiple cancers. *Nature* 608, 609–617
32. Turner, N. *et al.* (2010) Integrative molecular profiling of triple negative breast cancers identifies amplicon drivers and potential therapeutic targets. *Oncogene* 29, 2013–2023
33. Pearson, A. *et al.* (2016) High-Level Clonal FGFR Amplification and Response to FGFR Inhibition in a Translational Clinical Trial. *Cancer Discov* 6, 838–851
34. Gavine, P.R. *et al.* (2012) AZD4547: an orally bioavailable, potent, and selective inhibitor of the fibroblast growth factor receptor tyrosine kinase family. *Cancer Res* 72, 2045–2056
35. Lonic, A. *et al.* (2008) Fibroblast growth factor receptor 2 phosphorylation on serine 779 couples to 14-3-3 and regulates cell survival and proliferation. *Mol Cell Biol* 28, 3372–3385
36. Lonic, A. *et al.* (2013) Phosphorylation of serine 779 in fibroblast growth factor receptor 1 and 2 by protein kinase C(epsilon) regulates Ras/mitogen-activated protein kinase signaling and neuronal differentiation. *J Biol Chem* 288, 14874–14885
37. Derksen, P.W.B. *et al.* (2011) Mammary-specific inactivation of E-cadherin and p53 impairs functional gland development and leads to pleomorphic invasive lobular carcinoma in mice. *Dis Model Mech* 4, 347–358
38. Annunziato, S. *et al.* (2016) Modeling invasive lobular breast carcinoma by CRISPR/Cas9-mediated somatic genome editing of the mammary gland. *Genes Dev* 30, 1470–1480
39. Wakefield, L.M. *et al.* (2003) Spontaneous pituitary abnormalities and mammary hyperplasia in FVB/NCr mice: implications for mouse modeling. *Comp Med* 53, 424–432
40. Stüber, J.C. *et al.* (2019) High-Throughput Quantification of Surface Protein Internalization and Degradation. *ACS Chem Biol* 14, 1154–1163
41. Furdui, C.M. *et al.* (2006) Autophosphorylation of FGFR1 kinase is mediated by a sequential and precisely ordered reaction. *Mol Cell* 21, 711–717
42. Jumper, J. *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589
43. Fazioli, F. *et al.* (1993) Eps8, a substrate for the epidermal growth factor receptor kinase, enhances EGF-dependent mitogenic signals. *EMBO J* 12, 3799–3808

44. Cunningham, D.L. *et al.* (2013) Novel binding partners and differentially regulated phosphorylation sites clarify Eps8 as a multi-functional adaptor. *PLoS One* 8, e61513
45. Liu, P.C.C. *et al.* (2020) INCB054828 (pemigatinib), a potent and selective inhibitor of fibroblast growth factor receptors 1, 2, and 3, displays activity against genetically defined tumor models. *PLoS One* 15, e0231877
46. Hubbard, S.R. *et al.* (1994) Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature* 372, 746–754
47. Mol, C.D. *et al.* (2004) Structural basis for the autoinhibition and STI-571 inhibition of c-Kit tyrosine kinase. *J Biol Chem* 279, 31655–31663
48. Zhang, X. *et al.* (2006) An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell* 125, 1137–1149
49. Schiering, N. *et al.* (2003) Crystal structure of the tyrosine kinase domain of the hepatocyte growth factor receptor c-Met and its complex with the microbial alkaloid K-252a. *Proc Natl Acad Sci U S A* 100, 12654–12659
50. Khazaie, K. *et al.* (1988) Truncation of the human EGF receptor leads to differential transforming potentials in primary avian fibroblasts and erythroblasts. *EMBO J* 7, 3061–3071
51. Bardelli, A. *et al.* (1999) A peptide representing the carboxyl-terminal tail of the met receptor inhibits kinase activity and invasive growth. *J Biol Chem* 274, 29274–29281
52. Johnson, J.L. *et al.* (2023) An atlas of substrate specificities for the human serine/threonine kinome. *Nature* 613, 759–766
53. Hart, K.C. *et al.* (2000) Transformation and Stat activation by derivatives of FGFR1, FGFR3, and FGFR4. *Oncogene* 19, 3309–3320
54. Cunningham, D.L. *et al.* (2020) Differential responses to kinase inhibition in FGFR2-addicted triple negative breast cancer cells: a quantitative phosphoproteomics study. *Sci Rep* 10, 7950
55. Lanzetti, L. *et al.* (2000) The Eps8 protein coordinates EGF receptor signalling through Rac and trafficking through Rab5. *Nature* 408, 374–377
56. Luo, K. *et al.* (2021) Effects and mechanisms of Eps8 on the biological behaviour of malignant tumours (Review). *Oncol Rep* 45, 824–834
57. Maa, M.-C. *et al.* (2007) Eps8 facilitates cellular growth and motility of colon cancer cells by increasing the expression and activity of focal adhesion kinase. *J Biol Chem* 282, 19399–19409
58. Goyal, L. *et al.* (2017) Polyclonal Secondary FGFR2 Mutations Drive Acquired Resistance to FGFR Inhibition in Patients with FGFR2 Fusion-Positive Cholangiocarcinoma. *Cancer Discov* 7, 252–263
59. Krook, M.A. *et al.* (2020) Efficacy of FGFR Inhibitors and Combination Therapies for Acquired Resistance in FGFR2-Fusion Cholangiocarcinoma. *Mol Cancer Ther* 19, 847–857
60. Packer, L.M. *et al.* (2017) PI3K Inhibitors Synergize with FGFR Inhibitors to Enhance Antitumor Responses in FGFR2mutant Endometrial Cancers. *Mol Cancer Ther* 16, 637–648
61. Hafner, M. *et al.* (2016) Growth rate inhibition metrics correct for confounders in measuring sensitivity to cancer drugs. *Nat Methods* 13, 521–527
62. Gene Ontology Consortium (2021) The Gene Ontology resource: enriching a GOLD mine. *Nucleic Acids Res* 49, D325–D334
63. Aoki-Kinoshita, K.F. and Kanehisa, M. (2007) Gene annotation and pathway mapping in KEGG. *Methods Mol Biol* 396, 71–91
64. Liberzon, A. *et al.* (2015) The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* 1, 417–425
65. Cardiff, R.D. *et al.* (2000) The mammary pathology of genetically engineered mice: the consensus report and recommendations from the Annapolis meeting. *Oncogene* 19, 968–988



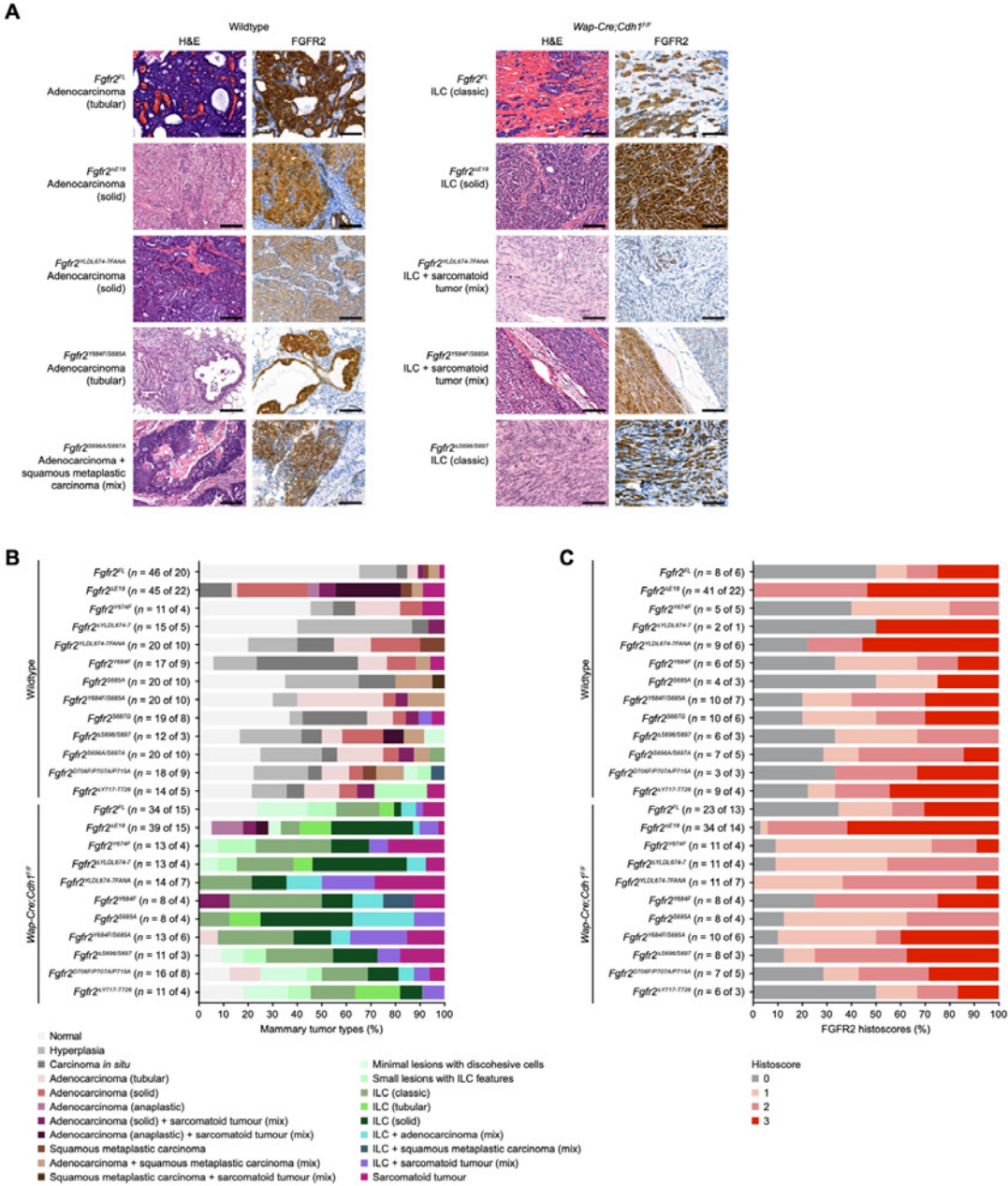
Supplementary Figure S1. Isolation and analysis of GFP^{hi} NMuMG cells expressing *Fgfr2*-GFP variants.

A, FACS gating strategy to isolate GFP^{hi} NMuMG cells expressing *GFP*, *Fgfr2^{FL}-GFP*, or *Fgfr2^{ΔE18}-GFP*. Cells were gated as follows: (i) FSC-A / SSC-A to select the bulk of NMuMG cells and exclude debris events, (ii) FSC-A / FSC-H to select single cells, (iii) propidium iodide (PI) / SSC-A to select live cells, and (iv) GFP / SSC-A to select GFP^{hi} cells. **B**, Western blots showing expression and phosphorylation of indicated proteins in GFP^{hi} NMuMG cells expressing *GFP*, *Fgfr2^{FL}-GFP*, or *Fgfr2^{ΔE18}-GFP* and cultured for 48 hr with or without 10% FBS. Data represent *GFP*, *Fgfr2^{ΔE18}-GFP*, *n* = 1; *Fgfr2^{FL}-GFP*, *n* = 2 replica of 4 independent experiments. **C** and **D**, MS-based proteomic data showing differentially enriched proteins after GFP-IP on serum-starved GFP^{hi} NMuMG cells expressing *GFP*, *Fgfr2^{FL}-GFP*, or *Fgfr2^{ΔE18}-GFP*. Proteins with fold change (FC) ≥ 3 and FDR-adjusted *P* values < 0.05 (calculated using one-tailed two-way ANOVA and FDR multiple-testing corrections using the two-stage step-up method from Benjamini, Krieger, and Yekutieli) for *Fgfr2^{FL}-GFP* PSMs versus *GFP* PSMs and/or *Fgfr2^{ΔE18}-GFP* PSMs versus *GFP* PSMs were considered enriched and used for the analyses. Principal components (PCs) in **C** were calculated on row unit variance scaled log₂[PSM + 1] values using singular value decomposition with imputation. Unsupervised hierarchical clustering of enriched proteins (*n* = 38) in **D** was performed on row mean-centered and unit variance scaled log₂[PSM + 1] values using Pearson's distance measure and Ward's minimum variance clustering. **E**, Heatmap depicting log₂[PSM] values for indicated RTKs in indicated GFP-IP samples. Grey squares indicate undetected proteins (PSM = 0). Unsupervised hierarchical clustering was performed on row mean-centered log₂[PSM + 1] values using Pearson's distance measure and Ward's minimum variance clustering. Data in **C-E** are represented as *n* = 4 independent replica per group of 2 independent experiments. **F** and **G**, Western blots showing expression and phosphorylation of indicated proteins in serum-starved GFP^{hi} NMuMG cells expressing *GFP*, *Fgfr2^{FL}-GFP*, or *Fgfr2^{ΔE18}-GFP* and treated for 20 min with indicated FGF10 doses (**F**) or treated with 20 ng/ml FGF-10 for indicated times (**G**). Separate blots in **G** stained with the same antibody were developed and recorded in parallel and subjected to equal post-imaging processing. **H** and **I**, Western blots showing expression and phosphorylation of indicated proteins in serum-starved GFP^{hi} NMuMG cells expressing *GFP*, *Fgfr2^{FL}-GFP*, or *Fgfr2^{ΔE18}-GFP* and treated for 15 min with 20 ng/ml FGF-10 and/or for 3 hr with 100 nM fexagratinib. β-Actin in **F-I** was run on a separate gel as sample processing control. Data in **F-I** represent 1 replica of 2 independent experiments. Weights of protein marker bands (kDa) in **B** and **F-I** are indicated.



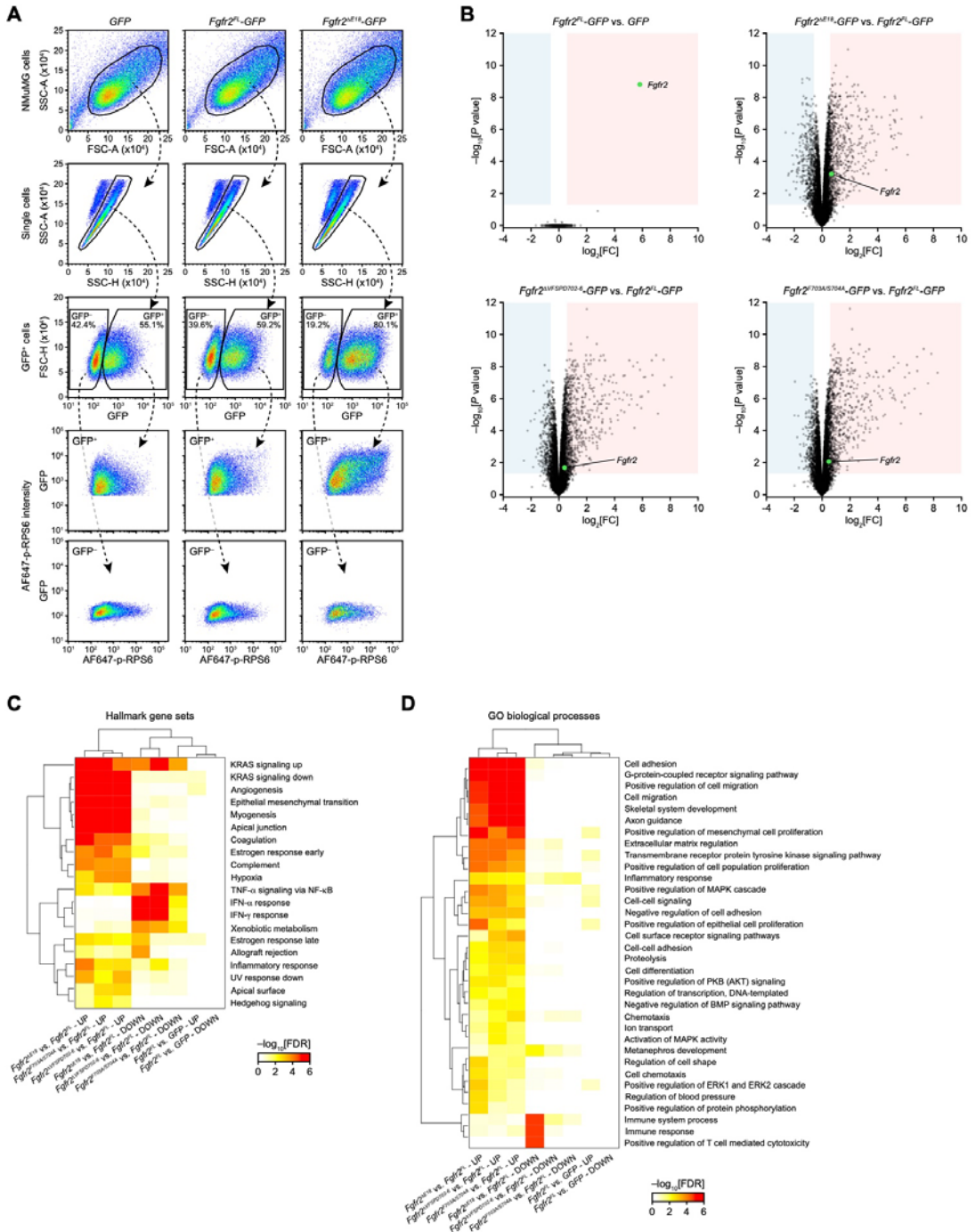
Supplementary Figure S2. Disruption of known C-terminal motifs has little effect on FGFR2 oncogenicity.

A, Western blots showing enrichment of indicated proteins in protein lysate input and FLAG-IP samples of serum-starved NMuMG cells expressing *GFP* or indicated *FLAG-Fgfr2* variants. Data represent 1 replica of 2 independent experiments. **B**, MS-based proteomic data showing differentially enriched proteins after GFP-IPs. Heatmap depicts $\log_2[\text{PSM}]$ values for FGFR2 and indicated 14-3-3 protein family members. Data are represented as $n = 4$ independent replica per group of 2 independent experiments. **C**, Western blots showing expression and phosphorylation of indicated proteins in serum-starved NMuMG cells expressing *GFP* or indicated *Fgfr2* variants. Data represent 1 replica of 3 independent experiments. Weights of protein marker bands (kDa) are indicated. **D-I**, Kaplan-Meier curves showing mammary tumor-specific survival of wildtype (**D-F**) or *Wap-cre;Cdh1^{fl/fl}* (**G-I**) female mice intraductally injected with lentiviruses encoding indicated *Fgfr2* variants. Cohort counts (n) are number of mice. The *Fgfr2^{EL}* and *Fgfr2^{ΔE18}* curves in **D** and **G** are duplicated in **E**, and **F** and in **H** and **I**, respectively. The *Fgfr2^{EL}*, *Fgfr2^{ΔE18}*, and *Fgfr2^{ΔY717-T726}* curves were previously published [31]. P values were calculated using log rank (Mantel-Cox) tests.



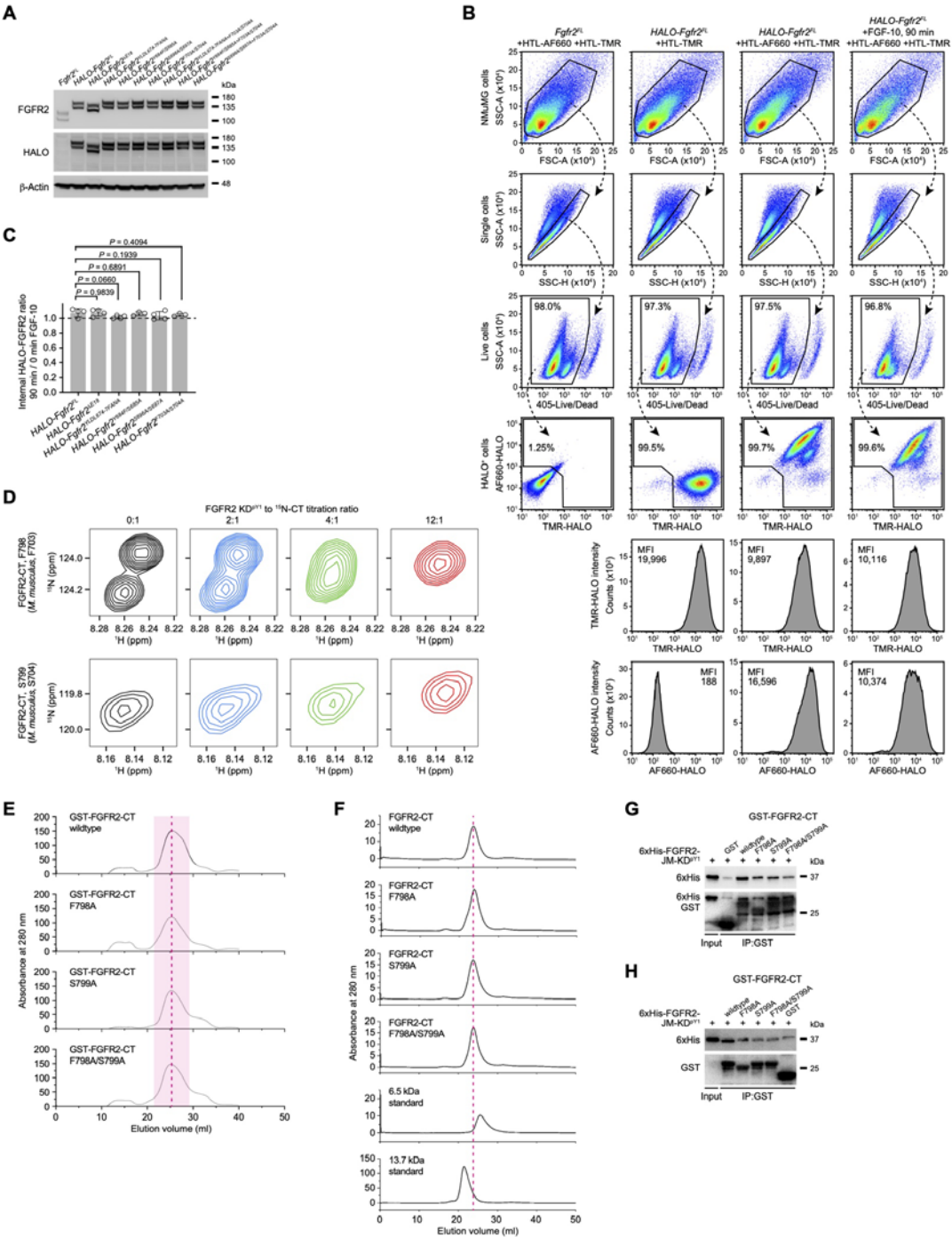
Supplementary Figure S3. Histological examination of mammary glands injected with *Fgfr2-E18* variants.

A, Representative H&E histochemistry and FGFR2 IHC stains on mammary tissue sections from indicated *Fgfr2* mouse models. Per mammary gland, one tissue section was stained and quantified for each of the indicated stains acquired in multiple independent randomized batches across all *Fgfr2* variants. Numbers of stained and quantified glands are in **B** and **C**. ILC, invasive lobular carcinoma. Scale bars, 100 μ m. **B**, Mammary tumor type classifications of *Fgfr2* mouse models based on H&E stains. Cohort counts (*n*) are evaluated mammary glands per number of mice. **C**, Histoscore of FGFR2 IHC stains on mammary tumors from *Fgfr2* mouse models. Cohort counts (*n*) are evaluated mammary tumors per number of mice. Parts of the histological evaluations of the *Fgfr2*^{FL}, *Fgfr2*^{ΔE18}, and *Fgfr2*^{ΔY717-7726} cohorts in **B** and **C** were previously published [31].



Supplementary Figure S4. F703–S704 in the C-tail suppresses FGFR2 signaling activity.

A, FACS gating strategy to analyze p-RPS6 MFI in GFP-positive and GFP-negative serum-starved NMuMG cells expressing *GFP* or indicated *Fgfr2-GFP* variants. Cells were gated as follows: (i) FSC-A / SSC-A to select the bulk of NMuMG cells and exclude debris events, (ii) SSC-H / SSC-A to select single cells, (iii) GFP / FSC-H to select GFP⁻ and GFP⁺ cells, and (iv) AF647-p-RPS6 / GFP to measure p-RPS6 MFI in GFP⁻ and GFP⁺ cells. p-RPS6 intensity histograms are shown in Fig. 4A. **B**, RNA-seq data of serum-starved GFP^{hi} NMuMG cells expressing *GFP* or indicated *Fgfr2-GFP* variants. Volcano plots show differentially expressed genes amongst indicated two-group comparisons. Significantly up or downregulated genes in each comparison were defined with $-0.58 > \log_2[FC] > 0.58$ and Benjamini-Hochberg-adjusted *P* values < 0.05 and are highlighted in blue and red, respectively. *Fgfr2* is indicated in green. **C** and **D**, Hallmark gene set (**C**) and GO biological processes (**D**) enrichment analyses for differentially expressed genes identified from indicated two-group comparisons of RNA-seq data in **B** using two-tailed Fisher's exact tests with FDR multiple-testing corrections using the Benjamini-Hochberg method. Terms that were significantly enriched (FDR < 0.01) by differentially expressed genes in at least one of the comparisons are represented in the heatmap. Negative $\log_{10}[\text{FDR}]$ values were clustered unsupervised using Pearson's distance measure with complete clustering.

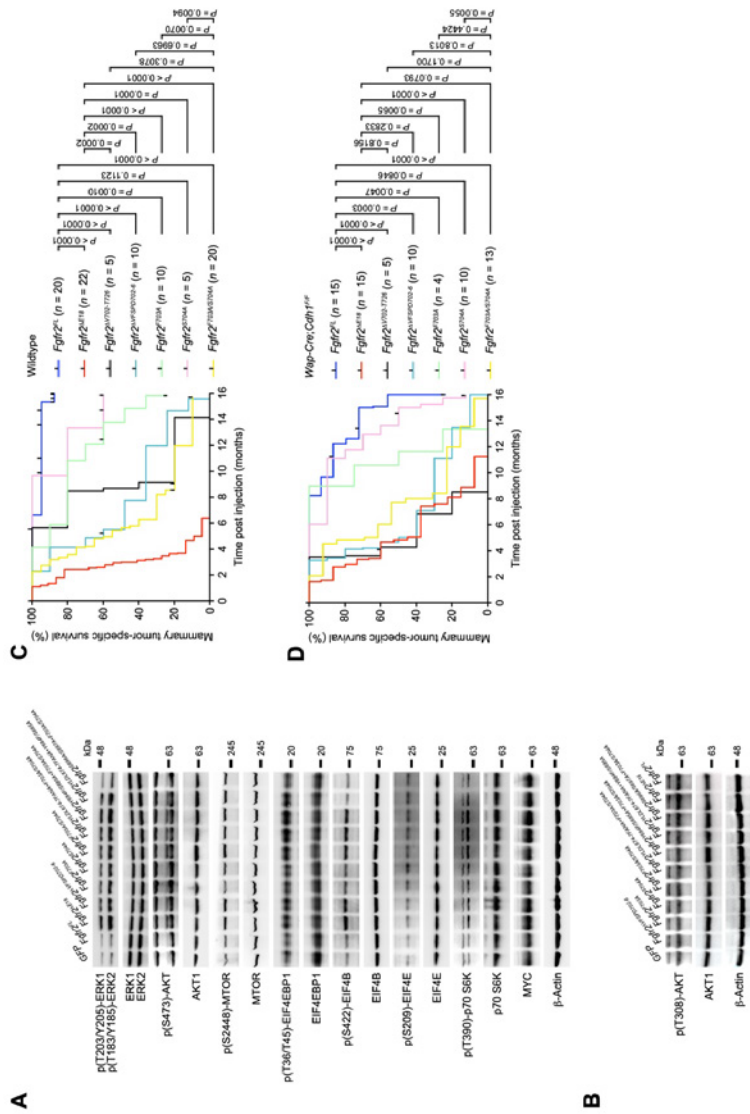


Supplementary Figure S5. Measurements of HALO-FGFR2 internalization and FGFR2-KD-FGFR2-CT interaction.

A, Western blots showing expression of HALO-FGFR2 in NMuMG cells expressing *Fgfr2* or indicated *HALO-Fgfr2* variants. Data represent 1 replica of 1 experiment. β -Actin was run on a separate gel as sample processing control, and weights of protein marker bands (kDa) are indicated. **B**, FACS gating strategy to analyze cell-surface and internalized HALO-tagged FGFR2 in serum-starved NMuMG cells expressing *Fgfr2* or *HALO-Fgfr2* variants following SPIDA [40]. Cells were, where indicated, treated with 20 ng/ml FGF-10 and subsequently labeled firstly with cell membrane impermeable HTL-AF660 and secondly with permeable HTL-TMR. Cells were gated as follows: (i) FSC-A / SSC-A to select the bulk of NMuMG cells and exclude debris events, (ii) SSC-H / SSC-A to select single cells, (iii) 405-Live/Dead / SSC-A to select live cells, (iv) TMR-HALO / AF660-HALO to select HALO⁺ cells, and (v) TMR-HALO and AF660-HALO intensity histograms to measure TMR-HALO and AF660-HALO MFI. **C**, Ratios of internal FGFR2 levels in NMuMG cells expressing indicated *HALO-Fgfr2* variants and stimulated with FGF-10 for 90 min versus unstimulated cells. The ratios are represented as mean $\pm$ s.d. and were calculated based on the data displayed in Fig. 5B. *P* values were calculated using one-tailed one-way ANOVA and FDR multiple-testing corrections using the two-stage step-up method from Benjamini, Krieger, and Yekutieli. **D**, ¹H-¹⁵N HSQC-2D-NMR spectroscopy used to quantify CSPs in parts per million (ppm) of recombinant human ¹⁵N-labeled FGFR2 C-terminal tail (CT) titrated with recombinant human unlabeled monophosphorylated FGFR2 kinase domain (FGFR2-KD^{py}). Visualizations of FGFR2-CT F798 and S799 CSPs using indicated KD^{py} to 15N-CT ratios are shown. **E**, Size exclusion chromatography (SEC) histograms of indicated recombinant GST-tagged FGFR2-CT variants. CT peaks highlighted in pink were collected during SEC purification. **F**, Analytical SEC of indicated SEC-purified and GST-tag-cleaved FGFR2-CT variants. Dotted pink lines in **E** and **F** indicate aligned peak maxima. **G** and **H**, Western blots showing enrichment of 6xHis-tagged FGFR2-JD-KD^{py} incubated with GST-tag alone or indicated GST-tagged FGFR2-CT variants followed by GST-IP. Membranes blotted for 6xHis were directly reblotted for GST (**G**) or stripped prior to GST blotting (**H**). Data represent 2 replica of 11 independent experiments.

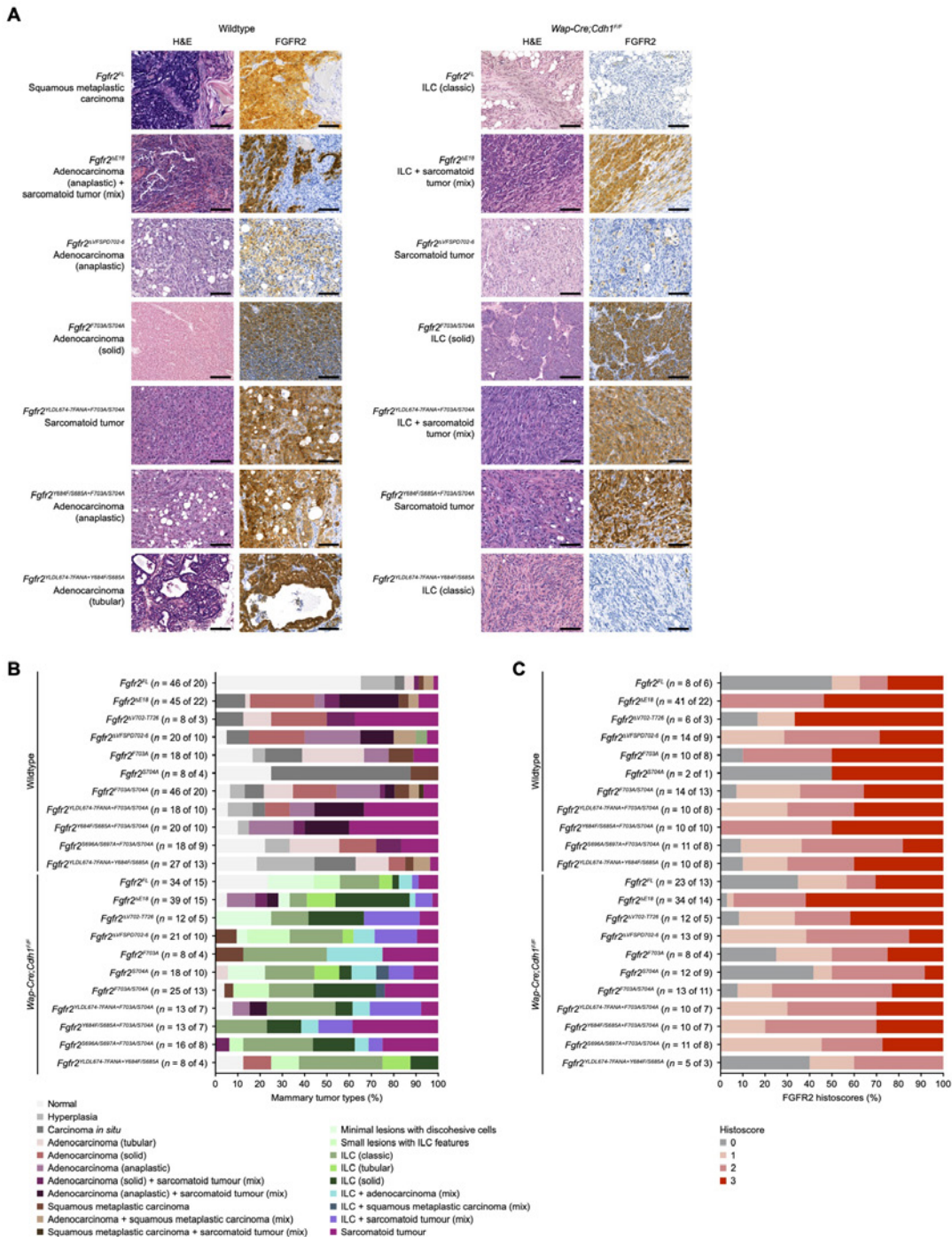


A and **B**, Amino acid sequence alignments of the C-tails of representative vertebrate FGFR2 orthologs (**A**) and the human FGFR paralogs (**B**). C-terminal binding sites and motifs previously associated with FGFR1 and FGFR2 signaling activity regulation as well as the here-identified KDBS motif (human, F798–S799; mouse, F703–S704) are indicated. Amino acid counts refer to the human FGFR2 NP_000132.3 isoform.



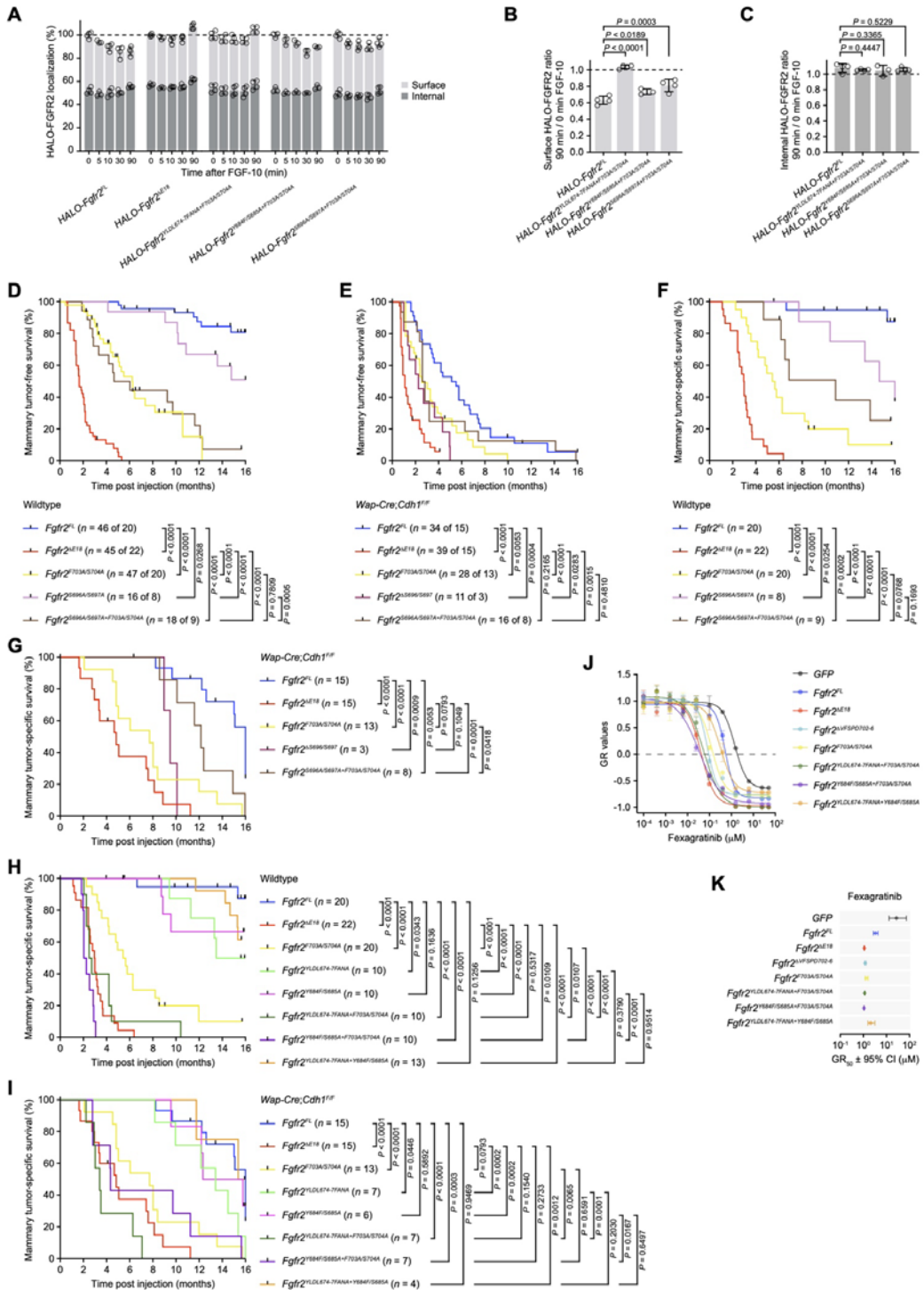
Supplementary Figure S7. C-terminal F703-S704 disruption drives FGFR2 oncogenicity.

A and **B**, Western blots showing expression and phosphorylation of indicated proteins in serum-starved NMuMG cells expressing GFP or indicated *Fgf2* variants. β -Actin was run on a separate gel as sample processing control, and weights of protein marker bands (kDa) are indicated. Data represent 1 replica of 3 independent experiments. **C** and **D**, Kaplan-Meier curves showing mammary tumor-specific survival of wildtype (**C**) or *Wap-cre;Cdh1^{fl/fl}* (**D**) female mice intraductally injected with lentiviruses encoding indicated *Fgf2* variants. Cohort counts (n) are number of mice. The *Fgf2^{WT}* and *Fgf2^{Δ1102-1128}* curves are duplicates from Supplementary Fig. S2D and G. The *Fgf2^{Δ1102-1128}* curves were previously published [31]. P values were calculated using log rank (Mantel-Cox) tests.



Supplementary Figure S8. Histological examination of mammary glands injected with *Fgfr2* KDBS motif variants.

A, Representative H&E and FGFR2 stains on mammary tissue sections from indicated *Fgfr2* mouse models. Per mammary gland, one tissue section was stained and quantified for each of the indicated stains acquired in multiple independent randomized batches across all *Fgfr2* variants. Numbers of stained and quantified glands are in **B** and **C**. Scale bars, 100 μ m. **B**, Mammary tumor type classifications of *Fgfr2* mouse models based on H&Es stains. Cohort counts (*n*) are evaluated mammary glands per number of mice. **C**, Histoscore of FGFR2 IHC stains on mammary tumors from *Fgfr2* mouse models. Cohort counts (*n*) are evaluated mammary tumors per number of mice. The data for *Fgfr2*^{FL} and *Fgfr2*^{ΔE18} cohorts in **B** and **C** are duplicates from Supplementary Fig. S3B and C. Parts of the histological evaluations of the *Fgfr2*^{ΔV702-T726} cohorts in **B** and **C** were previously published [31].



Supplementary Figure S9. Co-targeting of F703–S704 and C-terminal motifs phenocopies FGFR2^{ΔE18}.

A, Quantification of cell-surface and internalized HALO-tagged FGFR2 in serum-starved NMuMG cells expressing indicated HALO-Fgfr2 variants during FGF-10 stimulation using FACS and SPIDA. Cells were treated with 20 ng/ml FGF-10 for indicated times and subsequently labeled firstly with cell membrane impermeable HTLAF660 and secondly with permeable HTL-TMR. For FACS gating strategy, see Supplementary Fig. S5B. Data are represented as mean ± s.d. of HALO-Fgfr2^{FL}, HALO-Fgfr2^{ΔE18}, FGF-10, 0 min, 10 min, 30 min, 90 min, HALO-Fgfr2^{Y1DL674-7FANA+F703A/S704A}, HALO-Fgfr2^{S696A/S697A}, all timepoints, *n* = 4; other variants and timepoints, *n* = 3 independent replica. The HALO-Fgfr2^{FL} and HALO-Fgfr2^{ΔE18} values are duplicated from Fig. 5B. **B** and **C**, Ratios of cell surface (**B**) and internal (**C**) FGFR2 levels in NMuMG cells expressing indicated HALO-Fgfr2 variants and stimulated with FGF-10 for 90 min versus unstimulated cells. The ratios are represented as mean ± s.d. and were calculated based on the data displayed in **A**. The HALO-Fgfr2^{FL} values are duplicated from Fig. 5C and Supplementary Fig. S5C. **D–I**, Kaplan-Meier curves showing mammary tumor-free (**D** and **E**) or mammary tumor-specific (**F–I**) survival of wildtype (**D**, **F**, and **H**) or *Wap-cre;Cdh1^{cre}* (**E**, **G**, and **I**) female mice intraductally injected with lentiviruses encoding indicated Fgfr2 variants. Cohort counts (*n*) are injected mammary glands per number of mice (**D** and **E**) or number of mice (**F–I**). The Fgfr2^{FL}, Fgfr2^{ΔE18}, Fgfr2^{F703A/S704A}, Fgfr2^{S696A/S697A}, Fgfr2^{Y1DL674-7FANA}, and Fgfr2^{Y684F/S685A} curves are duplicates from Figs. 3H and K and 6E and F, and Supplementary Figs. S2D–I and S7C and D. *P* values were calculated using one-tailed one-way ANOVA and FDR multiple-testing corrections using the two-stage step-up method from Benjamini, Krieger, and Yekutieli (**B** and **C**) or log rank (Mantel-Cox) tests (**D–I**). **J**, GR-corrected dose-response curves of NMuMG cells expressing GFP or indicated Fgfr2 variants and treated for 4 days with fexagratinib. Data are represented as mean ± standard error of the mean of *n* = 5 replica per group. **K**, GR50 values for fexagratinib based on dose-response experiments in **J**. Data are represented as mean ± 95% CI. GR dose-response curves (**J**) and GR50 values (**K**) were modeled using [inhibitor] versus response with variable slope (four parameters), least-squares regressions, and asymmetrical CIs. Data in **J** and **K** represent 1 replica of 2 independent experiments.

5

CHAPTER 5

C-TERMINAL TRUNCATION AND FUSION PARTNER DETERMINE ONCOGENICITY OF FGFR3

Julia Yemelyanenko*, Jinhyuk Bhin*, Eline van der Burg, Anne Paulien Drenth, Jessica K. Lee, Catrin Lutz, Lea Dörner, Ellen Wientjens, Sjoerd Klarenbeek, Ji-Ying Song, Hyeonjin Moon, Stefano Annunziato, Natalie Proost, Bjørn Siteur, Jeffrey S. Ross, Marieke van de Ven, Olaf van Tellingen, Shridar Ganesan, Lodewyk F. A. Wessels, Daniel Zingg# & Jos Jonkers#

*equal contribution

#co-corresponding author

Cancer Research, 2026 Mar 16; 86(6):1372-1391
<https://doi.org/10.1158/0008-5472.CAN-24-2648>

ABSTRACT

Genomic alterations affecting components of the fibroblast growth factor (FGF) signaling axis can trigger aberrant pathway activation and tumor development. Genomic truncation of the FGF receptor 2 (*FGFR2*) exon 18 (E18) disrupts the *FGFR2* carboxy-terminal tail (C-tail), acting as a potent driver alteration across multiple tumor types. Here, we analyzed human oncogenomic datasets to reveal that E18 truncations are similarly prevalent in *FGFR3*, an *FGFR2* paralog. *FGFR3* E18 truncations primarily occurred due to rearrangements (REs) that involve transforming acidic coiled-coil containing protein 3 (*TACC3*), resulting in *FGFR3*^{ΔE18}-*TACC3* gene fusions. In contrast to E18-truncated *FGFR2*, functional *in vitro* and *in vivo* examination of *Fgfr3* variants demonstrated that the truncation of *Fgfr3* E18 is insufficient to promote oncogenic activity in cell lines or in the lungs and mammary glands of mice. Only the combination of an *Fgfr3* E18 truncation with a RE partner gene that encodes a receptor-dimerizing domain resulted in the development of tumors, which were sensitive to FGFR inhibition. Overall, these findings suggest that patients with cancers that are positive for rearranged *FGFR3* resulting in E18 truncation and a fusion to dimerizing partners should be considered for FGFR-targeted therapies.

Significance

FGFR3, unlike its paralog *FGFR2*, requires both a C-terminal truncation and fusion to a partner gene that retains the expression of a dimerizing domain to effectively drive oncogenic signaling and tumorigenesis.

INTRODUCTION

Genomic alterations affecting components of the fibroblast growth factor (FGF) / FGF receptor (FGFR) signaling pathway are frequent in human cancer. In fact, approximately 7% of cancer patients across a wide variety of tumor types harbor alterations in *FGFR* genes [1–3]. Gene amplifications (~60%) are the most prevalent alterations, followed by activating mutations (~30%) and gene fusions/rearrangements (REs, ~10%; refs. [1,3]). Notably, the incidence of alterations in the individual *FGFR* genes varies in a tissue-specific manner [1–3]. *FGFR1* amplifications, present as part of a large 8p11 amplicon, are the most common *FGFR* aberrations, primarily observed in breast, urothelial, ovarian, and squamous cell lung carcinomas. *FGFR2* activating mutations and amplifications are prevalent in endometrial cancer, whereas *FGFR2* fusions are frequently found in cholangiocarcinoma. *FGFR3* is most commonly altered in urothelial bladder and upper urinary tract cancers, with activating mutations and gene fusions being predominant.

The pan-FGFR inhibitors (FGFRi) erdafitinib and futibatinib as well as the FGFR1–3 inhibitors pemigatinib and infigratinib have recently received regulatory approvals in several *FGFR*-altered indications, and further small-molecule FGFRi and other biologic agents targeting FGFRs are currently being evaluated in clinical trials [4]. However, only 20–40% of patients with solid *FGFR*-altered tumors show objective responses to FGFRi [4]. This underscores the need to discern which FGFR alterations are tumor-driving and thus amenable to targeted therapeutic strategies.

We recently uncovered truncation of exon 18 (E18) of *FGFR2* as a potent driver alteration in cancer, irrespective of specific fusion partners [5]. Our work showed that the carboxy (C)-terminal tail of FGFR2, encoded by E18, can be subjected to a plethora of genomic alterations that lead to its loss and the subsequent initiation of oncogenic transformation. These alterations include canonical in-frame fusions with 3'-partner genes, non-canonical REs, partial E1-E17 amplifications, E18 splice-acceptor-site mutations and E18-truncating mutations. Importantly, analysis of FGFRi clinical trial data revealed that patients with cancers containing C-terminally-truncating *FGFR2* variants show clinically meaningful responses, irrespective of the truncation variant type [5,6].

Cancer genomic profiling has identified recurrent breakpoints (BPs) within intron 17 (I17) and E18 of *FGFR3* that lead to the generation of C-terminally truncated FGFR3 [5,7,8]. Given the structural and functional similarities between FGFR2 and FGFR3 [9], we hypothesized that, comparably to FGFR2, the loss of the FGFR3 C-terminal tail may suffice to generate an oncogene. However, oncogenomic analysis and functional annotation of *FGFR3* alterations revealed that FGFR3, as opposed to its paralog FGFR2, requires both C-terminal truncation and a fusion partner with retained dimerizing capacity to convert to an oncogene.

RESULTS

FGFR3 E18 truncations in human cancer

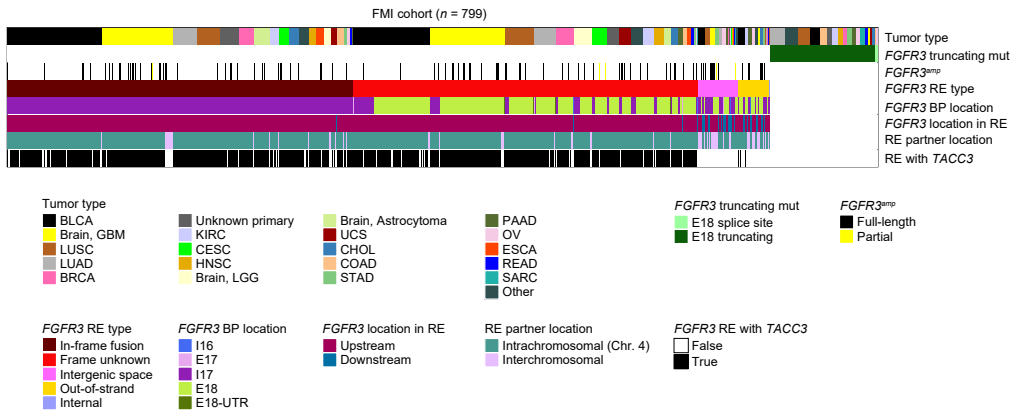
To study human cancer-associated *FGFR3* alterations causing E18-truncation, we analyzed 249,570 pan-cancer targeted tumor sequencing profiles derived from Foundation Medicine Inc. (FMI). We identified 799 samples (0.32% incidence) across a broad range of tumor indications that contain *FGFR3* alterations potentially resulting in the expression of *FGFR3* transcripts devoid of E18 (**Fig. 1A**; Supplementary Data S1). These could be classified into two mutually exclusive categories comprising *FGFR3* E18-truncating mutations ($n=98$, 12.3% of the 799 tumors; **Fig. 1A**; Supplementary Fig. S1A and S1B) and structural REs within the *FGFR3* locus ($n=701$, 87.7%). The REs resulted in canonical in-frame fusions with other partner genes ($n=318$, 39.8%) or non-canonical REs in which the reading frame of the partner gene could not be defined (frame-unknown; $n=317$, 39.7%), the partner sequence had its origin in intergenic space ($n=36$, 4.5%), the partner gene was out-of-strand ($n=29$, 3.6%) and internal gene REs ($n=1$, 0.1%). In most cases, *FGFR3* was located upstream of its RE partner ($n=674$, 96.1%), REs predominantly occurred intrachromosomally within chromosome 4 ($n=648$, 92.4%) and the majority of the *FGFR3* BPs were located either in I17 ($n=410$, 58.5%) or E18 ($n=283$, 40.2%). Notably, transforming acidic coiled-coil containing protein 3 (*TACC3*) was the single most recurrent RE partner gene of *FGFR3* ($n=574$, 81.9%). This finding contrasted with the diverse RE partners observed for *FGFR2* [5].

We next analyzed WGS profiles of 4,863 metastatic solid tumors from the Hartwig Medical Foundation (HMF; ref. [10]) and found 32 tumors (0.66% incidence) that had structural REs affecting *FGFR3* (Supplementary Data S2). Consistent with the results obtained from the FMI cohort, this analysis revealed a significant enrichment of RE BPs in I17 and E18 of *FGFR3* ($n=19$, 59.4%) (**Fig. 1B** and **C**). Amongst these, we found 3 in-frame (15.8%), 14 frame-unknown (73.7%), and 2 intergenic space (10.5%) REs (**Fig. 1D**). 10 samples containing *FGFR3* REs in I17 or E18 concomitantly presented full-length or partial amplifications of the *FGFR3* locus (*FGFR3*^{amp}; 43.8%; **Fig. 1D**).

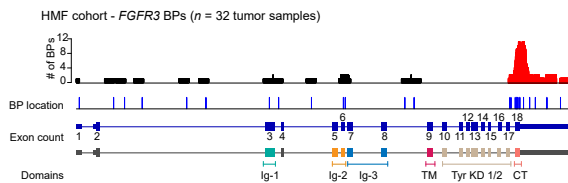
Figure 1. *FGFR3* E18 truncations in human cancer.

A, Analysis of 799 samples (0.32% incidence) carrying potential *FGFR3*-E18 truncations identified among 249,570 pan-cancer diagnostic hybrid-capture panel-seq profiles from FMI. These samples include *FGFR3* REs with BPs at the C-terminal tail, E17/18 or I16/17 and/or *FGFR3* E18-truncating mutations. *FGFR3* REs consist of I17 in-frame fusions ($n=318$, 39.8% incidence among 799 samples with potential *FGFR3*-E18 alterations), frame-unknown REs ($n=317$, 39.7% incidence), intergenic space REs ($n=36$, 4.5% incidence), out-of-strand REs ($n=29$, 3.6% incidence), internal RE ($n=1$, 0.1% incidence) and/or *FGFR3* E18-truncating mutations (mut; $n=98$, 12.3% incidence). Tumor type, *FGFR3* amplification status (*FGFR3*^{amp}; E1-E18 full-length and E1-E17 partial amplifications), and BP location of *FGFR3* and its corresponding RE partners are indicated. *FGFR3* REs with *TACC3* are highlighted. **B**, BPs generating *FGFR3* genomic REs identified in 32 out of 4,863 analyzed WGS profiles from the HMF cohort on metastatic solid tumors [10]. BP density was calculated using a 300 bp sliding window. The blue bars show BPs. Corresponding exon count and protein domains are indicated. Ig, immunoglobulin-like domains; TM, transmembrane domain; Tyr KD, tyrosine kinase domain; CT, C-terminal tail. **C**, Normalized frequency (top panel) and enrichment significance (*P* values, bottom panel) of *FGFR3* genomic RE BPs identified in 32 WGS profiles from the HMF cohort in each exon/intron. BP frequency was normalized by the kilobase of feature (exon/intron) length and the total number of BPs in *FGFR3*. *P* values were calculated with one-tailed binomial tests. **D**, *FGFR3* RE types and full-length and partial amplifications found in 32 WGS profiles from HMF. BP location of

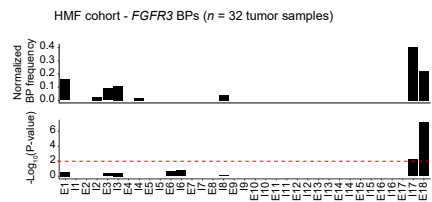
A



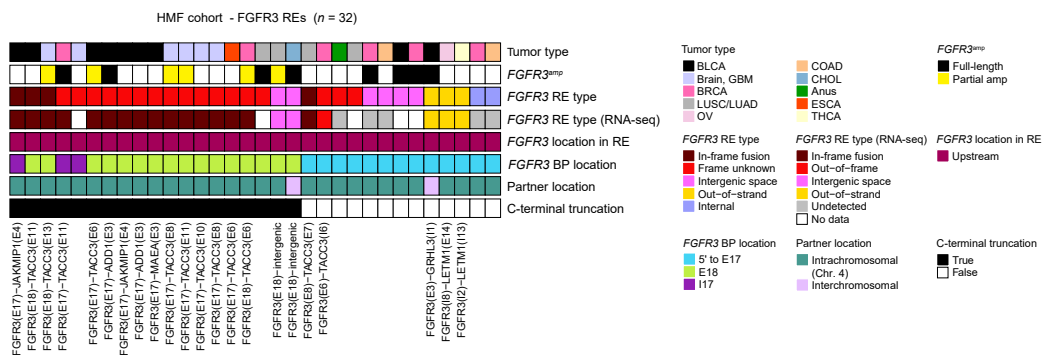
B



C



D



FGFR3 and its corresponding RE partners are indicated. RNA support indicates evidence for *FGFR3* REs in matching RNA-seq profiles, and cases with predicted C-terminal truncation are highlighted. Empty fields, no RNA-seq data available. BLCA, bladder urothelial carcinoma; GBM, glioblastoma multiforme; LUSC, lung squamous cell carcinoma; LUAD, lung adenocarcinoma; BRCA, breast invasive carcinoma; KIRC, kidney renal clear cell carcinoma; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; HNSC, head and neck squamous cell carcinoma; LGG, low-grade glioma; UCS, uterus carcinosarcoma; CHOL, cholangiocarcinoma; COAD, colon adenocarcinoma; STAD, stomach adenocarcinoma; PAAD, pancreatic adenocarcinoma; OV, ovarian serous cystadenocarcinoma; ESCA, oesophageal carcinoma; READ, rectum adenocarcinoma; SARC, sarcoma; THCA, papillary thyroid carcinoma.

Expression of E18-truncating *FGFR3* variants

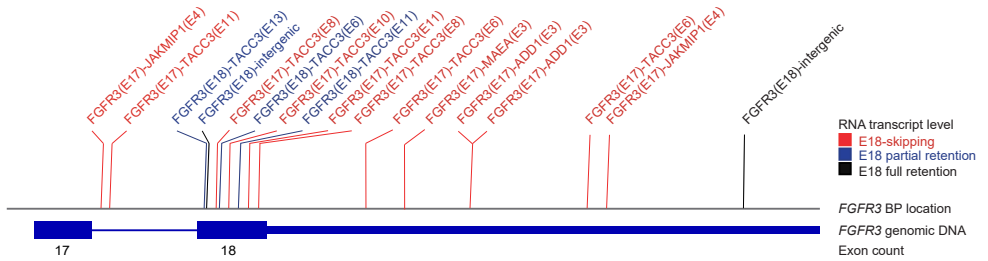
Next, we evaluated matching RNA-seq profiles that were available for 27 out of the 32 tumor samples with WGS profiles that revealed structural REs in *FGFR3* (**Fig. 1D** and **2A-E**; Supplementary Fig. S1C). For 9 samples predicted to have in-frame, out-of-strand, or intergenic space REs, we could confirm the expression of the respective *FGFR3* RE types. Furthermore, 12 out of the 14 tumors with frame-unknown *FGFR3* REs (85.7%) showed in-frame REs at the transcript level (**Fig. 1D** and **2B**). Interestingly, in most cases where BPs were observed in *FGFR3* I17 or E18, transcripts showed E18-skipping with junction reads connecting E17 to the RE partner genes (**Fig. 2A** and **C**; Supplementary Fig. S1C). In three cases with an *FGFR3* BP within E18, we observed expression of transcripts that show in-frame splicing originating within E18 and bridging to the RE partner (**Fig. 2A, B, D** and **E**; E18 partial retention in **Fig. 2A**, in blue). These transcripts encoded *FGFR3* derivatives devoid of a large portion of the C-terminal tail, thus likely resulting in a dysfunctional C-tail reminiscent of full E18 truncation.

In sum, we showed that genomic REs in human tumors cause the expression of *FGFR3* transcripts which either completely or partially lack E18 (*FGFR3^{ΔE18}*), consequently resulting in dysfunctional C-terminal tails. Based on these findings, we propose that the frame-unknown REs encountered in the FMI cohort (**Fig. 1A**) may likewise encode in-frame transcripts.

Figure 2. Expression of E18-truncated *FGFR3* variants.

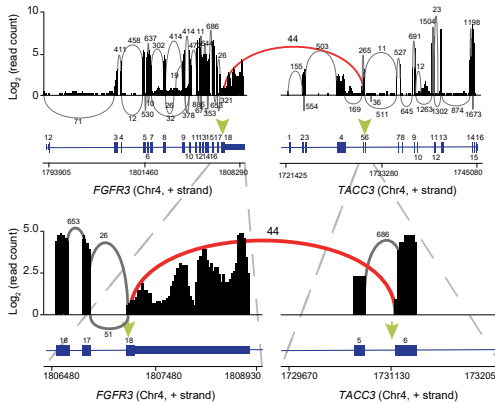
A, Overview of the BP locations on the C-terminal tail of *FGFR3* analyzed by WGS and the resulting transcripts (indicated by labels corresponding to each BP location) analyzed by matched RNA-seq data, from the HMF cohort. In red, transcripts lacking E17–E18 spanning reads which results in complete E18 skipping on RNA level; in blue, transcripts retaining a fraction of E18 comprised of the 5' exonic fragment prior to their corresponding BP; in black, transcripts with full E18 retention. **B**, Sashimi plot showing *FGFR3* read coverage and junction reads of the HMF sample CPCT022410009 (BRCA). *FGFR3*-E18 coding sequence (CDS) to *TACC3*-I5 frame-unknown RE identified with WGS and *FGFR3*-E18 to *TACC3*-E6 in-frame fusion found with RNA-seq. **C**, Sashimi plot showing *FGFR3* read coverage and junction reads of the HMF sample CPCT02050356 (BLCA). An in-frame *FGFR3*-I17 RE to *JAKMIP1*-I3 identified with WGS and an in-frame fusion *FGFR3*-E17 to *JAKMIP1*-E4 confirmed with RNA-seq. **D**, Sashimi plot showing *FGFR3* read coverage and junction reads of the HMF sample CPCT02020277 (BLCA). *FGFR3*-E18 coding sequence (CDS) RE to *TACC3*-E11 CDS identified with WGS and *FGFR3*-E18 to *TACC3*-E11 in-frame fusion confirmed with RNA-seq. **E**, Sashimi plot showing *FGFR3* read coverage and junction reads of the HMF sample CPCT02040255 (LUSC/LUAD). *FGFR3*-E18 CDS RE to intergenic space identified with WGS and *FGFR3*-E18 to intergenic region junctions found with RNA-seq. Green arrows indicate BPs identified with WGS.

A



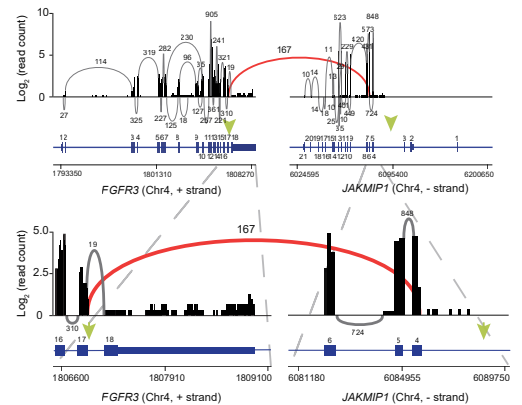
B

HMF-CPCT02410009 (BRCA)
WGS, *FGFR3*-E18 CDS RE with *TACC3*-I5, frame unknown;
RNA-seq, *FGFR3*-E18--*TACC3*-E6 in-frame fusion



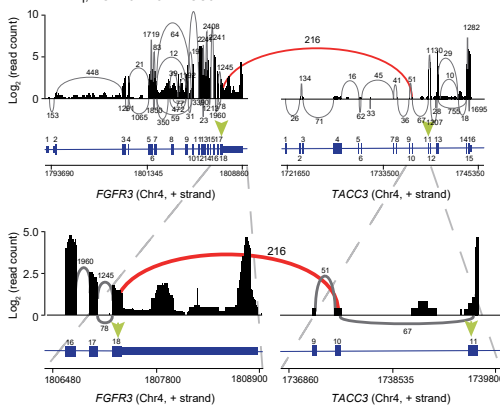
C

HMF-CPCT02050356 (BLCA)
WGS, *FGFR3*-I17 RE with *JAKMIP1*-I3, in-frame;
RNA-seq, *FGFR3*-E17--*JAKMIP1*-E4 in-frame fusion



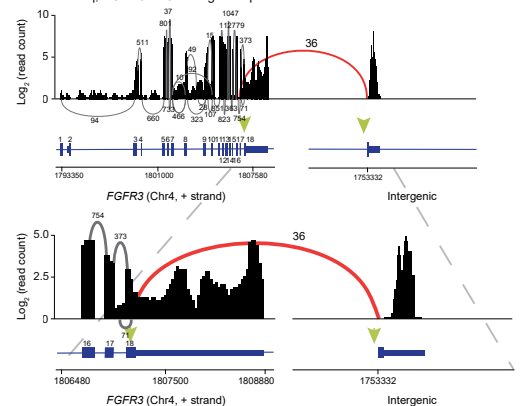
D

HMF-CPCT02020277 (BLCA)
WGS, *FGFR3*-E18 CDS RE with *TACC3*-E11, CDS, in-frame;
RNA-seq, *FGFR3*-E18--*TACC3*-E11 in-frame fusion



E

HMF-CPCT02040255 (LUSC/LUAD)
WGS, *FGFR3*-E18 CDS RE with intergenic space at chr4:1,753,329;
RNA-seq, *FGFR3*-E18--intergenic space fusion



***FGFR3* RE partners are enriched for proteins with self-interacting capacity**

Next, we classified the RE partner genes detected in the FMI cohort in more detail. *TACC3* was the single most recurrent partner gene, observed in 574 (85.3%) of all tumors carrying *FGFR3* E18-truncating REs (**Fig. 3A**; Supplementary Data S1). Intergenic space REs were the second most frequent (3.3%), while other known fusion partner genes were much less common (<2.0% each). Importantly, 94.8% of the proteins encoded by the RE partners contained self-interacting domains (**Fig. 3B**). The coiled-coil (CC) domain was found most frequently (90.5%), but other self-interacting domains, such as the HMG-box, SET, PWWP and Class II Aldolase domains, were also encountered (**Fig. 3B**; Supplementary Table S2). The ability to self-interact and the presence of CC domains was significantly enriched across the *FGFR3* fusion partners as compared to the human proteome (**Fig. 3C and D**).

Previous research examining *FGFR3*^{ΔE18}-*TACC3* fusions has predominantly relied on small cancer patient cohorts limited to specific tissue types [11,12]. In contrast, the extensive FMI data cohort used within this study allowed for detailed and unbiased analysis of the REs between *FGFR3*^{ΔE18} and *TACC3*. The 3'-BPs in *FGFR3* were mainly located in I17 and E18, whereas the 5'-BPs in *TACC3* were more diversely distributed across several introns and exons, particularly I10, I9 and I7 (**Fig. 3E and F**). Most *FGFR3*^{ΔE18}-*TACC3* fusions contained *TACC3* E11-E16, which encode two C-terminal CC domains (**Fig. 3E and F**). Together, these data suggested that the dimerizing capacity of the RE partners may be a relevant feature of oncogenic *FGFR3*^{ΔE18} fusions.

***Fgfr3*^{ΔE18}-*Tacc3* fusion induces oncogenic signaling and *in vitro* transformation**

FGFR3-*TACC3* fusions have been identified in several tumor types [1,11], most commonly in cancers of the brain [1,12–16], urinary tract [1,17], lung [1,18–22], and breast [23–25], in line with our comprehensive genomic analyses (**Fig. 1A and D**; Supplementary Data S1 and S2). Previous studies using mouse models of glioblastoma [13,15] and lung cancer [18] have shown that these fusions can act as potent oncogenes. However, the exact molecular mechanisms underlying their oncogenicity have remained incompletely understood.

To investigate the signaling pathways induced by *FGFR3*^{ΔE18} and *FGFR3*^{ΔE18}-*TACC3* fusions, we expressed GFP, full-length murine *Fgfr3* (*Fgfr3*^{FL}), *Fgfr3* lacking E18 (*Fgfr3*^{ΔE18}), *Fgfr3*^{FL} fused to *Tacc3* (*Fgfr3*^{FL}-*Tacc3*), or *Fgfr3*^{ΔE18}-*Tacc3* (Supplementary Methods S1) in NMuMG mouse mammary epithelial cells, with *Fgfr3*^{ΔE18}-*Tacc3* corresponding to *FGFR3*-*TACC3* fusions observed in the clinic. The *TACC3* BP in *FGFR3*-*TACC3* frequently locates to I7, resulting in a partial *TACC3* transcript encoded by E8-E16 (**Fig. 3E, F and 4A**). The corresponding murine *Tacc3* sequence that was used in the *Fgfr3* constructs consists of E7-E16. All constructs were expressed at comparable levels across engineered cell lines (**Fig. 4A**; Supplementary Fig. S2A). During serum starvation, expression of the *Fgfr3*^{ΔE18}-*Tacc3* fusion resulted in marked baseline activation of *FGFR3* and the mTOR signaling pathway effector proteins,

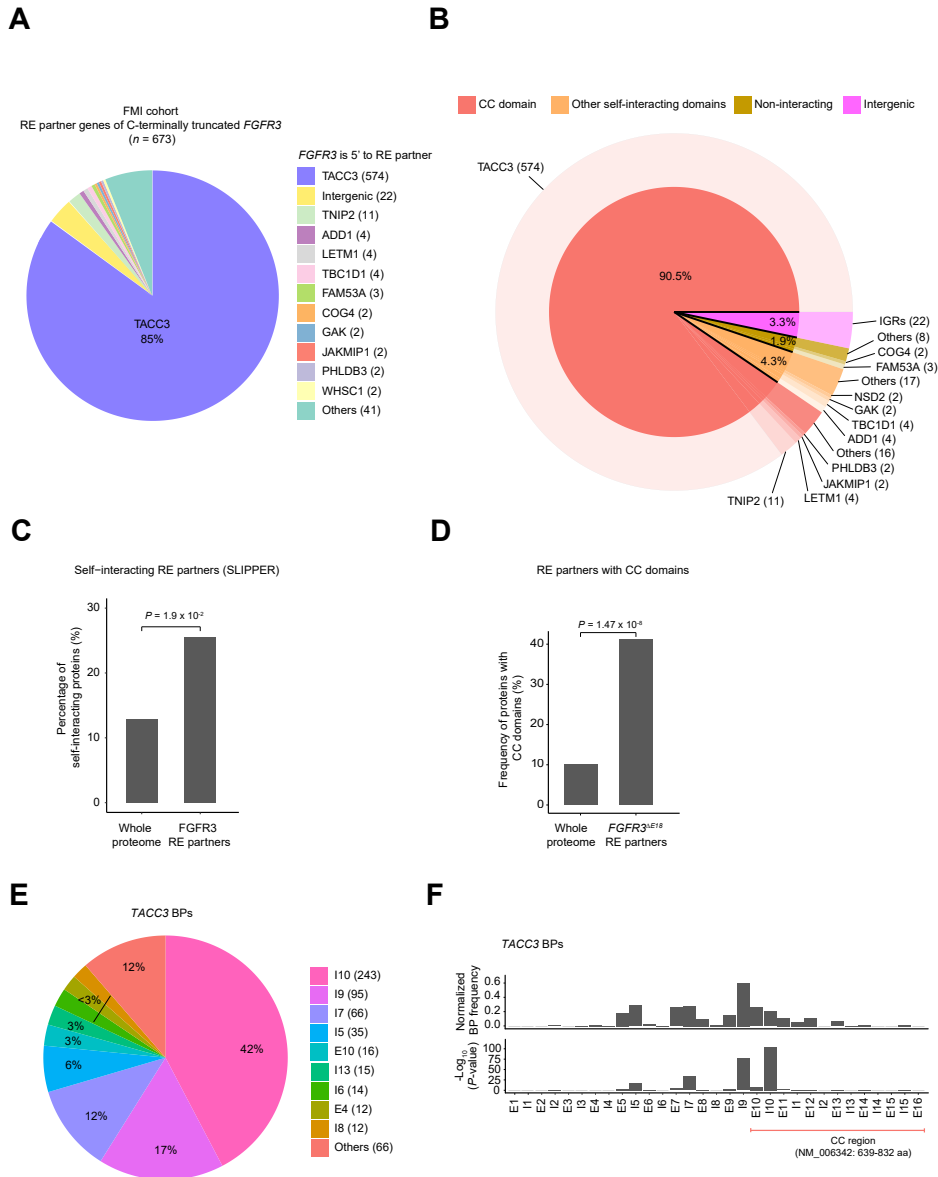


Figure 3. *FGFR3* RE partners are enriched for proteins with self-interacting capacity.

A, Recurrence of RE partners of C-terminally truncated *FGFR3* identified in the FMI cohort (*n*=673 tumors). **B**, Recurrence of *FGFR3* RE partners from **(A)** grouped by the presence or absence of self-interacting domains. Full list of RE partners and their corresponding self-interacting domains is disclosed in Supplementary Table S2. IGRs, intergenic regions; CC, coiled-coil. **C**, Frequency comparison between *FGFR3* RE partners that contain self-interacting domains and result in E18 truncation, based on the SLIPPER algorithm [27]. **D**, Frequency of proteins with CC domains among the E18-truncated *FGFR3* (21 out of 51) RE partners identified in the FMI cohort versus the human proteome (1993 out of 18975). *P* values were calculated with a two-tailed Fisher's exact test. **E**, Pie chart displaying the frequency of BPs in each exon/intron of *TACC3* identified in the 574 *FGFR3*^{ΔE18}-*TACC3* fusions found in the FMI cohort. **F**, Normalized frequency (top panel) and enrichment significance (*P* values, bottom panel) of *TACC3* genomic RE BPs identified in **(E)**. BP frequency was normalized by the kilobase of feature (exon/intron) length and total number of BPs. *P* values were calculated with a binomial test. Exons encompassing the CC region are indicated.

as evidenced by phosphorylation of FGFR3, RPS6, EIF4E, EIF4EBP1, EIF4B and stabilization of MYC. In contrast, all other *Fgfr3* variants tested showed no baseline activity (**Fig. 4B** and **C**). Stimulation with FGF-8b enhanced FGFR signaling via the mTOR axis in all FGFR3 variants (**Fig. 4B** and **C**). Similarly, activation of MAPK signaling hallmarked by the phosphorylation of ERK1/2 was predominantly induced through FGF-8b (**Fig. 4B** and **C**). Ligand-independent phosphorylation of STAT1, STAT3, FRS2 and PLC γ occurred exclusively in *Fgfr3^{ΔE18}-Tacc3*-expressing cells and was further boosted upon stimulation with FGF-8b (**Fig. 4C** and **D**). Treatment of cells with the pan-FGFRi fexagratinib [26] as well as the clinically approved pan-FGFRi erdafitinib and futibatinib [4], suppressed FGFR3^{ΔE18}-TACC3 signaling (**Fig. 4B-D**; Supplementary Fig. S2B-S2D), indicating that the constitutive activity of the FGFR3^{ΔE18}-TACC3 fusion relies on canonical FGFR3 kinase domain function.

Next, we assessed the transforming capacity of FGFR3 fusion variants *in vitro*. First, we made use of a 3D soft agar growth assay to measure cellular transformation in an anchorage-independent manner [28]. In line with its enhanced baseline signaling capacity, *Fgfr3^{ΔE18}-Tacc3*-expressing NMuMG cells showed considerable soft agar colony formation, as opposed to *GFP*, *Fgfr3^{FL}*, *Fgfr3^{ΔE18}* and *Fgfr3^{FL}-Tacc3* cells, which showed only marginal anchorage-independent growth (**Fig. 4E**; Supplementary Fig. S2E). These results were in striking contrast to what we previously observed for FGFR2, where loss of the C-terminal tail alone was sufficient to promote transformation and anchorage-independent growth of NMuMG cells [5]. We next expanded the functional evaluation to the Ba/F3 murine pro-B cell line that depends on IL-3 for survival and proliferation [29]. Consistent with our findings in NMuMG cells, only *Fgfr3^{ΔE18}-Tacc3*-expressing Ba/F3 cells were able to proliferate during IL-3 withdrawal (**Fig. 4F**; Supplementary Fig. S2F and S2G). The kinase-dead *Fgfr3^{ΔE18}-K502R-Tacc3* variant failed to promote IL-3-independent Ba/F3 cellular growth (**Fig. 4F**; Supplementary Fig. S2F and S2G) or activate oncogenic signaling in NMuMG cells with and without FGF-8b stimulation (Supplementary Fig. S2H and S2I), underscoring the critical role of kinase domain activity in driving oncogenicity of the FGFR3^{ΔE18}-TACC3 fusion [13,30].

Together, these data established that loss of the FGFR3 C-terminal tail is required yet insufficient to promote oncogenic signaling and anchorage- or IL-3-independent growth. In contrast, C-terminal truncation and fusion to a partner containing a self-interacting domain, such as TACC3, is required for full oncogenic transformation of FGFR3.

Impaired internalization does not define FGFR3^{ΔE18}-TACC3-driven transformation

Receptor tyrosine kinase (RTK) signaling is often negatively regulated through ligand-induced receptor internalization and subsequent lysosomal degradation, which limits signaling duration and intensity [31,32]. To assess whether impaired internalization contributes to FGFR3^{ΔE18}-TACC3 oncogenicity, we evaluated receptor trafficking using HALO-tagged FGFR3 variants and the SPIDA assay [33]. NMuMG cells expressing HALO-*Fgfr3* variants (Supplementary Fig. S3A) were sequentially labeled with a cell-impermeable

fluorescent HaloTag ligand (AF660-HTL) followed by a cell-permeable fluorescent ligand (TMR-HTL), and fluorescence intensities of both ligands were subsequently quantified by FACS (**Fig. 4G**; Supplementary Fig. S3B). The relative abundance of FGFR3 at the plasma membrane versus the intracellular compartment was determined over time during FGF-8b stimulation using the SPIDA quantification framework [33] (**Fig. 4H and I**). As expected, HALO-FGFR3^{FL} underwent ligand-induced internalization upon FGF-8b stimulation, accompanied by only a modest increase in intracellular signal, consistent with receptor degradation following endocytosis (**Fig. 4G-I**). In contrast, the kinase-dead HALO-FGFR3^{FL}-K502R variant failed to internalize (**Fig. 4G and H**), underscoring the requirement for kinase activity in this process. Notably, none of the other HALO-tagged FGFR3 variants, including the FGFR3^{ΔE18}-TACC3 fusion, displayed internalization upon ligand stimulation (**Fig. 4G and H**). This aligns with the fact that YXXΦ clathrin-mediated endocytosis motif [31] is located within E18 and therefore absent in truncated variants. While impaired internalization may contribute to sustained surface accumulation and aberrant signaling, it is not sufficient to drive transformation, as FGFR3^{ΔE18} alone – despite lacking the capacity to internalize (**Fig. 4G and H**) – did not induce oncogenic signaling (**Fig. 4B-D**; Supplementary Fig. S2B-S2D), provoke anchorage-independent 3D growth in NMuMG cells (**Fig. 4E**; Supplementary Fig. S2E), or confer IL-3-independent proliferation in Ba/F3 cells (**Fig. 4F**).

***Fgfr3*^{ΔE18}-*Tacc3* fusions drive mouse lung and mammary tumorigenesis**

Our *in vitro* findings prompted us to evaluate the *in vivo* oncogenic capacity of the different *Fgfr3* variants using somatic mouse modeling. To this end, we employed two different lentiviral delivery systems to the mouse lung and mammary gland, representing two human organ sites where *FGFR3*^{ΔE18}-*TACC3* fusions are recurrently detected in cancer (**Fig. 1A and D**; Supplementary Data S1 and S2).

FGFR3^{ΔE18}-*TACC3* fusions are found in 0.14-2.1% of non-small cell lung cancers (NSCLC) [7,8,19]. The FMI cohort contained 87 NSCLC cases with *FGFR3*^{ΔE18}-*TACC3* REs, of which 48 were diagnosed as lung squamous cell carcinoma (LUSC) and 39 as lung adenocarcinoma (LUAD) (**Fig. 1A**; Supplementary Data S1). The HMF cohort contained one NSCLC case with an *FGFR3*^{ΔE18}-*TACC3* fusion (**Fig. 1D and 2E**; Supplementary Data S2). *TP53* missense mutations are the most prevalent co-alterations in *FGFR3*-mutated lung cancer [8,18]. Consistently, 87% of all the LUSC/LUAD tumors in the FMI cohort with *FGFR3* alterations presented co-occurring *TP53* mutations ($n=107$ out of 123; Supplementary Fig. S4A).

To model the oncogenic potential of the different *Fgfr3* variants in the lung, we performed intratracheal intubation and delivery of lentiviral vectors encoding *Fgfr3*-P2A-Cre into wild-type and *Trp53*^{F/F} mice (**Fig. 5A**). In wild-type mice, *Fgfr3*^{ΔE18}-*Tacc3*-P2A-Cre expression resulted in lung tumor formation in a fraction of the intubated animals, as opposed to no tumor formation in the *Fgfr3*^{FL}-P2A-Cre, *Fgfr3*^{ΔE18}-P2A-Cre or *Fgfr3*^{FL}-*Tacc3*-P2A-Cre cohorts (**Fig. 5B**). Histopathological examination of the lungs intubated with *Fgfr3*^{ΔE18}-*Tacc3*-P2A-Cre

Genomic DNA (exons)

Homo sapiens (NM_000142.5)

Homo sapiens (NM_006342)

Protein

Experimental variants (Mus musculus)

Legend: ▼ Point mutation

Western blot analysis of protein expression and phosphorylation in cells transfected with GFP, FGF-8b, or FGF-8b-Tac3. The blots show levels of FGF-8b, p-FGF-8b, p-RPS6, RPS6, p-EIF4E, EIF4E, EIF4EBP1, p-EIF4B, EIF4B, MYC, p-ERK1, p-ERK2, ERK1, ERK2, and β-Actin. Molecular weight markers are indicated on the right in kDa.

The heatmap displays the Log₂(intensity) expression levels for 10 proteins across 16 conditions. The color scale ranges from -3 (blue) to 2 (red). The proteins are listed on the y-axis, and the conditions are listed on the x-axis.

	GFP	FoxO1	FoxO1+Tg3	FoxO1+Tg3	Tg3	Tg3	Tg3	Tg3	Tg3	Tg3	Tg3	Tg3	Tg3	Tg3	Tg3	Tg3	Tg3
FGF-8b	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Agmatrinib	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
p-FGR																	
p-RP96																	
p-EIF4E																	
eIF4EBP1																	
p-EIF4B																	
ERK1/2																	
p-STAT1																	
p-STAT3																	
p-FRS2																	
p-PLCγ																	
p-PKCα																	

Log₂(intensity)

Western blot analysis of protein levels in various tissues. The blot shows bands for AT1, AT3, S2S, S2L, Cy, Ca, and actin across different conditions. Molecular weight markers are indicated on the right.

3D colony count / well
NMuMG cells

2500
2000
1500
1000
500
0

GFP *Fgfr3^{G12R}* *Fgfr3^{G12R}+Tacc3* *Fgfr3^{G12R}+Tacc3^Δ*

$P < 0.0001$
 $P < 0.0001$
 $P < 0.0001$
 $P > 0.9948$

Figure 3: Bar/F3 viable cells/ml (x10⁶) vs Days post IL-3 withdrawal. The graph shows four data series: *Fgfr3*^{fl} (blue diamonds), *Fgfr3*^{fl};Tacc3 (green squares), *Fgfr3*^{fl};K502R-Tacc3 (red circles), and *Fgfr3*^{fl};E18 (black triangles). The y-axis ranges from 0 to 7 (x10⁶ cells/ml). The x-axis ranges from 0 to 15 days post IL-3 withdrawal. A dotted line is at 0.5 x10⁶ cells/ml. The *Fgfr3*^{fl};K502R-Tacc3 series shows a significant increase starting around day 10, reaching approximately 6.5 x10⁶ cells/ml by day 15. The other series remain below 0.5 x10⁶ cells/ml throughout the 15-day period.

Cell Line	Surface HALO-FGR3 ratio 30 min / 0 min FGF-8p
HALO-Fgr3 ^{+/+}	~0.75
HALO-Fgr3 ^{-/-} -K502R	~1.00
HALO-Fgr3 ^{-/-}	~1.00
HALO-Fgr3 ^{-/-} -Tacc3	~0.95
HALO-Fgr3 ^{-/-} -K502R-Tacc3	~1.05
HALO-Fgr3 ^{-/-} -Tacc3	~1.00

Genotype	Internal HALO-FGF3R ratio (approx.)
HALO-Fgf3 ^{+/+}	1.00
HALO-Fgf3 ^{+/+} -K502R	0.98
HALO-Fgf3 ^{+/+} -Tabc3	1.00
HALO-Fgf3 ^{+/+} -K502R-Tabc3	0.98
HALO-Fgf3 ^{-/-} -Tabc3	0.95

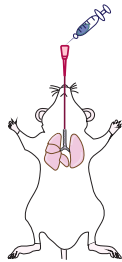
Figure 4. *Fgfr3^{ΔE18}-Tacc3* fusion induces oncogenic signaling and *in vitro* transformation.

A, Schematic representation of the genomic DNA and the resulting mature peptide of human *FGFR3* and *TACC3*. Protein domains of cloned mouse *Fgfr3* variants are indicated. Complete nucleotide and AA sequences for *Fgfr3^{ΔE18}-Tacc3* are disclosed in Supplementary Methods S1. **B**, Western blots showing expression and phosphorylation of indicated proteins in 24 hr serum-starved NMuMG cells stably expressing *GFP*, *Fgfr3^{FL}*, *Fgfr3^{ΔE18}*, *Fgfr3^{FL}-Tacc3*, or *Fgfr3^{ΔE18}-Tacc3* and treated for 10 min with 20 ng/ml FGF-8b and/or for 3 hr with 1 μM fexagratinib. β-Actin was run as sample processing control and blots were stained with Ponceau S to ensure equal loading of total protein. Molecular weights of protein marker bands are indicated in kilodalton (kDa). **C**, Quantifications of indicated phosphoprotein band intensities based on Western blots in (**B**, **D**). Phospho band intensities were normalized to their individual lane backgrounds and to their corresponding total protein band intensities and β-Actin bands. **D**, Western blots generated as in (**B**), showing expression and phosphorylation of additional proteins. β-Actin is duplicated from (**B**). **E**, Colony quantifications of 3D soft agar colony formation assays using NMuMG cells stably expressing *GFP*, *Fgfr3^{FL}*, *Fgfr3^{ΔE18}*, *Fgfr3^{FL}-Tacc3*, or *Fgfr3^{ΔE18}-Tacc3* and grown for 20 days. Data are represented as mean ± SD of *n*=30 independent replica from 5 individual experiments. **F**, Growth rates of parental and Ba/F3 cells stably expressing *GFP*, *Fgfr3^{FL}*, *Fgfr3^{ΔE18}*, *Fgfr3^{FL}-Tacc3*, *Fgfr3^{ΔE18}-Tacc3*, or *Fgfr3^{ΔE18}-K502R-Tacc3* upon IL-3 withdrawal. Data are represented as mean ± SEM of *n*=3 independent replica from 1 representative experiment. Assay was performed at least 3 independent times. **G-I**, FACS to quantify HALO-FGFR3 internalization in serum-deprived NMuMG cells expressing indicated HALO-*Fgfr3* variants using the SPIDA assay [33]. Cells were treated with 20 ng/ml FGF-8b for 5, 10 or 30 min and sequentially labeled with cell membrane impermeable HaloTag ligand (HTL)-AF660 and with permeable HTL-TMR. AF660-HALO-FGFR3 (**G**) and TMR-HALO-FGFR3 MFIs were used to calculate ratios of surface (**H**) and internal (**I**) FGFR3 levels in cells stimulated with FGF-8b for 30 min versus unstimulated cells. Ratios were calculated based on the data displayed in Supplementary Fig. S3B. Histograms in (**G**) show 1 representative replica of data quantified in (**H**, **I**), which are represented as mean ± SD of HALO-*Fgfr3^{FL}*, HALO-*Fgfr3^{FL}-K502R*, HALO-*Fgfr3^{ΔE18}*, HALO-*Fgfr3^{FL}-Tacc3*, HALO-*Fgfr3^{ΔE18}-Tacc3*, FGF-8b, 0 min, 30 min, *n*=4; *Fgfr3^{ΔE18}-K502R-Tacc3*, *n*=3 independent replica. *P* values in (**E**, **H**, **I**) were calculated with ordinary one-way ANOVA and FDR multiple testing corrections.

revealed a high prevalence of adenocarcinomas, papillary carcinomas and mesotheliomas, in contrast to the mainly healthy tissue or low-grade adenoma lesions observed in lungs of the other *Fgfr3* cohorts (Supplementary Fig. S4B). The transforming capacity of *Fgfr3^{ΔE18}-Tacc3-P2A-Cre* was markedly exacerbated when lungs of *Trp53^{F/F}* animals were intubated (**Fig. 5C**), with >90% of lungs presenting high-grade lesions including mesotheliomas and mixed mesothelioma and adenocarcinoma histotypes (Supplementary Fig. S4B). In contrast, expression of *Fgfr3^{FL}-P2A-Cre*, *Fgfr3^{ΔE18}-P2A-Cre* or *Fgfr3^{FL}-Tacc3-P2A-Cre* in the lungs of *Trp53^{F/F}* mice only marginally increased tumor formation as compared to *Cre*-injected animals (**Fig. 5C**; Supplementary Fig. S4B). In sum, in both wild-type and *Trp53^{F/F}* mice, only the *Fgfr3^{ΔE18}-Tacc3* fusion had the competence to effectively drive lung tumorigenesis.

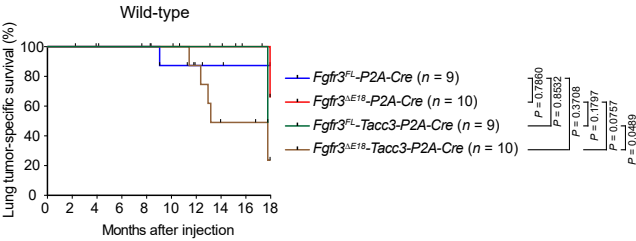
Apart from lung cancers, *FGFR3^{ΔE18}-TACC3* fusions have also been reported in cancers of the breast [23–25]. We identified 25 breast cancer cases harboring *FGFR3^{ΔE18}-TACC3* REs in the FMI cohort (**Fig. 1A**; Supplementary Data S1) and 3 cases in the HMF cohort (**Fig. 1D** and **2B**; Supplementary Data S2). In the context of breast cancer, *FGFR3-TACC3* fusions have been assessed *in vitro* [23,24], yet, to our best knowledge, no *in vivo* studies have been conducted. We evaluated the oncogenicity of *Fgfr3* variants in mouse mammary tumor models representing different breast cancer subtypes, including invasive lobular carcinoma (ILC) that is characterized by the loss of E-cadherin, which is encoded by *CDH1* [34,35].

A

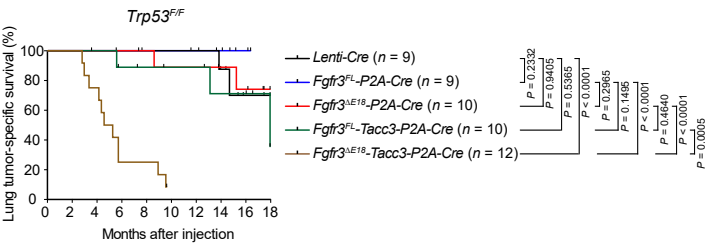


Intratracheal lentiviral injections

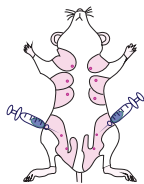
B



C

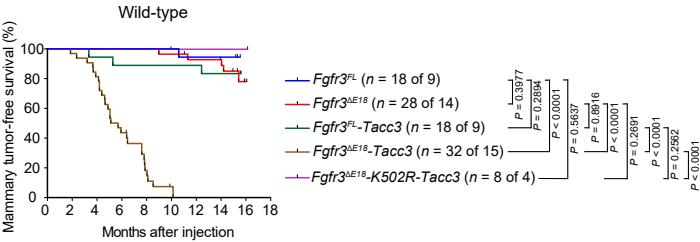


D



Intraductal lentiviral injections

E



F

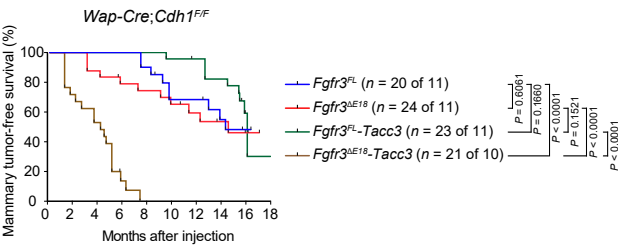


Figure 5. *Fgfr3^{ΔE18}-Tacc3* fusions drive mouse lung and mammary tumorigenesis.

A, Overview of the somatic mouse model for lung cancer formation consisting of intratracheal intubation and delivery of lentiviral vectors encoding the *Fgfr3* variants of interest. **B**, Kaplan-Meier curves showing lung tumor-specific survival of female wild-type mice administered with lentiviruses encoding the indicated *Fgfr3* variants via intratracheal intubation. *Fgfr3^{FL}-P2A-Cre*, *n*=9; *Fgfr3^{ΔE18}-P2A-Cre*, *n*=10; *Fgfr3^{FL}-Tacc3 -P2A-Cre*, *n*=9; *Fgfr3^{ΔE18}-Tacc3-P2A-Cre*, *n*=10 mice. **C**, Kaplan-Meier curves showing lung tumor-specific survival of female *Trp53^{FF}* mice administered with lentiviruses encoding the indicated *Fgfr3* variants via intratracheal intubation. *Lenti-Cre* control, *n*=9; *Fgfr3^{FL}-P2A-Cre*, *n*=9; *Fgfr3^{ΔE18}-P2A-Cre*, *n*=10; *Fgfr3^{FL}-Tacc3 -P2A-Cre*, *n*=10; *Fgfr3^{ΔE18}-Tacc3 -P2A-Cre*, *n*=12 mice. **D**, Overview of the somatic mouse model for mammary gland-specific overexpression of indicated *Fgfr3* variants via intraductal injection of lentiviruses. Two glands per mouse were injected. **E**, Kaplan-Meier curves showing mammary tumor-free survival of female wild-type mice intraductally injected with lentiviruses encoding indicated *Fgfr3* variants. *Fgfr3^{FL}*, *n*=18 of 9; *Fgfr3^{ΔE18}*, *n*=28 of 14; *Fgfr3^{FL}-Tacc3*, *n*=18 of 9; *Fgfr3^{ΔE18}-Tacc3*, *n*=32 of 15; *Fgfr3^{ΔE18}-K502R-Tacc3*, *n*=8 injected mammary glands of 4 mice. **F**, Kaplan-Meier curves showing mammary tumor-free survival of female *Wap-Cre;Cdh1^{FF}* mice intraductally injected with lentiviruses encoding indicated *Fgfr3* variants. *Fgfr3^{FL}*, *n*=20 of 11; *Fgfr3^{ΔE18}*, *n*=24 of 11; *Fgfr3^{FL}-Tacc3*, *n*=23 of 11; *Fgfr3^{ΔE18}-Tacc3*, *n*=21 injected mammary glands of 10 mice. *P* values were calculated with log-rank (Mantel-Cox) tests (**B**, **C**, **E**, **F**).

5

Using intraductal injections (**Fig. 5D**; refs. [5,36]), we somatically delivered lentiviruses containing *Fgfr3* variants to the mammary glands of female wild-type (**Fig. 5E**; Supplementary Fig. S5A) and *Wap-Cre;Cdh1^{FF}* (**Fig. 5F**; Supplementary Fig. S5B) mice. Mammary glands injected with *Fgfr3^{FL}*, *Fgfr3^{ΔE18}*, *Fgfr3^{FL}-Tacc3* or *Fgfr3^{ΔE18}-K502R-Tacc3* variants displayed no (**Fig. 5E**) or slow (**Fig. 5F**) tumorigenesis, whereas those injected with the *Fgfr3^{ΔE18}-Tacc3* fusion showed marked tumor formation in both genotypes (**Fig. 5E** and **F**). Reminiscent of the histopathological observations in lung, mammary glands of wild-type mice injected with *Fgfr3^{FL}*, *Fgfr3^{ΔE18}* and *Fgfr3^{FL}-Tacc3* variants revealed mostly no or low-grade lesions, whereas glands injected with *Fgfr3^{ΔE18}-Tacc3* contained high-grade tumors, such as adenocarcinomas and squamous metaplastic carcinomas (Supplementary Fig. S5C). Similarly, mammary glands of *Wap-Cre;Cdh1^{FF}* mice injected with *Fgfr3^{FL}*, *Fgfr3^{ΔE18}* and *Fgfr3^{FL}-Tacc3* variants mostly revealed healthy tissue, low-grade lesions, or classic ILCs. Notably, the low-grade lesions and classic ILCs observed in the *Wap-Cre;Cdh1^{FF}* cohorts likely represent spontaneous tumor events owing to the ILC-predisposition of this strain [5,37]. In contrast, mammary glands injected with *Fgfr3^{ΔE18}-Tacc3* displayed tumors of higher grade, especially solid ILCs (Supplementary Fig. S5C).

Together, our *in vivo* modeling efforts uncovered that the *Fgfr3^{ΔE18}-Tacc3* fusion is a potent driver in mouse models of both lung and breast tumorigenesis, whereas the loss of the C-terminal tail of FGFR3 on its own was insufficient to induce tumor formation. Finally, we explored driver gene alterations diagnosed by FMI's oncogenomic profiling and their incidence in cancers with *FGFR3^{ΔE18}* alterations. Tumors that harbored *FGFR3* E18 truncations devoid of a partner gene (i.e., E18-truncating mutations, intergenic or internal REs, or partial amplifications) showed enrichment for alterations in *KRAS*, *ERBB2*, *MYC*, *PIK3CA*, and *PTEN*, as compared to tumors that contained a *FGFR3^{ΔE18}* RE/fusion with a partner gene (Supplementary Fig. S4A and S5D-S5G), suggesting that additional tumor drivers may be required to enable oncogenicity of *FGFR3^{ΔE18}* alterations without a partner gene.

Self-interacting capacity but not mitotic function of TACC3 is required for FGFR3^{ΔE18}-TACC3 oncogenicity

The high prevalence of self-interacting domains among all FGFR3 fusion partners (**Fig. 3B-D**) and the preservation of the two TACC3-CC domains in most of the FGFR3^{ΔE18}-TACC3 fusions (**Fig. 3E** and **F**), suggested that the presence of a self-interacting domain is a prerequisite for the oncogenic activity of FGFR3^{ΔE18}. To assess the relevance of the TACC3-CC domains for the oncogenic competence of FGFR3^{ΔE18}-TACC3, we substituted key AA residues critical for the integrity of either of the two CC domains, as identified by the Paircoil2 algorithm [38], with prolines (**Fig. 6A**; Supplementary Methods S2 and S3). Intraductal injection of *Fgfr3*^{ΔE18}-*Tacc3*^{CC1-Pro} or *Fgfr3*^{ΔE18}-*Tacc3*^{CC2-Pro} variants that encode FGFR3^{ΔE18}-TACC3 derivatives predicted to lack dimerization capacity, did not induce tumor formation in wild-type mice, in contrast to parental FGFR3^{ΔE18}-TACC3 that retains dimerization and drives marked tumorigenesis (**Fig. 6B**; Supplementary Fig. S5H). These findings underscored the requirement of functional TACC3-CC domains for full oncogenic capacity of FGFR3^{ΔE18}-TACC3.

TACC3 facilitates centrosome-dependent microtubule assembly [39], mediates kinetochore-microtubule attachment [40], and ensures spindle-dependent chromosome alignment [41–43] throughout mitosis. These functions require phosphorylation of TACC3 at S558 by Aurora-kinase A (AURKA), followed by binding of clathrin at L566/L567 [39,44–48]. Both the AURKA phosphorylation site as well as the clathrin binding site are located to a structural unit that is encoded by E8 of human TACC3 and are retained in the clinical *FGFR3*^{ΔE18}-*TACC3*^{E8-E16} fusion and in the corresponding murine *Fgfr3*^{ΔE18}-*Tacc3* used in this study. To test whether the two TACC3 regulatory sites are critical for FGFR3^{ΔE18}-TACC3 oncogenicity, we generated the *Fgfr3*^{ΔE18}-*Tacc3*^{S764A} and the *Fgfr3*^{ΔE18}-*Tacc3*^{L772A/L773A} variants with mutated AURKA phosphorylation or clathrin binding sites, respectively (**Fig. 6A**; Supplementary Methods S4 and S5). We also generated the *Fgfr3*^{ΔE18}-*Tacc3*^{E10-E16} variant that corresponds to the human *FGFR3*^{ΔE18}-*TACC3*^{E11-E16} fusion recurrently observed in clinical samples (**Fig. 3E** and **F**) and that lacks both the AURKA phosphorylation and the clathrin binding sites in TACC3 (**Fig. 6A**; Supplementary Methods S6). Intraductal injection of the *Fgfr3*^{ΔE18}-*Tacc3*^{E10-E16} fusion into wild-type mice modestly increased tumor latency as compared to the *Fgfr3*^{ΔE18}-*Tacc3* cohort, and similarly to that of *Fgfr3*^{ΔE18}-*Tacc3*^{S764A} and the *Fgfr3*^{ΔE18}-*Tacc3*^{L772A/L773A} (**Fig. 6C**). However, there was no statistically significant difference in the overall survival among the four mouse cohorts (Supplementary Fig. S5I), suggesting that the assessed TACC3-regulatory sites are negligible for the oncogenicity of the FGFR3^{ΔE18}-TACC3 fusion, in agreement with previous *in vitro* work [49].

Oncogenicity of *Fgfr3^{ΔE18}-Add1*, *Fgfr3^{ΔE18}-Tnip2*, and *Fgfr3^{ΔE18}-Jakmip1* fusions

To assess whether FGFR3^{ΔE18} fusions other than FGFR3^{ΔE18}-TACC3 had oncogenic capacity, we functionally assessed three previously uncharacterized *FGFR3* fusions identified in the FMI and HMF cohorts (**Fig. 1A, 1D, 2C, 3A and 3B**; Supplementary Data S1 and S2). These were *FGFR3^{ΔE18}-ADD1^{E3-E15}* (6 cases), *FGFR3^{ΔE18}-TNIP2^{E2-E6}* (11 cases), and *FGFR3^{ΔE18}-JAKMIP1^{E4-E13}* (4 cases; **Fig. 7A**). Notably, all three fusion proteins contained self-dimerizing domains from ADD1, TNIP2 and JAKMIP1, respectively (**Fig. 3B and 7A**).

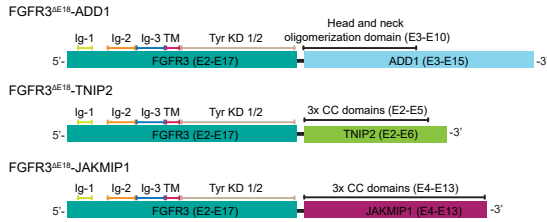
Stable expression of the murine orthologs *Fgfr3^{ΔE18}-Add1*, *Fgfr3^{ΔE18}-Tnip2*, and *Fgfr3^{ΔE18}-Jakmip1* in NMuMG cells (Supplementary Fig. S2A and S6A) led to activation of downstream FGFR3 signaling under serum-deprived conditions, which was abrogated by treatment with fexagratinib (**Fig. 7B**; Supplementary Fig. S6A). Similar to *Fgfr3^{ΔE18}-Tacc3* and hotspot mutant *Fgfr3^{K644E}* (corresponding to human *FGFR3^{K650E}*; ref. [49]), all three fusions triggered robust phosphorylation of FGFR3, EIF4EBP1, ERK1/2 and FRS2, consistent with activation of FGFR, mTOR, and MAPK signaling. *Fgfr3^{ΔE18}-Add1* additionally induced strong phosphorylation of RPS6, STAT1, and STAT3, resembling levels of *Fgfr3^{ΔE18}-Tacc3*.

FGFR3 fusions involving *ADD1*, *TNIP2* and *JAKMIP1* were predominantly identified in urothelial cancer (**Fig. 1A and 1D**; Supplementary Data S1 and S2). Therefore, we made use of the human bladder cancer 5637 cell line that harbors an activating *ERBB2^{S310F}* mutation and EGFR overexpression and is sensitive to the dual ERBB2/EGFR tyrosine kinase inhibitors (TKIs) afatinib and dacomitinib (**Fig. 7C and 7D**; Supplementary Fig. S6B-S6E). Expression of *Fgfr3^{ΔE18}-Tacc3*, *Fgfr3^{ΔE18}-Add1*, *Fgfr3^{ΔE18}-Tnip2*, *Fgfr3^{ΔE18}-Jakmip1*, as well as *Fgfr3^{K644E}* conferred marked resistance to both TKIs, and sensitivity to afatinib was restored when cells were co-treated with the FGFRi fexagratinib (**Fig. 7C-F**; Supplementary Fig. S6D-S6G).

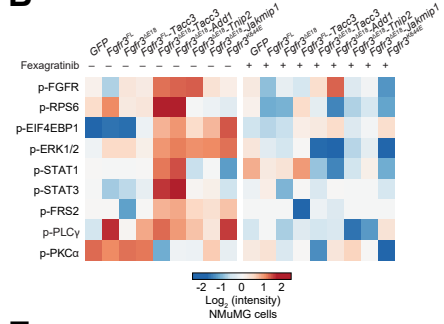
Figure 7. *In vitro* activity and *in vivo* tumorigenicity of *Fgfr3^{ΔE18}-Add1*, *Fgfr3^{ΔE18}-Tnip2* and *Fgfr3^{ΔE18}-Jakmip1* gene fusions.

A, Schematic representation of the FGFR3^{ΔE18}-ADD1, FGFR3^{ΔE18}-TNIP2 and FGFR3^{ΔE18}-JAKMIP1 fusions detected in the FMI and HMF cohorts. Complete nucleotide and AA sequences for the corresponding murine experimental constructs used are disclosed in Supplementary Methods S7-S9. **B**, Quantifications of indicated phosphoprotein band intensities based on Western blots in Supplementary Fig. S6A, where 24 hr serum-starved NMuMG cells stably expressing GFP, *Fgfr3^{ΔE18}*, *Fgfr3^{ΔE18}-Tacc3*, *Fgfr3^{ΔE18}-Add1*, *Fgfr3^{ΔE18}-Tnip2*, *Fgfr3^{ΔE18}-Jakmip1*, or *Fgfr3^{K644E}* were treated for 3 hr with either vehicle or 1 μM fexagratinib. Phospho band intensities were normalized to their individual lane backgrounds and to their corresponding total protein band intensities and β-actin bands, and unit variance scaling is applied to the rows. **C**, Dose-response curves of 5637 bladder cancer cells, either untransduced (parental) or stably expressing the indicated *Fgfr3* variants via lentiviral transduction, treated for 3 days with different concentrations of afatinib. Cell viability was determined using CellTiter-Blue. Data were normalized to untreated conditions within each cell line and are presented as mean ± SD of n=2 independent replica from 1 representative experiment. Assays were performed at least 2 independent times. **D**, Representative well images of (**C**), obtained after fixing and staining the cells with crystal violet. **E**, Dose-response curves of 5637 bladder cancer cells as in (**C**), treated for 3 days with different concentrations of afatinib in combination with 2.5 μM fexagratinib, corresponding to the inhibitory concentration (IC)10-IC20. **F**, Representative well images of (**E**). **G**, Kaplan-Meier curves showing mammary tumor-free survival of female wild-type mice intraductally injected with lentiviruses encoding indicated *Fgfr3* variants. *Fgfr3^{ΔE18}*, n=28 of 14; *Fgfr3^{ΔE18}-Tacc3*, n=32 of 15; *Fgfr3^{ΔE18}-Add1*, n=28 of 14; *Fgfr3^{ΔE18}-Tnip2*, n=28 of 14; *Fgfr3^{ΔE18}-Jakmip1*, n=24 injected mammary glands of 12 mice. The *Fgfr3^{ΔE18}* and the *Fgfr3^{ΔE18}-Tacc3* curves are duplicated from (**Fig. 5E**). P values were calculated with log-rank (Mantel-Cox) tests.

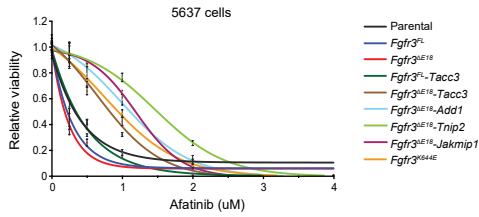
A



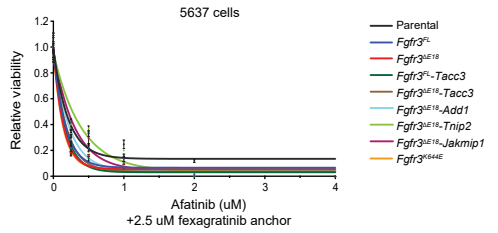
B



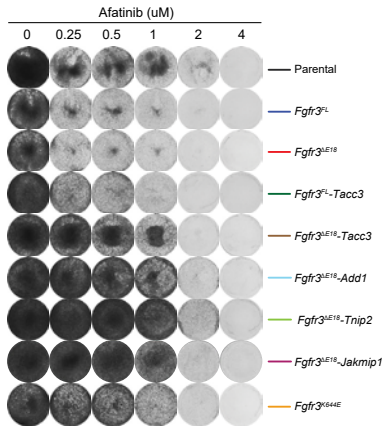
C



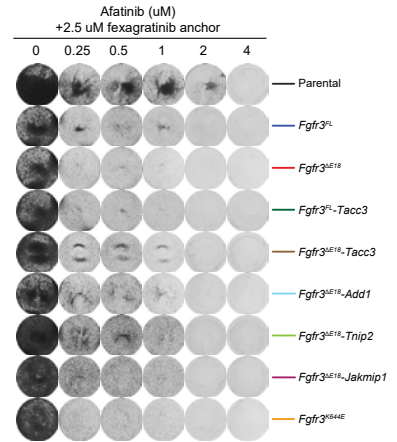
E



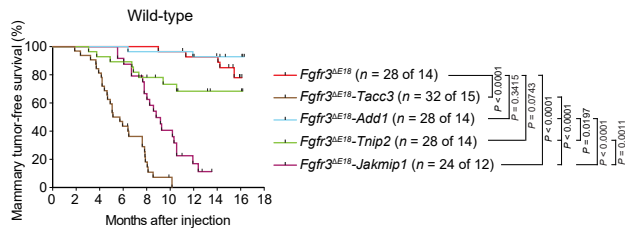
D



F



G



Intraductal injection of *Fgfr3^{ΔE18}-Add1*, *Fgfr3^{ΔE18}-Tnip2* and *Fgfr3^{ΔE18}-Jakmip1* fusions into the mammary glands of wild-type mice revealed that *Fgfr3^{ΔE18}-Jakmip1* robustly induced mammary tumorigenesis, comparable to *Fgfr3^{ΔE18}-Tacc3* (**Fig. 7G**; Supplementary Fig. S6H). In contrast, *Fgfr3^{ΔE18}-Tnip2* induced tumors in only a fraction of the injected mammary glands and *Fgfr3^{ΔE18}-Add1* completely failed to induce mammary tumorigenesis despite displaying considerable activity *in vitro* (**Fig. 7B-D** and **7G**; Supplementary Fig. S6). These differences in *in vivo* oncogenicity in wild-type mouse mammary epithelium suggest that the oncogenic competence of FGFR3 fusions may be affected by the mutational context as well as the tissue of origin – *ADD1* fusions were mostly found in urothelial cancer. In sum, our results extend the tumor- and drug resistance-driving properties of FGFR3^{ΔE18} fusions beyond TACC3, underscoring the potential of FGFR-targeted therapies in these settings.

***Fgfr3^{ΔE18}-Tacc3*-driven mammary tumors are sensitive to FGFRi**

Finally, we tested whether mammary tumors driven by *Fgfr3^{ΔE18}-Tacc3* would be sensitive to FGFR-targeted therapy, by orthotopically transplanting tumor pieces into the mammary fat pad of syngeneic recipient female mice. When tumors reached a size of 62.5 mm³, animals were subjected to an intermittent dosing regimen using 12.5 mg/kg/day of fexagratinib (**Fig. 8A**; refs. [26,50]). Fexagratinib treatment in mice engrafted with *Fgfr3^{ΔE18}-Tacc3*-induced tumors derived from either wild-type (**Fig. 8B**; Supplementary Fig. S7A) or *Wap-Cre;Cdh1^{F/F}* (**Fig. 8C**; Supplementary Fig. S7B) donors was efficacious by significantly suppressing tumor growth as compared to the respective vehicle-treated controls, although complete tumor regression was not achieved. We next subjected tumor-bearing mice to a PK/PD study using a daily 12.5 mg/kg dose of fexagratinib for 2 or 5 consecutive days. Analysis of plasma and tumor samples 1 hr post dosing confirmed high fexagratinib exposure, with plasma and intratumoral concentrations reaching several orders of magnitude above the reported *in vitro* IC₅₀ of ~1.8 nmol/L for FGFR3 (Supplementary Fig. S7C-S7E; ref. [26]), ruling out insufficient drug exposure as a limiting factor in the efficacy study. PD biomarker assessment demonstrated substantial on-target FGFR3 inhibition as well as FGFR3 signaling pathway suppression (Supplementary Fig. S7F and S7G), confirming that the dosing regimen achieved effective target engagement. To further investigate the lack of complete tumor eradication upon fexagratinib treatment, we analyzed protein lysates from endpoint tumors in both wild-type and *Wap-Cre;Cdh1^{F/F}* donors that had progressed during fexagratinib treatment (Supplementary Fig. S7H-S7K). Western blot analysis revealed reactivation of FGFR3 downstream signaling in a fraction of the fexagratinib-treated tumors. Notably, the remaining tumors did not show pathway reactivation during treatment outgrowth, suggesting that resistance may also arise through FGFR3 signaling-independent mechanisms, such as tumor cell plasticity or stromal changes including the immune compartment [4,50,51]. Nonetheless, our data demonstrated that tumors driven by the *Fgfr3^{ΔE18}-Tacc3* oncogene are sensitive to FGFRi, thus validating the use of FGFR-targeted therapy for FGFR3^{ΔE18}-TACC3 fusion-positive cancers.

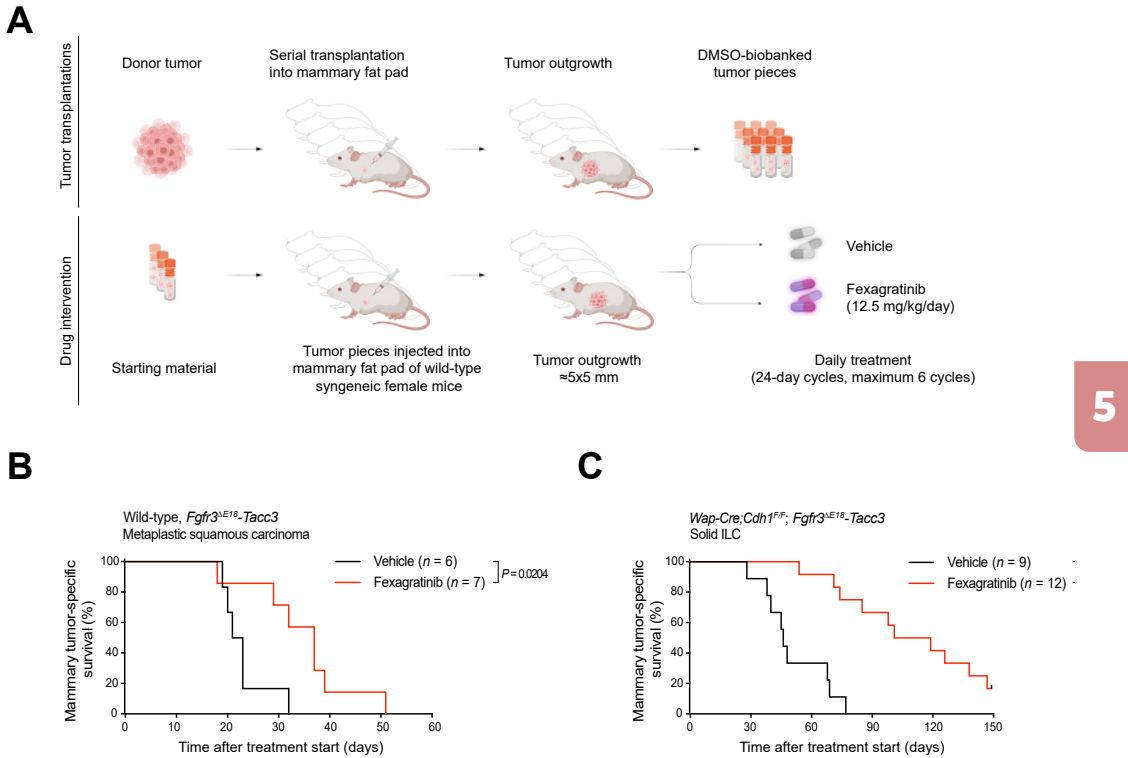


Figure 8. *Fgfr3^{ΔE18}-Tacc3*-driven mammary tumors are sensitive to FGFRi.

A, Schematic overview of the mammary tumor transplantation setup and subsequent FGFRi intervention study fexagratinib (partially created with BioRender.com). **B**, **C**, Kaplan-Meier curves showing mammary tumor-specific survival of female syngeneic wild-type (**B**) or *Wap-Cre;Cdh1^{FF}* (**C**) mice bearing mammary fat pad transplants derived from indicated tumor donors and treated daily orally with vehicle or 12.5 mg/kg fexagratinib using an intermittent dosing regimen [50]. *P* values were calculated with log-rank (Mantel-Cox) tests.

DISCUSSION

Our previous work uncovered E18-truncated *FGFR2* to be a potent driver alteration in cancer that is highly sensitive to FGFR2-targeting [5]. This finding has substantially broadened the spectrum of cancer patients who benefit from FGFR-targeted therapies [5,6]. Here we set out to elucidate the landscape of E18-truncating *FGFR3* alterations in cancer to understand whether our previous findings in *FGFR2* translate to its paralog *FGFR3*. In-depth oncogenomic analyses of *FGFR3* REs enabled the comprehensive detection of alterations that likely promote the expression of *FGFR3^{ΔE18}*, either by *FGFR3* E18-truncating mutations or structural REs. Although many *FGFR3* REs were initially classified as having an unknown

reading frame at the DNA level, analysis of matched RNA-seq data showed that most of these events produce in-frame RE transcripts (**Fig. 1D** and **2B**). The predominance of self-interacting domains in the FGFR3^{ΔE18} RE partners (**Fig. 3B** and **C**), indicated that a RE partner with a conserved self-interacting domain might be a prerequisite for the oncogenic activity of FGFR3^{ΔE18}. Among structural REs resulting in FGFR3^{ΔE18}, TACC3 emerged as the most frequent RE partner (85%), with most cases conserving its two CC domains intact as part of the fusion (**Fig. 3**). FGFR3-TACC3 fusions have been identified in multiple solid tumor types, particularly in glioblastoma, urothelial, lung and breast cancers [8,11,13,23,24], consistent with our own analyses (**Fig. 1A** and **D**).

Notably, 12.3% of all FGFR3 alterations were E18-truncating mutations potentially producing C-terminally truncated FGFR3 devoid of any fusion partner (**Fig. 1A**). In the case of FGFR2, the oncogenic competence of different alteration types depends on distinct co-occurring driver mutations. For instance, FGFR2 hotspot mutations and FGFR2^{ΔE18} alterations cooperate with PIK3CA hotspot mutations to drive PI3K signaling [5,52]. Conversely, FGFR2 amplifications co-occur with TP53 mutations, PTEN loss-of-function, and MYC, CCND1, and FGF3/4/19 amplification [5,53,54], and FGFR2 functionally cooperates with these driver alterations to promote tumorigenesis [5]. The oncogenic competence of FGFR3 E18-truncating mutations and FGFR3 REs devoid of a partner gene might likewise depend on cooperation with specific co-drivers, as suggested by the co-occurrence of KRAS, ERBB2, MYC, PIK3CA, and PTEN alterations in this subset of FGFR3 alterations (Supplementary Fig. S5D-S5G).

We set out to model identified FGFR3 alterations in disease-relevant animal models; however, somatic gene delivery to tissues of the urothelial tract has remained technically challenging due to a resistant urothelial barrier and the lack of efficient delivery methods. We therefore focused our efforts on testing the *in vivo* oncogenicity of Fgfr3 variants in lung and mammary gland tumorigenesis. We demonstrated a causal role of FGFR3^{ΔE18}-TACC3 fusions in tumor formation in both settings (**Fig. 5**; Supplementary Fig. S4B and S5A-S5C). The oncogenic capacity of FGFR3-TACC3 fusions has been previously addressed in glioblastoma and lung tumor formation *in vivo* [13,18]. This is in line with our own results for lung tumorigenesis, where we employed additional controls to confirm that both the loss of the C-terminal tail of FGFR3 (FGFR3^{ΔE18}) as well as the fusion of a dimerization domain-containing TACC3 portion are required for complete tumor driving potential (**Fig. 5A-C**; Supplementary Fig. S4B). Our data showed that the oncogenic capacity of FGFR3^{ΔE18}-TACC3 is fully dependent on intact CC domains of TACC3, since CC disruption by key proline substitutions hindered FGFR3^{ΔE18}-TACC3 oncogenicity (**Fig. 6A** and **B**; Supplementary Fig. S5H). These findings are supported by previous *in vitro* work suggesting considerable relevance of the TACC3 CC domains for cell transformation and aberrant downstream signaling induction by FGFR3-TACC3 fusions [30,55]. Moreover, the majority of the FGFR3 fusions found in cancer involve partner genes with retained expression of a dimerization

domain, which can facilitate ligand-independent receptor dimerization and signaling [5,56].

In this study, we characterized three additional, previously untested *FGFR3^{ΔE18}* fusions: *FGFR3^{ΔE18}-ADD1*, *FGFR3^{ΔE18}-TNIP2*, and *FGFR3^{ΔE18}-JAKMIP1*. All three variants activated downstream FGFR3 signaling in a kinase-dependent manner and conferred resistance to ERBB2/EGFR TKIs *in vitro*, yet remained sensitive to FGFRi (**Fig. 7**; Supplementary Fig. S6), underscoring a therapeutic opportunity in clinical settings. Intriguingly, the *in vivo* oncogenicity of these fusions in wild-type mouse mammary epithelium diverged: *Fgfr3^{ΔE18}-Jakmip1* acted as a *bona fide* tumor driver, whereas *Fgfr3^{ΔE18}-Tnip2* displayed an intermediate phenotype, and *Fgfr3^{ΔE18}-Add1* failed to induce tumors (**Fig. 7G**; Supplementary Fig. S6H). These differences suggest that *in vivo* tumorigenicity of *FGFR3^{ΔE18}* fusions may be dependent on the target tissue, as *FGFR3^{ΔE18}-ADD1* and *FGFR3^{ΔE18}-TNIP2* fusions are prevalent in urothelial cancer rather than breast cancer. Additionally, these fusions might require specific genomic co-alterations to achieve full oncogenic competence, as proposed for *FGFR3^{ΔE18}* alterations without a fusion partner. Moreover, ADD1 contains a dimerizing head-and-neck domain, as opposed to the CC domains in TACC3, TNIP2, and JAKMIP1 (**Fig. 7A**). These different structural features might impact *in vivo* dimerization capacity of *FGFR3^{ΔE18}-ADD1* and ultimately oncogenicity.

Previous research has suggested that the contribution of TACC3 to the oncogenic potential of *FGFR3^{ΔE18}-TACC3* fusions might extend beyond merely providing a self-interacting domain. Singh *et al.* [13] proposed that the *FGFR3-TACC3* fusion locates to the mitotic spindle poles via the retained CC domain, leading to chromosomal segregation defects, aneuploidy, and eventually oncogenic transformation. However, others [57] have not observed direct localization of the fusion protein at the mitotic spindle and attribute the observed mitotic defects to the fusion's CC domain-mediated sequestration of endogenous TACC3 from the mitotic spindle. Our results show that the physiological role of TACC3 in mitotic spindle assembly and microtubule stability is dispensable for the oncogenic potential of the *FGFR3^{ΔE18}-TACC3* fusion (**Fig. 6C**; Supplementary Fig. S5I). Additionally, while impaired internalization of the *FGFR3^{ΔE18}-TACC3* fusion upon ligand stimulation may contribute to sustained kinase signaling, it is not the primary driver of its oncogenicity. Both *FGFR3^{ΔE18}* and *FGFR3^{ΔE18}-TACC3* lack the C-terminal tail that contains the YXXΦ endocytosis internalization motif and accordingly show impaired receptor internalization (**Fig. 4G-I**), yet only the fusion variant is capable of inducing tumor formation (**Fig. 5**) and *in vitro* transformation (**Fig. 4B-F**; Supplementary Fig. S2B-S2E). This indicates that receptor accumulation at the cell membrane alone is insufficient to promote oncogenic signaling.

We recently identified a conserved C-terminal phenylalanine-serine motif in FGFR2 that mediates binding of the C-terminal tail to the kinase domain and suppresses RTK signaling transduction [58]. However, in FGFR3 only the phenylalanine is conserved whereas the

serine is substituted by an alanine, raising the possibility that the C-terminal tail in FGFR3 plays a less significant role in suppressing RTK activity as compared to FGFR2. In support of this notion, truncations resulting in loss of the C-terminal tail lead to constitutive activation of FGFR2 [5,58] but have only marginal impact on FGFR3 activation (**Fig. 4B-F**; Supplementary Fig. S2B-S2E). As delineated in this study, FGFR3 requires both C-tail truncation and fusion to a partner gene that retains the expression of a dimerizing domain to effectively drive oncogenic signaling and tumorigenesis. The presence of a fusion partner may also be necessary for FGFR3^{ΔE18} to retain the ability to activate PLCγ in a FGFR3 kinase domain-dependent manner, despite lacking the canonical PLCγ-binding site that locates to the proximal C-terminal tail (**Fig. 7B**; Supplementary Fig. S2I and S6A). Taken together, our work highlights how subtle structural differences between the two RTKs may dictate distinct paths towards oncogenic activation.

Collectively, our findings demonstrate divergent behavior of the FGFR2 and FGFR3 paralogs in cancer biology following the loss of the C-terminal tail. Importantly, C-terminally truncated FGFR2 as well as truncated FGFR3 fusions rely on canonical kinase function and are sensitive to FGFRi. Thus, we propose that patients with tumors harboring E18-truncating *FGFR2* alterations or *FGFR3* fusions that engage dimerizing partner genes should be considered for FGFR-targeted therapies.

MATERIALS AND METHODS

Mouse models

To somatically model *Fgfr3* variants in the mouse mammary gland, 6 to 7-week-old female FVB/NRj wild-type (Janvier Labs, RRID:IMSR_RJ:FVB) or in-house bred *Wap-cre;Cdh1^{f/f}* (RRID:MGI:6296605) mice were intraductally injected as previously described [36]. In short, 20 μl of high-titer concentrated lentiviruses (2×10^8 - 2×10^9 viral particles/ml) encoding *Fgfr3* variants mixed with 0.2% Evans blue dye in PBS were injected into the fourth and/or third mammary glands. Mice were monitored and palpated twice weekly for tumor development. Mammary tumor-free survival was assessed individually for each injected gland upon detection of a palpable tumor (event), while mammary glands that remained tumor-free were censored. Tumor growth was measured in 2D using calipers by applying the following formula: length x width² x 0.5. Endpoint was reached when the total mammary tumor burden was a volume of either 1,500mm³ for a single tumor or 2,000 mm³ for cumulative tumors or the mice suffered from clinical signs of distress, such as respiratory distress or rapid weight loss. Mice that were euthanized due to clinical signs of distress were censored.

To somatically model *Fgfr3* variants in the mouse lung, high-titer lentiviral vectors encoding *Fgfr3* variants and a *Cre* recombinase were intratracheally delivered via an intubation cannula [59] into the trachea of in-house bred 8 to 12-week-old female/male

FVB *Wap-cre;mTmG* or *Wap-cre;mTmG;P53F* mice. Intratracheal delivery is considered the most direct and consistent method for the virus to reach the lungs. To increase infection efficacy, mice were treated with cyclosporin A orally in the drinking water one week prior to lentiviral infection up until two weeks post infection. Animals were monitored for tumor development every two weeks using non-invasive CT scan-based imaging technology, starting six weeks post lentiviral infection. Mice were sacrificed upon appearance of tumor burden-related clinical symptoms, such as respiratory distress or acute weight loss, which is when a lung tumor-specific survival event was scored.

All tumor measurements and post-mortem analyses were performed in a blinded fashion. The mouse colonies were housed in a certified animal facility with a 12 h light/dark cycle in a temperature-controlled room. Mice were kept in individually ventilated cages, and food and water were provided *ad libitum*. All animal experiments were approved by the Animal Ethics Committee of the Netherlands Cancer Institute under work protocol number 9.2.9125; 9.2.9168; 9.1.9592; 9.1.10022 (AVD3010020172464), 34.1.1140 (AVD30100202316772), 24.1.10061 (AVD30100202011584) and conducted in accordance with institutional, national and European guidelines for Animal Care and Use.

***In vivo* fexagratinib efficacy and PK/PD studies**

DMSO-preserved 1 mm³ tumor fragments derived from spontaneous tumors driven by *Fgfr3* variants that emerged in the mammary glands of either FVB/NRj wild-type or *Wap-cre;Cdh1^{F/F}* mice were serially transplanted into the mammary fat pad of 8 to 12-week-old FVB/NRj wild-type female mice. Animals were monitored weekly for tumor development and, as soon as tumor pieces showed outgrowth, tumors were measured with a caliper 1-2 times per week. Mice were sacrificed when total tumor burden of 1,500 mm³ was reached, and tumors were harvested for viable DMSO-biobanking. These DMSO-preserved donor tumor pieces were then used as the starting material for fexagratinib intervention studies and serially transplanted into the mammary fat pad of 8 to 12-week-old female FVB/NRj wild-type mice. For the efficacy study, tumors were grown to a volume of 62.5 mm³ (5 × 5 mm, measured in 2D using calipers; volume = length × width² × 0.5), and animals were randomly allocated to treatment arms. Mice received either vehicle control (1% Tween-80 in demineralized water) or an optimized intermittent dosing schedule of the FGFRi fexagratinib (12.5 mg/kg, AstraZeneca; ref. [26]) as previously described [50]. The treatments were performed daily via oral gavage and mice were euthanized 1 h post the last dosing. For the efficacy or pharmacokinetic-pharmacodynamic (PK/PD) study, tumors were grown to a volume of approximately 170 mm³ (7 × 7 mm), and animals were randomized to treatment arms. Mice were dosed daily with either vehicle or 12.5 mg/kg fexagratinib for either 2 or 5 consecutive days. Mice were euthanized 1 h post the last dosing, and blood samples were obtained by cardiac puncture using heparinized syringes. Plasma was obtained after centrifugation for 5 min at 5000 × g and kept frozen until analysis. Tumors were excised, weighed and snap-frozen. Prior to analyses, tumors

were fractioned and one part was homogenized in 10 volumes of 1% bovine serum albumin (BSA) in water. Fexagratinib concentrations were measured using an LC-MS/MS method. Calibration samples (range 10 to 10,000 nM) were prepared in human plasma. Volumes of 5 μ l of human plasma and 5 μ l of mouse plasma samples were vortex-mixed with 60 μ l of acetonitrile:formic acid (100:1; v/v), containing 100 nM of buparlisib as internal standard and centrifuged 5 min (20,000 g, 4°C). For tumors, 5 μ l of human plasma and 50 μ l of tumor homogenate were vortex-mixed with 600 μ l of acetonitrile:formic acid (100:1; v/v). Volume of 40 μ l of the supernatants were diluted with 160 μ l of water and 20 μ l was injected on a Zorbax Extend C18 column (100 x 2 mm; ID) using a linear gradient from 20% to 95% methanol in 0.1% formic acid in water at a flow rate of 0.4 ml/min. MS detection was performed in multiple reaction monitoring (MRM) mode using ion pairs 464.3/217.2 for fexagratinib and 411.2/367.2 for buparlisib. The second tumor part was subjected to Western blot PD biomarker analysis (see “Protein isolation and Western blotting”).

Histology and immunohistochemistry

Mouse mammary gland and lung tissues were formalin-fixed and paraffin-embedded (FFPE), sectioned and processed for hematoxylin and eosin (H&E) histochemical staining using routine procedures. For mammary gland tissues, E-cadherin immunohistochemistry (IHC) stainings were additionally performed. First, antigen retrieval was achieved with Tris-EDTA at pH 9.0. Next, sections were incubated with anti-E-cadherin primary antibody (see Supplementary Table S1 for details) overnight at 4°C, labeled with EnVision+ HRP Labelled Polymer Anti-Rabbit System (#K4003, Dako; RRID: AB_2630375), visualized with the Liquid DAB+ Substrate Chromogen System (#K3468, Dako), and counterstained with hematoxylin. H&E slides were used to classify mammary and lung lesion types according to the international pathology consensus, and E-cadherin stainings were used to confirm the mammary lesion types. Certified comparative pathologists (S.K., J-Y.S.) evaluated all slides in a blinded fashion.

Lentiviral vectors and virus production

The SIN.LV.SF, SIN.LV.SF-T2A-Puro, SIN.LV.SF-GFP-T2A-Puro and SIN.LV.SF-P2A-Cre lentiviral vectors were previously described [5]. Custom-synthesized gBlocks gene fragments of mouse *Fgfr3* (NM_001163215.2), *Tacc3* (NM_001310541.1), *Add1* (NM_001024458), *Tnip2* (NM_139064) and *Jakmip1* (NM_178394) were purchased (Integrated DNA Technologies). gBlocks gene fragments or parts thereof were assembled in the respective lentivectors using PCR amplification of backbone and insert(s) and the In-Fusion HD Cloning Plus kit (#638911, Takara Bio) according to the manufacturer’s recommendations. Point mutations S764A and L772A/L773A in *Tacc3*, and K502R and K644E in *Fgfr3* were generated using the QuikChange Lightning Site-Directed Mutagenesis Kit (#210519, Agilent). The Paircoil2 algorithm [38] was used to annotate the two main CC motifs in the TACC3 protein sequence and to predict the likelihood of their disruption by substituting key leucine (L), valine (V), and/or glutamine (Q) amino acid (AA) residues by prolines (P). Custom-

synthesized gBlocks gene fragments of proline-substituted *Tacc3* variants were purchased (Integrated DNA Technologies) and assembled in the respective lentivectors as described above. Detailed nucleotide and AA sequences are provided in Supplementary Methods S1-S9. In-Fusion and site-directed mutagenesis primers were designed using SnapGene (v.7.0.3; RRID: SCR_015052) and QuikChange Primer Design (<https://www.agilent.com/store/primerDesignProgram.jsp>; RRID: SCR_025323), respectively. The lentiviral vectors used in this study were all verified by Sanger sequencing. Concentrated (*in vivo*) and unconcentrated (*in vitro*) lentiviral stocks were produced by transient co-transfection of four plasmids in HEK293T cells as previously described [60]. Viral titers were measured with the qPCR Lentivirus Titer Kit (#LV900, Applied Biological Materials).

Cell culture and lentiviral transduction

HEK293T cells (#CRL-3216, ATCC; RRID: CVCL_0063) were cultured in Iscove's Modified Dulbecco Medium (#21980-032, Thermo Fisher Scientific) supplemented with 10% Fetal Bovine Serum (FBS, #FBS-12A, Capricorn Scientific) and 100 IU/ml penicillin and streptomycin (Pen Strep, #15070, Thermo Fisher Scientific). NMuMG mouse mammary epithelial cells (#CRL-1636, ATCC; RRID: CVCL_0075) were cultured in Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F-12, #31331-028, Thermo Fisher Scientific) containing 10% FBS and 100 IU/ml Pen Strep. Ba/F3 murine pro B cells (#ACC-300, DSMZ; RRID: CVCL_0161) were cultured in RPMI 1640 medium (#21875-034, Thermo Fisher Scientific) supplemented with 10% FBS, 100 IU/ml Pen Strep and 10 ng/ml recombinant murine IL-3 (#213-13, PeproTech). 5637 human bladder cancer cells (#ACC-35, DSMZ; RRID: CVCL_0126) were cultured in RPMI 1640 medium supplemented with 10% FBS and 100 IU/ml Pen Strep. No re-authentication of the abovementioned cell lines was performed for this study beyond that initially carried out by providers.

To stably express the lentiviral *GFP-T2A-Puro* and *Fgfr3-T2A-Puro* constructs, NMuMG, Ba/F3 and 5637 cells were transduced with lentiviral supernatants at equal multiplicity of infection (MOI) in the presence of 8 µg/ml Polybrene (#H9268, Sigma-Aldrich) for 24 h. Transduced cells were selected with 2 µg/ml Puromycin (#A11138, Thermo Fisher Scientific) for at least 7 days and subsequently grown in medium containing 1 µg/ml Puromycin, with additional 10 µM Y-27632 dihydrochloride (#M1817, AbMole) for the NMuMG cells. All cell lines were cultured in standard incubators at 37 °C with 5% CO₂ and routinely tested for mycoplasma contamination using the MycoAlert Mycoplasma Detection Kit (#LT07-218, Lonza). Experiments were performed with cells cultured for a maximum of 10 passages post lentiviral transduction.

Soft agar anchorage-independent 3D growth assay

Twelve-well plates were precoated with 1 ml of 0.6% low-gelling-temperature agarose (#A9414, Sigma-Aldrich) by diluting 3% agarose solution (in PBS) in DMEM/F-12 supplemented with 3% FBS and Pen Strep. The bottom layer was solidified at 4 °C for 30

min. A single-cell suspension of 30,000 NMuMG cells per well was suspended in 1 ml of 0.3% low-gelling-temperature agarose in DMEM/F-12 supplemented with 3% FBS and Pen Strep, and plated on top. The top layer was solidified at 4 °C for 30 min before transferring the plates to the incubator. The plates were imaged after 20 days using the GelCount colony counter and anchorage-independent growth was quantified using the integrated GelCount colony counting platform (v.1.1.2, Oxford Optronix).

Ba/F3 growth assay upon IL-3 withdrawal

Exponentially growing Ba/F3 cells were washed three times with PBS and seeded in triplicate at 2×10^5 cells/ml in the absence of IL-3 supplementation in 12-well plates. Viable cell counts at each time point were obtained using 0.4% trypan blue (#T8154, Sigma-Aldrich) exclusion on a LUNA-II automated cell counter (Logos Biosystems).

Cell viability assays

5637 parental cells and those expressing *Fgfr3-T2A-Puro* variants were seeded in 12-well plates at 30,000 cells per well. After 24 hr, cells were treated with fexagratinib, afatinib (#HY-10261, MedChem Express), dacomitinib (#HY-13272, MedChem Express) or a combination of afatinib with an anchor concentration of 2.5 μ M fexagratinib for 72 hr. Cell viability was measured using CellTiter-Blue reagent (#G808A, Promega) after 1.5-2 hr incubation, and fluorescence was recorded with an Infinite M Plex plate reader (Tecan; RRID:SCR_026522) operated with Tecan i-control software (v.3.9.1; RRID: SCR_024562). Fluorescence values were normalized to untreated control to determine relative viability. Drug-response curves were fitted in GraphPad Prism (v.10.4.1, RRID: SCR_002798) using a four-parameter logistic regression model ([Inhibitor] vs. response, variable slope). For endpoint visualization, cells were fixed with 4% formaldehyde and stained with 0.5% crystal violet in 25% methanol, washed with deionized water, air-dried, and imaged with GelCount (Oxford Optronix; RRID: SCR_023219).

RNA isolation and RT-qPCR

To confirm lentiviral construct overexpression, transduced cells were lysed in RLY buffer (#93594, Bioline) containing 1% of β -mercaptoethanol and total RNA was extracted using the ISOLATE II RNA Mini Kit (#BIO-52073, Bioline) as per manufacturer's instructions. Purified RNA was quantified using the DS-11 Series Spectrophotometer (DeNovix) and 1 μ g of total RNA was subjected to reverse transcriptase reaction using the Tetro cDNA Synthesis Kit (#BIO-65042, Bioline) with random hexamer primers. Reverse transcription quantitative PCR (RT-qPCR) was performed using the SensiFAST SYBR Hi-ROX Kit (#BIO-92005, Bioline) and the QuantStudio 6 Flex Real-Time PCR System (#4485691, Thermo Fisher Scientific; RRID: SCR_020239) operated with the QuantStudio Real-Time PCR Software (v.1.7.2, Thermo Fisher Scientific). Primers used were designed using Primer-BLAST (RRID: SCR_003095). Relative quantified cDNA was normalized using the mouse *Usf1* as the housekeeping transcript.

Protein isolation and Western blotting

NMuMG cells stably expressing *GFP* or *Fgfr3* variants were cultured under starvation conditions (0% FBS) for 24 h and, where indicated, treated with 20 ng/ml recombinant human/murine fibroblast growth factor 8b (FGF-8b, #100-25, PeproTech) for 10 min, or 1 μ M fexagratinib (AstraZeneca; ref. [26]), 100 nM erdafitinib (#HY-18708, MedChem Express) or 100 nM futibatinib (#HY-100818, MedChem Express) for 3 h. Cells and tumor pieces were lysed in RIPA buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA, 1% NP40, 0.5% deoxycholate, and 0.1% SDS in milli-Q water) containing Halt Protease and Phosphatase Inhibitor Cocktail (#78440, Thermo Fisher Scientific) either by pipetting (cells) or using a mechanical homogenizer (tumor pieces). Lysates were cleared by centrifugation at 14,000 rpm for 10 min at 4°C, and protein concentrations were determined with the Pierce BCA Protein Assay Kit (#23225, Thermo Fisher Scientific) using an Infinite 200 PRO plate reader. BlueEye Prestained Protein Marker (#PS-104, Jena Bioscience) and equal amounts of protein were separated on NuPAGE 4-12% Bis-Tris Mini Protein Gels (#NP0323, Thermo Fisher Scientific) using NuPAGE MOPS SDS Running Buffer (#NP0001, Thermo Fisher Scientific), and transferred overnight at 4°C onto Nitrocellulose Membranes (#1620112; Bio-Rad) in transfer buffer (25 mM Tris, 2 M Glycine, 20% methanol in demineralized water). Membranes were stained with Ponceau S (#P7170, Sigma-Aldrich) and imaged with Fusion FX (Vilber), blocked in 5% BSA (#A8022; Sigma-Aldrich) in PBS-T (0.05% Tween-20), and incubated in 5% BSA in PBS-T overnight at 4°C. Membranes were washed in PBS-T and incubated with secondary antibodies in 5% BSA in PBS-T for 1 h at room temperature. Membranes were then washed three times in PBS-T, once in PBS and developed using SuperSignal West Pico PLUS Chemiluminescent Substrate (#34580, Thermo Fisher Scientific), and imaged using Fusion FX operated with the Fusion FX7 Edge imaging system (v.18.05, Vilber). Protein band intensities were measured with mean grey value in Fiji (v.2.14.0, RRID: SCR_002285)[61] and normalized to β -Actin, with phosphoprotein bands being further normalized to corresponding total protein bands. A list of the antibodies used is provided in Supplementary Table S1.

Fluorescence-activated cell sorting

To measure FGFR3 internalization, NMuMG cells expressing *HALO-Fgfr3-T2A-Puro* variants were cultured in DMEM/F12 medium lacking FBS supplementation for 24 hr and treated with 20 ng/ml recombinant human/murine FGF-8b for indicated times. Next, following the surface protein internalization and degradation assay (SPIDA; ref. [33]), cells were sequentially labeled for 15 min each at 37 °C with 350 nM cell membrane impermeable HaloTag ligand (HTL-AF660; #G8471, Promega) and 250 nM cell membrane permeable ligand (HTL-TMR; #G8251, Promega). Cells were then collected with 2 mM EDTA in PBS and labeled with the LIVE/DEAD Fixable Violet Dead Cell Stain Kit (405 nm excitation; #L34955, Thermo Fisher Scientific) and fixed with BD Phosflow Fix Buffer I (#557870, BD Biosciences), at 4 °C for 30 min each. Fluorescence-activated cell sorting (FACS) was performed using a BD LSR II flow cytometer (BD Biosciences, RRID: SCR_002159) operated with BD FACSDiva

Software (v.8.0.1, RRID: SCR_001456). Fluorescence was detected using 450/50, 670/14, and 585/15 bandpass filters to measure LIVE/DEAD viability dye, AF660-HALO-FGFR3, and TMR-HALO-FGFR3, respectively. Data were analyzed with FlowJo (v.10.10.0, BD Biosciences; RRID: SCR_008520) and GraphPad Prism softwares.

Analysis of CGP data from Foundation Medicine

Comprehensive genomic profiling (CGP) was performed on FFPE tumor tissue or blood samples prospectively collected during routine clinical care, as previously described [5]. Briefly, testing was conducted at Foundation Medicine Inc. (FMI), a CLIA-certified, CAP-accredited, New York State-regulated reference laboratory (RRID: SCR_025628). The study was approved by the Western Institutional Review Board (protocol 20152817), with waivers for informed consent and HIPAA authorization. For 217,017 tumor tissue specimens, DNA (>50 ng) was extracted from FFPE specimens, and next-generation sequencing was performed using FoundationOne or FoundationOne CDx testing. This provided high, uniform coverage (>500x) of 315 or 324 cancer-related genes and selected introns of 28 or 36 genes frequently rearranged in cancer [62,63]. Additionally, 26,289 samples were assayed with DNA sequencing for 406 genes and selected introns of 31 genes involved in RE events, along with RNA sequencing (RNA-seq) for 265 genes [64]. For 6,264 liquid samples, plasma was isolated from 20 ml of peripheral whole blood, and at least 20 ng of circulating tumor DNA was extracted [65]. Adapted sequencing libraries were created for the coding exons of 70 genes before hybrid-capture and sample-multiplexed sequencing. Results were analyzed for base substitutions, short insertions and deletions, copy number gains or losses, and REs. The companion diagnostic tests included probes against all *FGFR3* exons and *FGFR3*-I17.

To annotate REs, the location and orientation of both BP sides (*FGFR3* and its partner) were used to determine RE types, as previously described [5]. The gene encoding the longer protein sequence was used as the RE classification backbone. The RE types were defined as follows: (i) in-frame fusion, both BP sides were in the intronic regions of coding genes with in-frame adjacent exons, or both BP sides were in the exonic regions with an in-frame fused sequence; (ii) frame unknown RE, both BP sides were in the intronic regions with at least one out-of-frame adjacent exon, or one or both BP sides were in the exonic regions with an out-of-frame fused sequence; (iii) RE with intergenic space, one BP side was in *FGFR3* and the other in a non-coding intergenic region; (iv) out-of-strand RE, both BP sides were in coding regions, with sense-oriented reads for *FGFR3* and antisense-oriented reads for the partner gene; (v) internal RE, both BP sides were within the genomic region of *FGFR3*; (vi) unresolved RE, the gene upstream to the BP was supported by antisense-oriented read sequences or the REs contained single breakends, which were excluded from the downstream analysis. *FGFR3* amplifications were called if $\geq 80\%$ of the *FGFR3* targets were at an amplified copy number (CN), defined as $\geq 4 + \text{median ploidy of the sample}$. Differential CN gains of E18 targets < E1-E17 targets were defined as *FGFR3*-E1-E17 partial

amplifications. In samples with *FGFR3* REs and co-amplification, low-level REs were discarded at a read threshold dependent on the amplification CN. Nonsense, frameshift, and splice-site mutations in E18 were defined as *FGFR3* E18-truncating mutations.

Analysis of self-interacting capacity among *FGFR3* RE partners

The SLIPPER algorithm was used to obtain proteins with self-interacting capacity [27]. The SLIPPER algorithm, trained on seven proteome databases (DIP, IntAct, MINT, BioGRID, PDB, MatrixDB and 12D), established the SLIPPER golden standard dataset of potentially self-interacting proteins. Using this dataset, the proportion of self-interactors among unique proteins encoded by *FGFR3* RE partner genes in the FMI dataset was calculated. To identify specific self-interacting domains in E18-truncated *FGFR3* RE partners, domain-domain interaction data were obtained from 3did [66] and PPIDM [67] databases. Proteins containing coiled-coil domains were obtained from the UniProt database (release 2021_04). Two-tailed Fisher's exact test was performed to statistically compare the proportion of self-interacting proteins and proteins with coiled-coil domains.

Analysis of WGS data from the HMF

Preprocessed whole genome sequencing (WGS) data on metastatic solid tumors were obtained from the Hartwig Medical Foundation (HMF; data access request DR-205) via their Google Cloud computing platform. The data were preprocessed using their bioinformatics pipeline, designed to detect all types of somatic alterations, including structural variants and CN alterations (CNAs; <https://github.com/hartwigmedical/pipeline5>). In brief, sequencing reads were mapped to the human reference genome GRCh37 using Burrows-Wheeler Alignment (BWA-MEM, v.0.7.5a; ref. [68]). Somatic structural variants were identified with GRIDSS (v.1.8.0; ref. [69]), and CNAs and tumor purity were estimated using PURPLE (v.2.43, RRID: SCR_022999; ref. [10]). LINX (v.1.9; ref. [70]) was then used to annotate events and construct derivative chromosome structures. Based on the PURPLE output, samples containing structural variant BPs within the *FGFR3* genomic region were selected for further structural variant analysis. To annotate structural variants, the location and orientation of both BP sides (*FGFR3* and its partner) were used to determine RE types as described for the FMI panel-seq data. *FGFR3* CN gains of >5 were defined as amplifications. Among the samples with *FGFR3* CN segment BPs at I17/E18, samples with E1-E17 CN ($CN_{E1-E17} > 5$) and $CN_{E1-E17} - CN_{E18} > 2$ were defined as *FGFR3*-E1-E17 partially amplified.

Analysis of RNA-seq data from the HMF

Raw RNA-seq data (fastq files) on metastatic solid tumors were obtained from the HMF (data access request DR-205). Sequencing reads were mapped to the human reference genome GRCh38 (Gencode v32 CTAT) using STAR (v.2.7.2; ref. [71]) with the recommended parameters. These mapped reads were then used to run STAR-Fusion (v.1.8.1; RRID: SCR_025853; ref. [72]). STAR-Fusion was executed with the chimeric alignment information (Chimeric.out.junction) obtained from STAR and GRCh38 Gencode v32 CTAT. Chimeric

alignments from STAR and gene fusions from STAR-Fusion were inspected to identify RNA-seq alignments supporting the REs identified in WGS. For samples with in-frame fusions identified in WGS, the upstream and downstream exon numbers of the fusion gene, inferred from STAR-Fusion, were matched to the fusion found in WGS. For samples with other types of REs identified in WGS, chimeric reads spanning the upstream exon and the downstream exon (out-of-frame REs), the downstream intergenic sequence (REs with intergenic space), or the downstream antisense gene sequence (out-of-strand REs) were mined from the 'Chimeric.out.junction' file and matched to the REs found in WGS. Genome coordinates were converted from GRCh37 to GRCh38 using UCSC Lift Genome Annotations (<https://genome.ucsc.edu/cgi-bin/hgLiftOver>). IGV (v.1.11.0, RRID: SCR_011793; ref. [73]) was used to generate sashimi plots.

Statistical analysis and reproducibility

Data of *in vitro* and *in vivo* experiments were analyzed using Prism (v.10.0.3, GraphPad Software; RRID: SCR_002798) and ClustVis (v.2.0, RRID: SCR_017133; ref. [74]). Genomic data were analyzed using R (v. 4.3.3, RRID: SCR_001905). *In vitro* experiments were independently repeated at least twice, and all attempts at replication were successful. Sample sizes of mouse cohorts were based on previous calculations [5] or determined using the Sample Size Calculator (<http://powerandsamplesize.com/Calculators>) and were large enough to measure the effect sizes. Data were reproducible across batches analyzed, and all attempts at replication were successful. The statistical tests used are described in the corresponding figure legends. $P < 0.05$ was considered statistically significant. Exact P values are always disclosed in the corresponding figure panels.

Data availability

Preprocessed WGS and raw RNA-seq data from the HMF were downloaded from their Google Cloud computing platform under data-sharing agreement DR-205. These data can be obtained through standardized procedures and request forms available online (<https://www.hartwigmedicalfoundation.nl/en/>). CGP data can be obtained from FMI on reasonable request (<https://www.foundationmedicine.com/service/genomic-data-solutions>). The human reference genome (GRCh38 Gencode v32 CTAT) used for RNA-seq data analysis is available in CTAT Genome Lib data resources (https://data.broadinstitute.org/Trinity/CTAT_RESOURCE_LIB). Domain-domain interaction information from 3did and PPIDM are available online (<https://3did.irbbarcelona.org> and <http://ppidm.loria.fr>, respectively). The code used in this study can be found at Code Ocean (<https://codeocean.com/capsule/5476052/tree/v1>). All other raw data generated in this study are available upon request from the corresponding authors.

Code availability

WGS data from the HMF were analyzed using published computer codes (<https://github.com/hartwigmedical/pipeline5>) including BWA-MEM, GRIDSS, PURPLE and LINX to detect

all types of somatic alterations, including structural variants and CNAs. Codes used for analyzing *FGFR3* RE types and partners are deposited at https://github.com/bynjh007/FGFR3_RE_analysis.

Conflicts of interest

J.K.L. is an employee of Foundation Medicine, Inc. and has equity in Roche. J.S.R. is an employee of Foundation Medicine, Inc., has equity in Roche, and is a board member of Celsius Therapeutics. S.G. has received research funding from M2Gen, has equity in Inspirata and Silagene, and has consulted for Foghorn Therapeutics, Foundation Medicine, Inspirata, Merck, Novartis, Roche and Silagene; his spouse is an employee of and has equity in Merck. D.Z. is an employee of Genentech, Inc. and has equity in Roche. The other authors declare no potential conflicts of interest.

Acknowledgments

We thank H. van der Gulden for support with cloning, F. van Diepen for assistance with FACS, and the NKI animal facility, animal pathology facility and Mouse Cancer Clinic for their technical support in performing the animal experiments. This publication and the underlying study have been made possible in part on the basis of data provided by HMF and the Center of Personalised Cancer Treatment (CPCT).

Research at the NKI is supported by institutional grants of the Dutch Cancer Society and the Dutch Ministry of Health, Welfare and Sport. This work is part of the OncoCode Institute, which is partly financed by the Dutch Cancer Society. This work was funded by the Dutch Cancer Society (KWF grants 7835 and 12894). This work was also supported by National Research Foundation of Korea (NRF) grants (MSIT; RS-2024-00342142 and RS-2023-00261820) and a faculty research grant of Yonsei University College of Medicine (6-2024-0041).

Author contributions

J.Y., J.B., D.Z., and J.J. conceived the ideas and designed the experiments. J.Y., C.L., L.D., E.W., S.A., and O.v.T. performed the laboratory experiments. J.Y., A.P.D., E.v.d.B., N.P., B.S., and D.Z. conceived and performed the animal experiments. S.K. and J.-Y.S. assessed the histology of mouse tumors. J.B. and H.M. performed the bioinformatic analyses. J.K.L. and J.S.R. provided the Foundation Medicine data. J.Y., J.B., S.G., L.F.A.W., D.Z. and J.J. supervised the bioinformatic analyses and the laboratory experiments. M.v.d.V. supervised the animal experiments. J.Y., J.B., D.Z., and J.J. wrote the manuscript.

Online content

The supplementary data for this article, including Supplementary Tables S1-S2, Supplementary Data S1-S2 and Supplementary Methods S1-S9 are available at <https://doi.org/10.1158/0008-5472.CAN-24-2648>.

REFERENCES

1. Helsten, T. *et al.* (2016) The FGFR landscape in cancer: Analysis of 4,853 tumors by next-generation sequencing. *Clin. Cancer Res.* 22, 259–267
2. Murugesan, K. *et al.* (2022) Pan-tumor landscape of fibroblast growth factor receptor 1-4 genomic alterations. *ESMO Open* 7, 100641
3. Sun, Y. *et al.* (2020) A comprehensive pan-cancer study of fibroblast growth factor receptor aberrations in Chinese cancer patients. *Ann. Transl. Med.* 8, 1290–1290
4. Katoh, M. *et al.* (2024) FGFR-targeted therapeutics: clinical activity, mechanisms of resistance and new directions. *Nat. Rev. Clin. Oncol.* DOI: 10.1038/s41571-024-00869-z
5. Zingg, D. *et al.* (2022) Truncated FGFR2 is a clinically actionable oncogene in multiple cancers. *Nature* 608, 609–617
6. Rodón, J. *et al.* (2024) Pemigatinib in previously treated solid tumors with activating FGFR1–FGFR3 alterations: phase 2 FIGHT-207 basket trial. *Nat. Med.* DOI: 10.1038/s41591-024-02934-7
7. Wu, Y.M. *et al.* (2013) Identification of targetable FGFR gene fusions in diverse cancers. *Cancer Discov.* 3, 636–647
8. Qin, A. *et al.* (2019) Detection of Known and Novel FGFR Fusions in Non–Small Cell Lung Cancer by Comprehensive Genomic Profiling. *J. Thorac. Oncol.* 14, 54–62
9. Turner, N. and Grose, R. (2010) Fibroblast growth factor signalling: From development to cancer. *Nat. Rev. Cancer* 10, 116–129
10. Priestley, P. *et al.* (2019) Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature* 575, 210–216
11. Costa, R. *et al.* (2016) FGFR3-TACC3 fusion in solid tumors: Mini review. *Oncotarget* 7, 55924–55938
12. Mata, D.A. *et al.* (2020) Genetic and epigenetic landscape of IDH-wildtype glioblastomas with FGFR3-TACC3 fusions. *Acta Neuropathol. Commun.* 8, 186
13. Singh, D. *et al.* (2012) Transforming fusions of FGFR and TACC genes in human glioblastoma. *Science* 337, 1231–1235
14. Di Stefano, A.L. *et al.* (2015) Detection, characterization, and inhibition of FGFR-TACC fusions in IDH wild-type glioma. *Clin. Cancer Res.* 21, 3307–3317
15. Parker, B.C. *et al.* (2013) The tumorigenic FGFR3-TACC3 gene fusion escapes miR-99a regulation in glioblastoma. *J. Clin. Invest.* 123, 855
16. Bao, Z.-S. *et al.* (2014) RNA-seq of 272 gliomas revealed a novel, recurrent PTPRZ1-MET fusion transcript in secondary glioblastomas. *Genome Res.* 24, 1765–1773
17. Williams, S.V. *et al.* (2013) Oncogenic FGFR3 gene fusions in bladder cancer. *Hum. Mol. Genet.* 22, 795–803
18. Best, S.A. *et al.* (2018) FGFR3-TACC3 is an oncogenic fusion protein in respiratory epithelium. *Oncogene* 37, 6096–9104
19. Majewski, I.J. *et al.* (2013) Identification of recurrent FGFR3 fusion genes in lung cancer through kinome-centred RNA sequencing. *J. Pathol.* 230, 270–276
20. Capelletti, M. *et al.* (2014) Identification of recurrent FGFR3-TACC3 fusion oncogenes from lung adenocarcinoma. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 20, 6551–6558
21. Wang, R. *et al.* (2014) FGFR1/3 tyrosine kinase fusions define a unique molecular subtype of non-small cell lung cancer. *Clin. Cancer Res.* 20, 4107–4114
22. Kim, Y. *et al.* (2014) Integrative and comparative genomic analysis of lung squamous cell carcinomas in East Asian patients. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 32, 121–128
23. Shaver, T.M. *et al.* (2016) Diverse, Biologically Relevant, and Targetable Gene Rearrangements in Triple-Negative Breast Cancer and Other Malignancies. *Cancer Res.* 76, 4850–4860

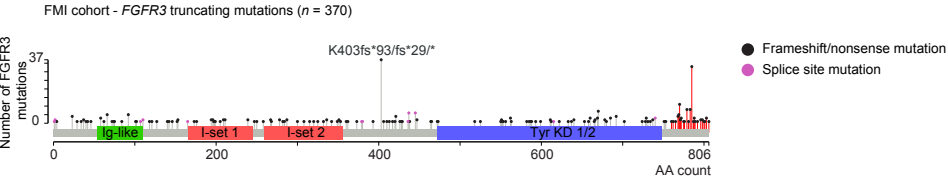
24. Chew, N.J. *et al.* (2020) FGFR3 signaling and function in triple negative breast cancer. *Cell Commun. Signal.* 18, 1–17
25. Nozad, S. *et al.* (2017) Comprehensive genomic profiling of malignant phyllodes tumors of the breast. *Breast Cancer Res. Treat.* 162, 597–602
26. Gavine, P.R. *et al.* (2012) AZD4547: an orally bioavailable, potent, and selective inhibitor of the fibroblast growth factor receptor tyrosine kinase family. *Cancer Res.* 72, 2045–2056
27. Liu, Z. *et al.* (2013) Proteome-wide prediction of self-interacting proteins based on multiple properties. *Mol. Cell. Proteomics MCP* 12, 1689–1700
28. Horibata, S. *et al.* (2015) Utilization of the Soft Agar Colony Formation Assay to Identify Inhibitors of Tumorigenicity in Breast Cancer Cells. *J. Vis. Exp. JoVE* DOI: 10.3791/52727
29. Warmuth, M. *et al.* (2007) Ba/F3 cells and their use in kinase drug discovery. *Curr. Opin. Oncol.* 19, 55–60
30. Nelson, K.N. *et al.* (2016) Oncogenic gene fusion FGFR3-TACC3 Is regulated by tyrosine phosphorylation. *Mol. Cancer Res.* 14, 458–469
31. Szybowska, P. *et al.* (2021) Negative Regulation of FGFR (Fibroblast Growth Factor Receptor) Signaling. *Cells* 10, 1342
32. Bache, K.G. *et al.* (2004) Defective downregulation of receptor tyrosine kinases in cancer. *EMBO J.* 23, 2707–2712
33. Stüber, J.C. *et al.* (2019) High-Throughput Quantification of Surface Protein Internalization and Degradation. *ACS Chem. Biol.* 14, 1154–1163
34. Ciriello, G. *et al.* (2015) Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell* 163, 506–519
35. Derksen, P.W.B. *et al.* (2011) Mammary-specific inactivation of E-cadherin and p53 impairs functional gland development and leads to pleomorphic invasive lobular carcinoma in mice. *Dis. Model. Mech.* 4, 347–358
36. Annunziato, S. *et al.* (2016) Modeling invasive lobular breast carcinoma by CRISPR/Cas9-mediated somatic genome editing of the mammary gland. *Genes Dev.* 30, 1470–1480
37. Kas, S.M. *et al.* (2017) Insertional mutagenesis identifies drivers of a novel oncogenic pathway in invasive lobular breast carcinoma. *Nat. Genet.* 49, 1219–1230
38. McDonnell, A.V. *et al.* (2006) Paircoil2: improved prediction of coiled coils from sequence. *Bioinforma. Oxf. Engl.* 22, 356–358
39. Kinoshita, K. *et al.* (2005) Aurora A phosphorylation of TACC3/maskin is required for centrosome-dependent microtubule assembly in mitosis. *J. Cell Biol.* 170, 1047–1055
40. Fu, W. *et al.* (2013) Self-assembly and sorting of acentrosomal microtubules by TACC3 facilitate kinetochore capture during the mitotic spindle assembly. *Proc. Natl. Acad. Sci. U. S. A.* 110, 15295–15300
41. Schneider, L. *et al.* (2007) The transforming acidic coiled coil 3 protein is essential for spindle-dependent chromosome alignment and mitotic survival. *J. Biol. Chem.* 282, 29273–29283
42. Gergely, F. *et al.* (2003) The ch-TOG/XMAP215 protein is essential for spindle pole organization in human somatic cells. *Genes Dev.* 17, 336–341
43. Yao, R. *et al.* (2007) TACC3 is required for the proper mitosis of sclerotome mesenchymal cells during formation of the axial skeleton. *Cancer Sci.* 98, 555–562
44. Peset, I. *et al.* (2005) Function and regulation of Maskin, a TACC family protein, in microtubule growth during mitosis. *J. Cell Biol.* 170, 1057–1066
45. Lioutas, A. and Vernos, I. (2013) Aurora A kinase and its substrate TACC3 are required for central spindle assembly. *EMBO Rep.* 14, 829–836
46. Ulisse, S. *et al.* (2007) Transforming acidic coiled-coil 3 and Aurora-A interact in human thyrocytes and their expression is deregulated in thyroid cancer tissues. *Endocr. Relat. Cancer* 14, 827–837

47. Ding, Z.M. *et al.* (2017) The role of TACC3 in mitotic spindle organization. *Cytoskeleton* 74, 369–378
48. Hood, F.E. *et al.* (2013) Coordination of adjacent domains mediates TACC3-ch-TOG-clathrin assembly and mitotic spindle binding. *J. Cell Biol.* 202, 463–478
49. Nelson, K.N. *et al.* (2018) Oncogenic driver FGFR3-TACC3 is dependent on membrane trafficking and ERK signaling. *Oncotarget* 9, 34306–34319
50. Kas, S.M. *et al.* (2018) Transcriptomics and transposon mutagenesis identify multiple mechanisms of resistance to the FGFR inhibitor AZD4547. *Cancer Res.* 78, 5668–5679
51. Katoh, M. (2019) Fibroblast growth factor receptors as treatment targets in clinical oncology. *Nat. Rev. Clin. Oncol.* 16, 105–122
52. Juric, D. *et al.* (2025) PI3Ka inhibitor and degrader inavolisib can co-opt FGFR2 to enhance responses in patients with PIK3CA-mutated solid tumors and in preclinical models. *Clin. Cancer Res.* DOI: 10.1158/1078-0432.CCR-25-1459
53. Silverman, I.M. *et al.* (2021) Clinicogenomic Analysis of FGFR2-Rearranged Cholangiocarcinoma Identifies Correlates of Response and Mechanisms of Resistance to Pemigatinib. *Cancer Discov.* 11, 326–339
54. D'Souza, N. *et al.* (2024) Comutational patterns in FGFR3-altered non-muscle invasive bladder cancer influence recurrence and progression. *Urol. Oncol. Semin. Orig. Investig.* 42, S55
55. Wang, C.G. *et al.* (2023) Oncogenic driver FGFR3-TACC3 requires five coiled-coil heptads for activation and disulfide bond formation for stability. *Oncotarget* 14, 133
56. Parker, B.C. *et al.* (2014) Emergence of FGFR family gene fusions as therapeutic targets in a wide spectrum of solid tumours. *J. Pathol.* 232, 4–15
57. Sarkar, S. *et al.* (2017) FGFR3-TACC3 cancer gene fusions cause mitotic defects by removal of endogenous TACC3 from the mitotic spindle. *Open Biol.* 7
58. Zingg, D. *et al.* (2025) The C-Terminal Kinase Domain–Binding and Suppression Motif Prevents Constitutive Activation of FGFR2. *Cancer Res.* 85, 3234–3257
59. DuPage, M. *et al.* (2009) Conditional mouse lung cancer models using adenoviral or lentiviral delivery of Cre recombinase. *Nat. Protoc.* 4, 1064–1072
60. Dull, T. *et al.* (1998) A third-generation lentivirus vector with a conditional packaging system. *J. Virol.* 72, 8463–8471
61. Schindelin, J. *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682
62. Frampton, G.M. *et al.* (2013) Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. *Nat. Biotechnol.* 31, 1023–1031
63. Milbury, C.A. *et al.* (2022) Clinical and analytical validation of FoundationOne®CDx, a comprehensive genomic profiling assay for solid tumors. *PloS One* 17, e0264138
64. He, J. *et al.* (2016) Integrated genomic DNA/RNA profiling of hematologic malignancies in the clinical setting. *Blood* 127, 3004–3014
65. Clark, T.A. *et al.* (2018) Analytical Validation of a Hybrid Capture-Based Next-Generation Sequencing Clinical Assay for Genomic Profiling of Cell-Free Circulating Tumor DNA. *J. Mol. Diagn. JMD* 20, 686–702
66. Stein, A. *et al.* (2005) 3did: interacting protein domains of known three-dimensional structure. *Nucleic Acids Res.* 33, D413–417
67. Alborzi, S.Z. *et al.* (2021) PPIDomainMiner: Inferring domain-domain interactions from multiple sources of protein-protein interactions. *PLoS Comput. Biol.* 17, e1008844
68. Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760
69. Cameron, D.L. *et al.* (2017) GRIDSS: sensitive and specific genomic rearrangement detection using positional de Bruijn graph assembly. *Genome Res.* 27, 2050–2060

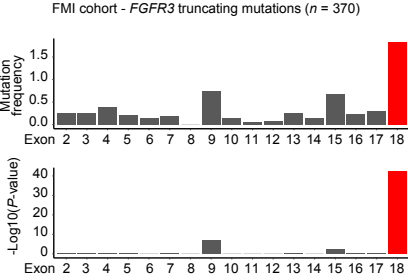
70. Shale, C. *et al.* (2022) Unscrambling cancer genomes via integrated analysis of structural variation and copy number. *Cell Genomics* 2, 100112
71. Dobin, A. *et al.* (2013) STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21
72. Haas, B.J. *et al.* (2019) Accuracy assessment of fusion transcript detection via read-mapping and de novo fusion transcript assembly-based methods. *Genome Biol.* 20, 213
73. Robinson, J.T. *et al.* (2011) Integrative genomics viewer. *Nat. Biotechnol.* 29, 24–26
74. Metsalu, T. and Vilo, J. (2015) ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res.* 43, W566–570

SUPPLEMENTARY FIGURES

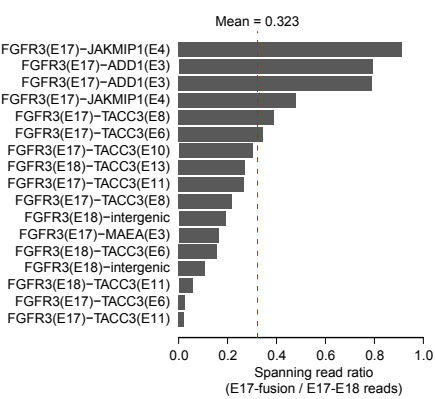
A



B

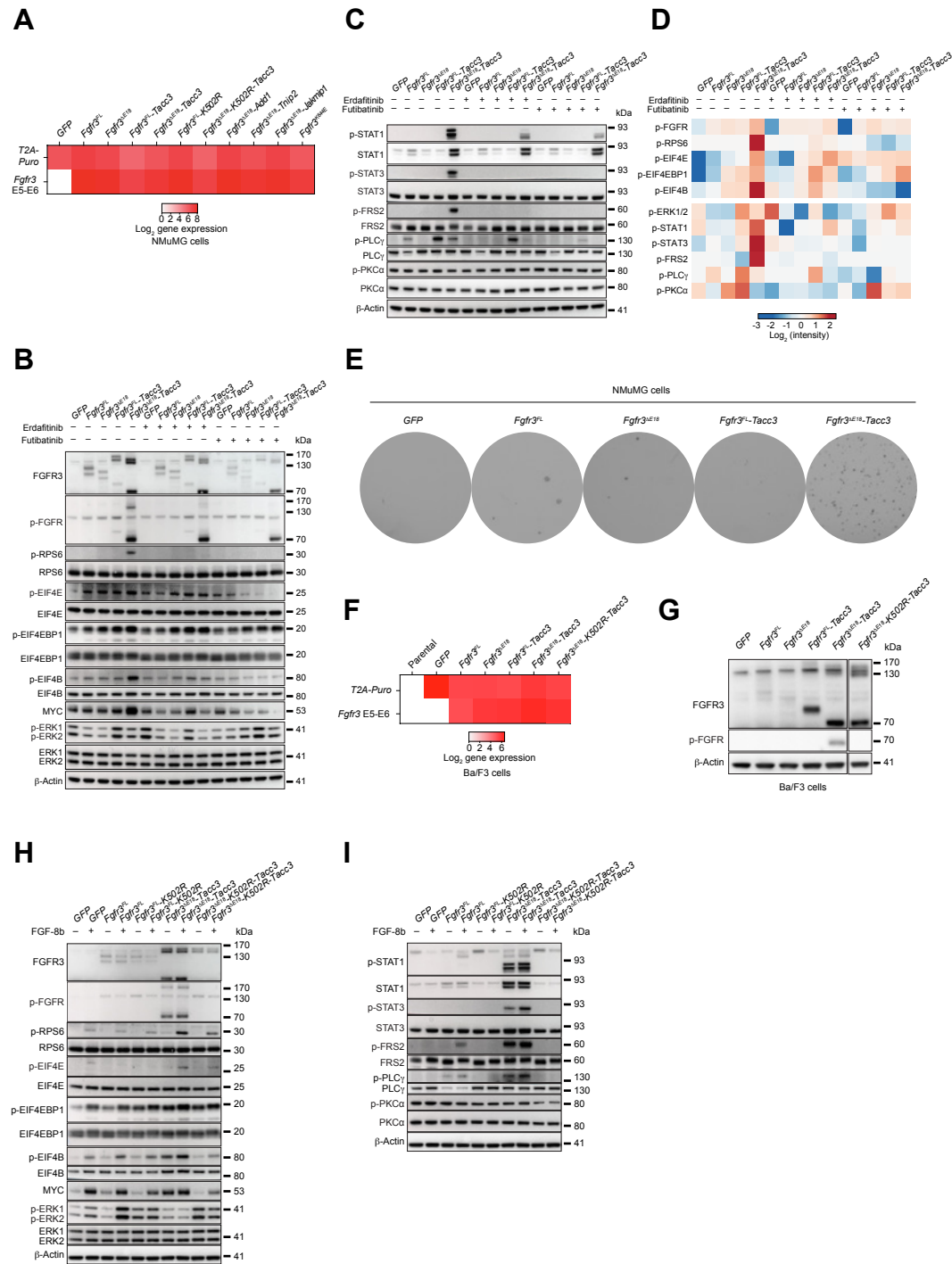


C



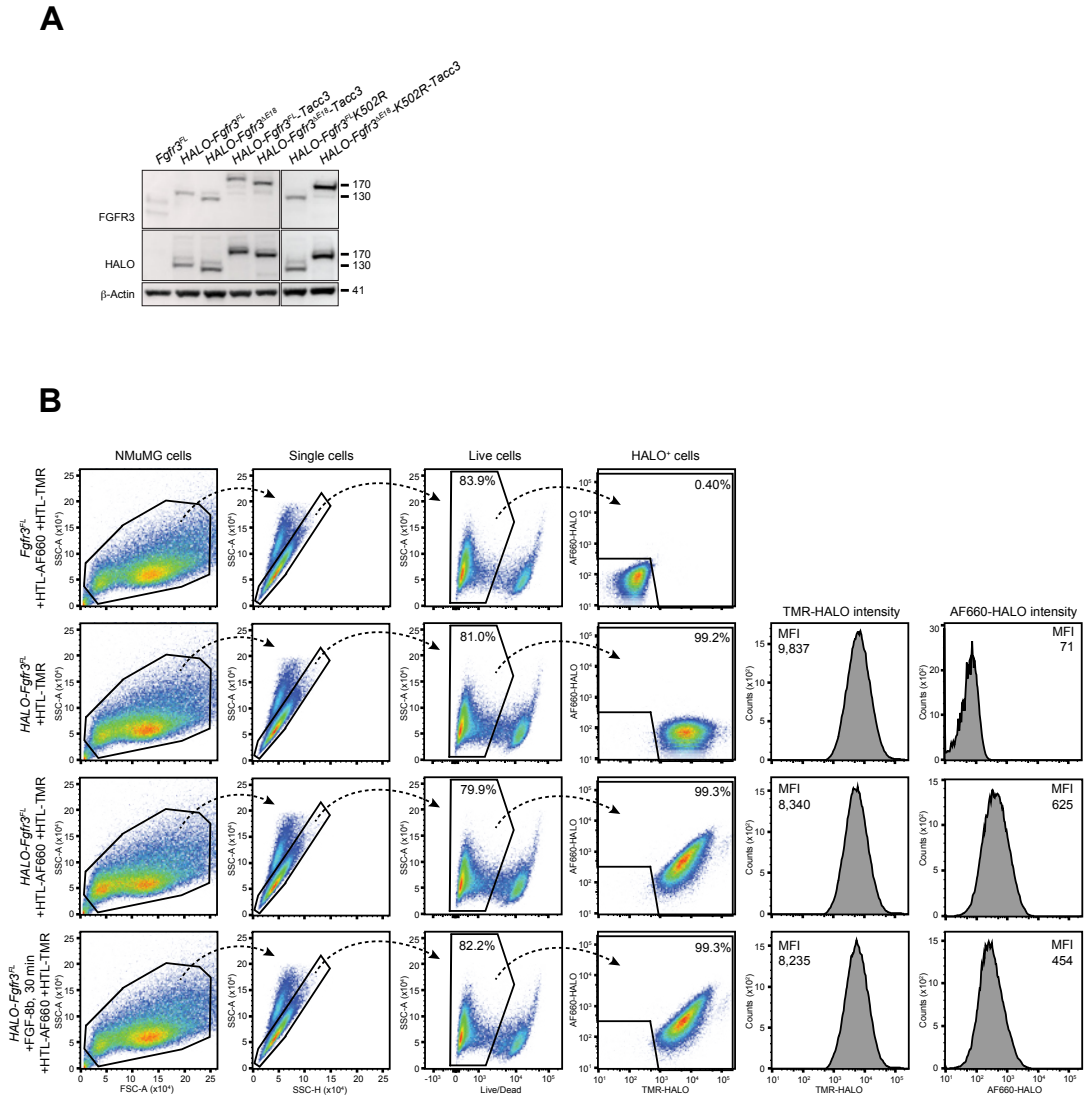
Supplementary Figure S1. Genomic alterations affecting *FGFR3*.

A,B, Lollipop plot (**A**) and normalized frequency (top panel) and enrichment significance (*P* values, bottom panel) (**B**) of *FGFR3*-truncating mutations identified in the FMI cohort. Mutation frequency was normalized by the kilobase of exon length and total number of mutations. AA, amino acid; in red, mutations affecting E18. **C,** Ratio of *FGFR3* junction reads spanning from E17 to E18 (E17-E18) versus E17-fusion partners (E17-fusion) for 17 of the RNA-seq profiles available from the HMF cohort.



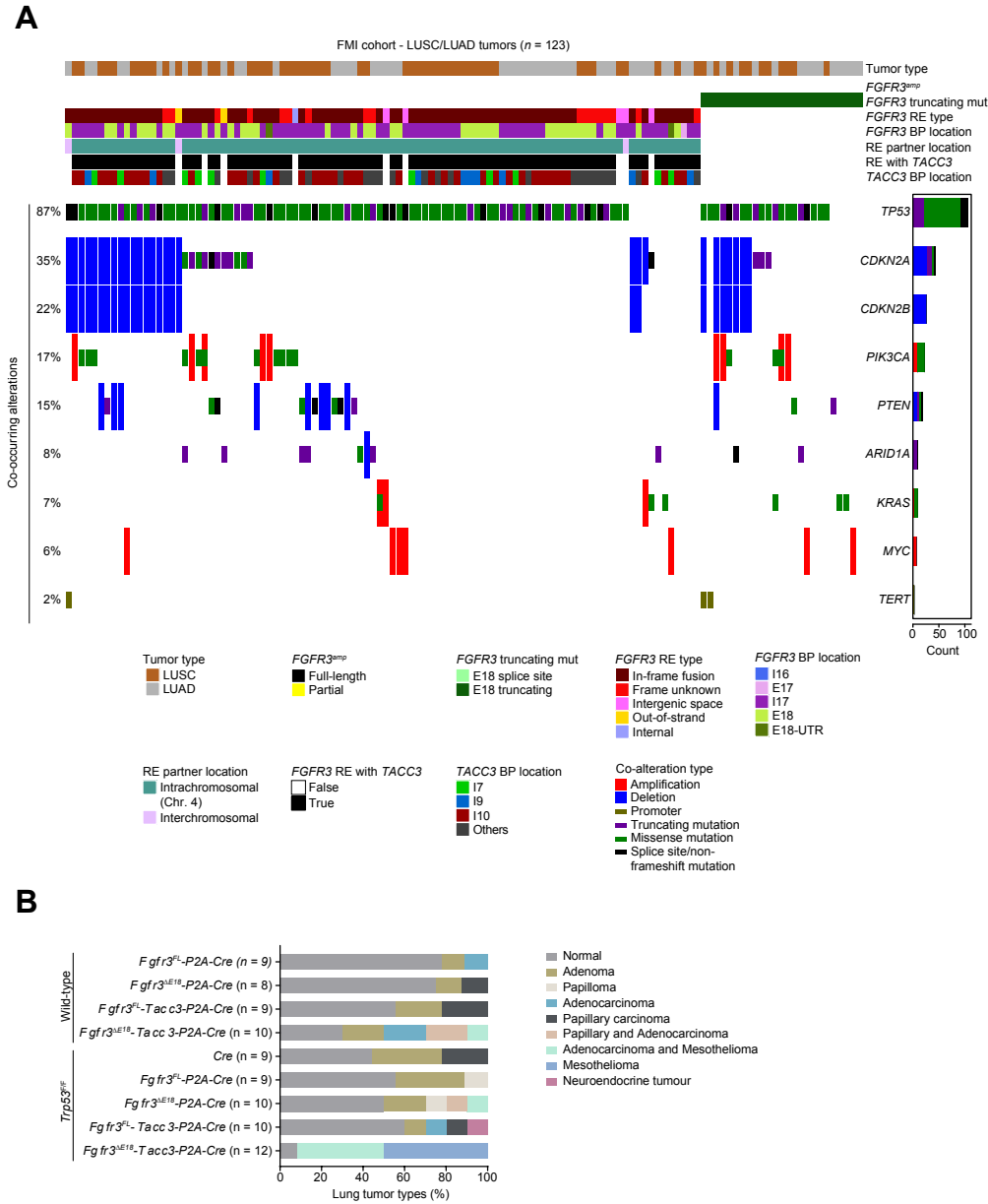
Supplementary Figure S2. Expression, FGFR3 pathway inhibition/activation, and transforming potential of *Fgfr3* variants in NMuMG and Ba/F3 cell models.

A, Validation of the overexpression of *Fgfr3* variants in NMuMG cells using RT-qPCR. T2A-Puro primers specifically detect *Fgfr3* transcripts derived from lentiviral constructs, *Fgfr3* E4-E5 primers (spanning the exon junction of exons 4 and 5 of *Fgfr3*) detect both endogenous transcripts and those derived from lentiviral constructs. Log2-transformed expression of each cDNA segment is normalized to *Usf1* housekeeping gene expression. Data are represented as mean of $n = 3$ technical replica per group of 1 experiment. **B, C**, Western blots showing expression and phosphorylation of indicated proteins in 24 hr serum-starved NMuMG cells stably expressing *GFP*, *Fgfr3^{FL}*, *Fgfr3^{ΔE18}*, *Fgfr3^{FL}-Tacc3*, or *Fgfr3^{ΔE18}-Tacc3* and treated for 3 hr with 100 nM erdafitinib or 100 nM futibatinib. β -Actin was run as sample processing control and is duplicated in (**B, C**). Weights of protein marker bands (kDa) are indicated. **D**, Quantifications of indicated phosphoprotein band intensities based on Western blots in (**B, C**). Phospho band intensities were normalized to their individual lane backgrounds and to their corresponding total protein band intensities and β -actin bands. **E**, Representative images of 12-well plate wells of 3D soft agar colony formation assay using NMuMG cells stably expressing *GFP*, *Fgfr3^{FL}*, *Fgfr3^{ΔE18}*, *Fgfr3^{FL}-Tacc3*, or *Fgfr3^{ΔE18}-Tacc3* and grown for 20 days. **F**, Validation of the overexpression of *Fgfr3* variants in Ba/F3 cells using RT-qPCR. Primers and analysis strategy used was the same as in (**A**). **G**, Western blots showing FGFR3 and phosphorylated FGFR protein expression in Ba/F3 cells stably expressing *GFP*, *Fgfr3^{FL}*, *Fgfr3^{ΔE18}*, *Fgfr3^{FL}-Tacc3*, *Fgfr3^{ΔE18}-Tacc3* or *Fgfr3^{ΔE18}-K502R-Tacc3*. β -Actin was run as sample processing control. Intermediate lanes containing samples not relevant to this figure were removed for clarity during image processing. A gap indicates where the blot was cropped, but all samples shown were run on the same gel and processed under identical conditions. No other alterations were made. **H, I**, Western blots showing expression and phosphorylation of indicated proteins in 24 hr serum-starved NMuMG cells stably expressing *GFP*, *Fgfr3^{FL}*, *Fgfr3^{FL}-K502R*, *Fgfr3^{ΔE18}-Tacc3* or *Fgfr3^{ΔE18}-K502R-Tacc3* and treated for 10 min with 20 ng/ml FGF-8b. β -Actin was run as sample processing control and is duplicated in (**H, I**). Molecular weights of protein marker bands in (**B, C, G-I**) are indicated in kDa, and all blots were stained with Ponceau S to ensure equal loading of total protein.



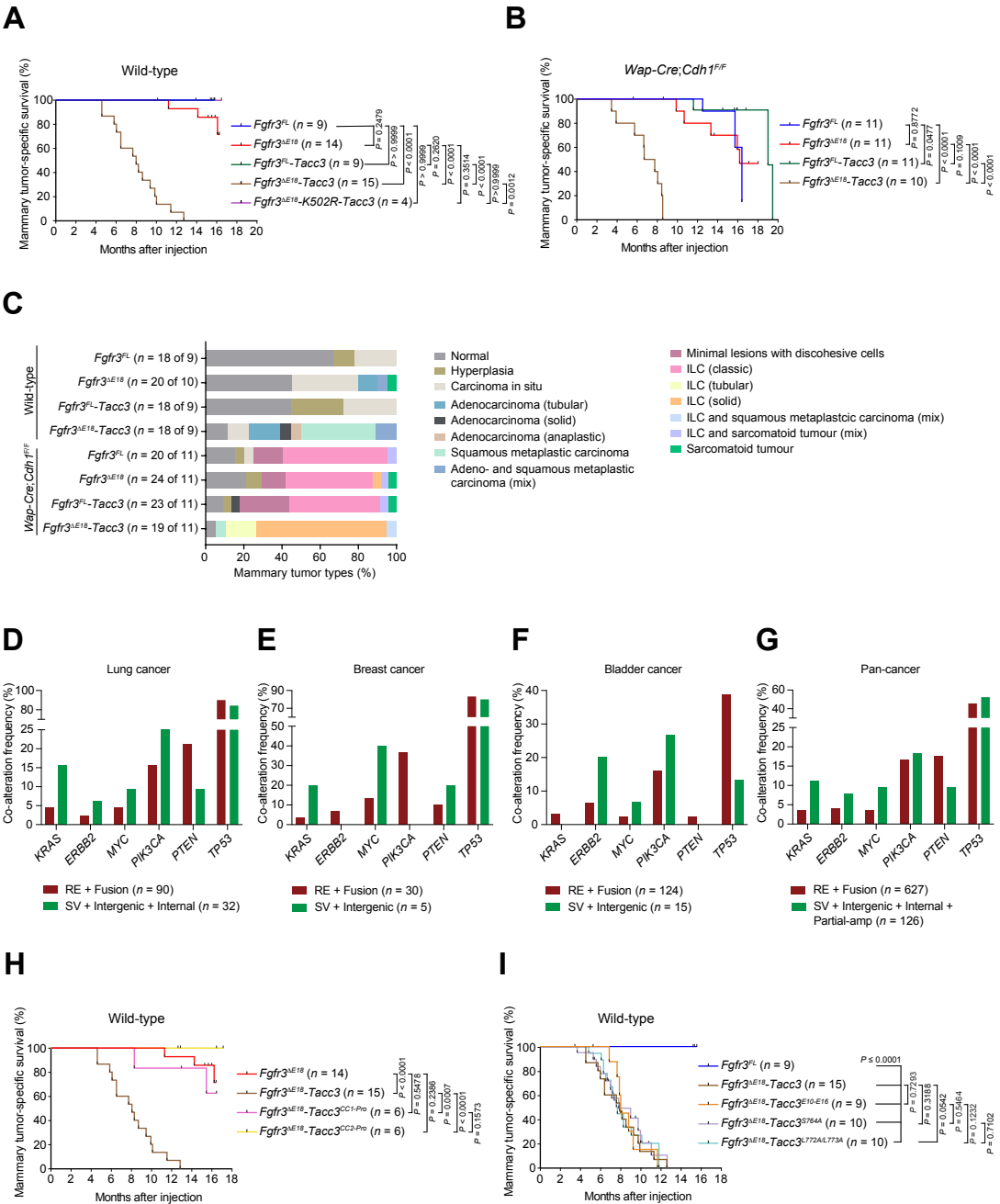
Supplementary Figure S3. Measurements of HALO-FGFR3 internalization.

A, Western blots showing expression of HALO-FGFR3 in NMuMG cells expressing *Fgfr3* or indicated HALO-*Fgfr3* variants used to perform the SPIDA assay. β -Actin was run as sample processing control on the same blot as FGFR3, and weights of protein marker bands (kDa) are indicated. **B**, FACS gating strategy used to quantify surface and internal HALO-tagged FGFR3 in serum-starved NMuMG cells expressing *Fgfr3*^{RL} or HALO-*Fgfr3*^{RL} constructs following the SPIDA assay. Where applicable, cells were stimulated with 20 ng/ml FGF-8b and sequentially labeled with a cell-impermeable HaloTag ligand conjugated to AF660, followed by a cell permeable HaloTag ligand conjugated to TMR. Gating was performed as follows: (i) FSC-A vs. SSC-A to identify the main cell population and exclude debris, (ii) SSC-H vs. SSC-A to isolate single cells, (iii) Live/Dead vs. SSC-A to gate live cells, and (iv) TMR-HALO vs. AF660-HALO to select for HALO-positive cells. Median fluorescence intensity (MFI) of AF660 and TMR signals was then quantified to assess surface-localized and internalized HALO-tagged FGFR3, respectively.



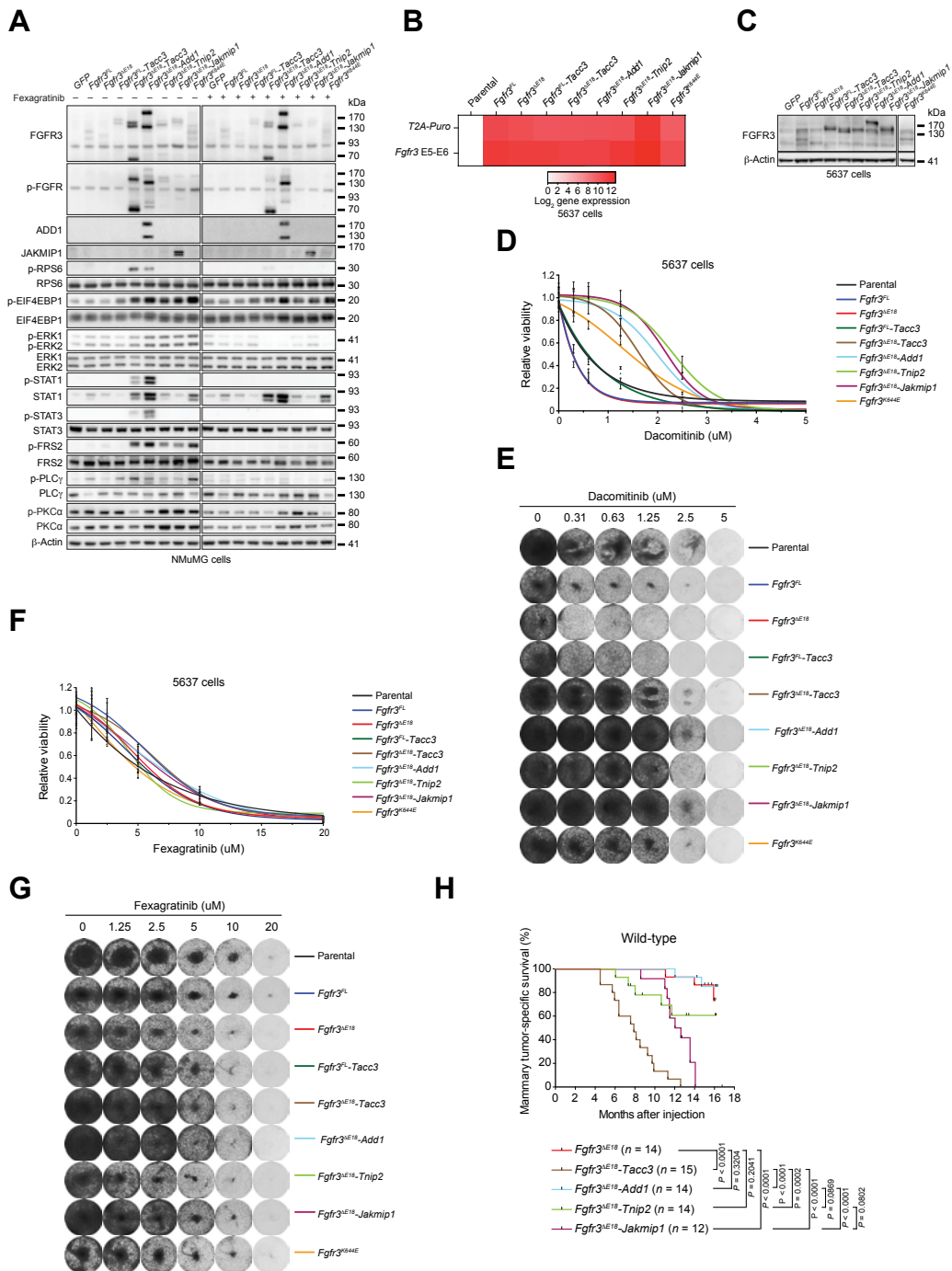
Supplementary Figure S4. Top driver genes co-occurring in human lung cancers with *FGFR3* E18 truncations, and histological examination of mouse lungs injected with *Fgfr3* variants.

A, Analysis of 123 LUSC/LUAD tumors (0.049% incidence) carrying potential *FGFR3* E18 truncations identified among 249,570 pan-cancer diagnostic hybrid-capture panel-seq profiles derived from FMI. These samples include *FGFR3* REs with BPs at the C-terminal tail, E17/18 or I16/17, or *FGFR3*-E18 truncating mutations. *FGFR3* REs consist of in-frame fusions, frame-unknown REs, intergenic space REs, out-of-strand REs, and internal REs. Tumor type, *FGFR3* amplification status, BP location of *FGFR3*, RE partner location, and REs with *TACC3* and its corresponding BP location are indicated. Top-9 co-occurring tumor driver genes and their corresponding alteration type are shown. **B**, Lung tumor type classifications of *Fgfr3* somatic mouse models based on hematoxylin and eosin (H&E) stainings. Cohort counts (n) represent evaluated mammary glands per number of mice.



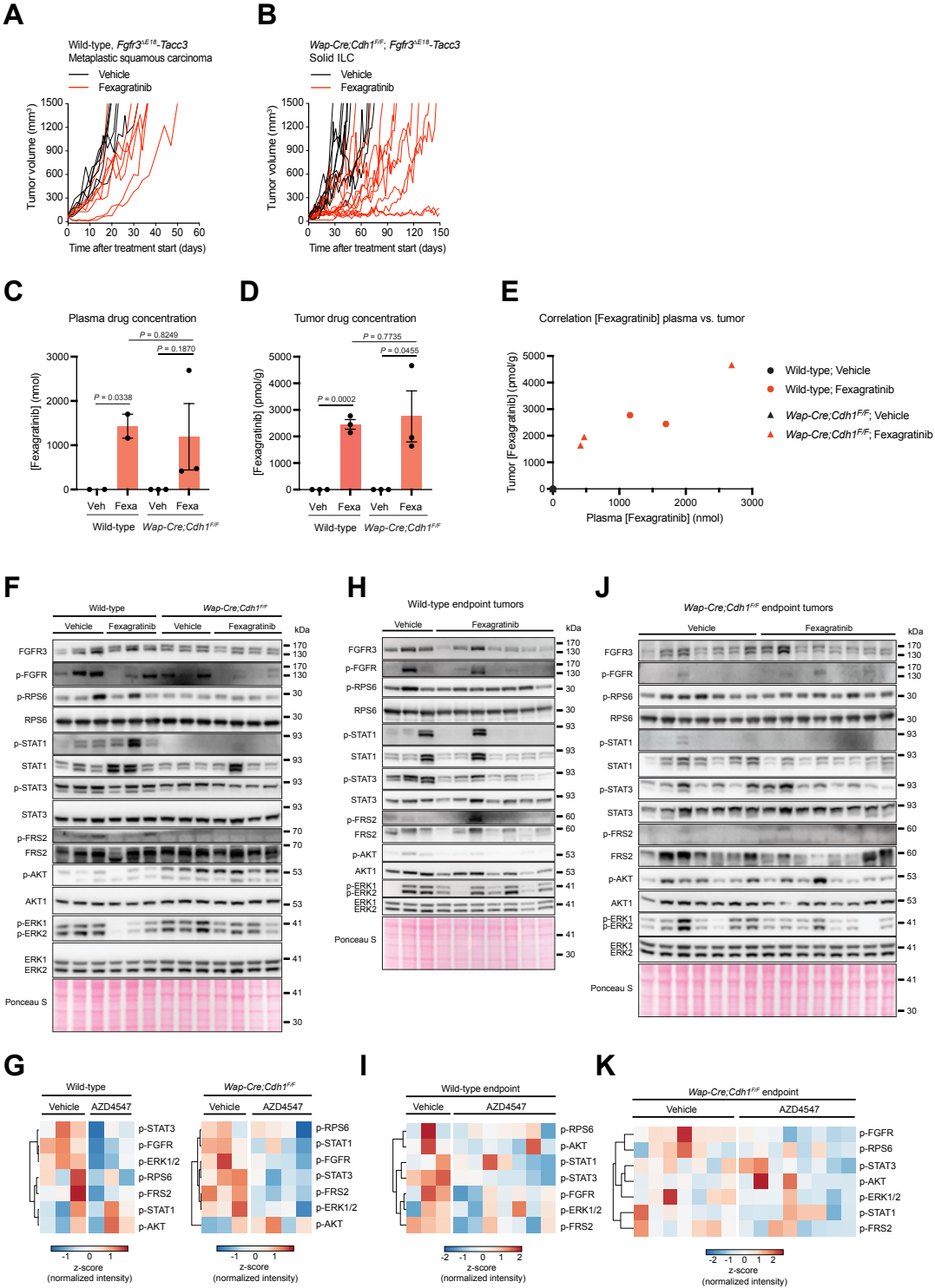
Supplementary Figure S5. Mammary tumor-specific survival and histological examinations of wild-type mice intraductally injected with *Fgfr3* variants, and co-mutational context of *FGFR3* E18 truncations.

A, Kaplan-Meier curves showing mammary tumor-specific survival of female wild-type mice intraductally injected with lentiviruses encoding indicated *Fgfr3* variants. *Fgfr3^{FL}*, *n* = 9; *Fgfr3^{ΔE18}*, *n* = 14; *Fgfr3^{FL}-Tacc3*, *n* = 9; *Fgfr3^{ΔE18}-Tacc3*, *n* = 15; *Fgfr3^{ΔE18}-K502R-Tacc3*, *n* = 4 mice. **B**, Kaplan-Meier curves showing mammary tumor-specific survival of female *Wap-Cre;Cdh1^{fl/fl}* mice intraductally injected with lentiviruses encoding indicated *Fgfr3* variants. *Fgfr3^{FL}*, *n* = 11; *Fgfr3^{ΔE18}*, *n* = 11; *Fgfr3^{FL}-Tacc3*, *n* = 11; *Fgfr3^{ΔE18}-Tacc3*, *n* = 10 mice. **C**, Mammary tumor type classifications of *Fgfr3* somatic mouse models based on H&E and E-cadherin immunohistochemistry (IHC) stainings. Cohort counts (*n*) represent evaluated mammary glands per number of mice. **D-G**, Enrichment of co-alterations in key tumor driver genes (*KRAS*, *ERBB2*, *MYC*, *PIK3CA*, *PTEN*, and *TP53*) across the indicated *FGFR3* E18-truncating alteration categories in the FMI lung (**D**), breast (**E**), bladder (**F**), and pan-cancer (**G**) cohorts. SV, short variants. **H**, Kaplan-Meier curves showing mammary tumor-specific survival of female wild-type mice intraductally injected with lentiviruses encoding indicated *Fgfr3* variants. *Fgfr3^{ΔE18}*, *n* = 14; *Fgfr3^{ΔE18}-Tacc3*, *n* = 15; *Fgfr3^{ΔE18}-Tacc3^{CC1-Pro}*, *n* = 6; *Fgfr3^{ΔE18}-Tacc3^{CC2-Pro}*, *n* = 6 mice. **I**, Kaplan-Meier curves showing mammary tumor-specific survival of female wild-type mice intraductally injected with lentiviruses encoding indicated *Fgfr3* variants. *Fgfr3^{FL}*, *n* = 9; *Fgfr3^{ΔE18}-Tacc3*, *n* = 15; *Fgfr3^{ΔE18}-Tacc3^{E10-E16}*, *n* = 9; *Fgfr3^{ΔE18}-Tacc3^{S764A}*, *n* = 10; *Fgfr3^{ΔE18}-Tacc3^{L772A/L773A}*, *n* = 10 mice. The *Fgfr3^{ΔE18}* curve in (**H**), the *Fgfr3^{FL}* curve in (**I**), and the *Fgfr3^{ΔE18}-Tacc3* curves in (**H**, **I**) are duplicated from (**A**). *P* values in (**A**, **B**, **H**, **I**) were calculated using log-rank (Mantel-Cox) tests.



Supplementary Figure S6. Functional *in vitro* and *in vivo* characterization of *Fgfr3*^{ΔE18}-*Add1*, *Fgfr3*^{ΔE18}-*Tnip2*, and *Fgfr3*^{ΔE18}-*Jakmip1* fusions.

A, Western blots showing expression and phosphorylation of the indicated proteins in NMuMG cells stably expressing GFP, *Fgfr3*^{FL}, *Fgfr3*^{ΔE18}, *Fgfr3*^{FL}-*Tacc3*, *Fgfr3*^{ΔE18}-*Tacc3*, *Fgfr3*^{ΔE18}-*Add1*, *Fgfr3*^{ΔE18}-*Tnip2*, *Fgfr3*^{ΔE18}-*Jakmip1* or *Fgfr3*^{K644E}, and treated for 3 hr with either vehicle or 1 μM fexagratinib following serum starvation for 24 hr. Intermediate lanes containing samples not relevant to this figure were removed for clarity. A gap indicates where the blot was cropped; samples shown were run on two different gels but imaged together and processed under identical conditions. No other alterations were made. **B**, Validation of the overexpression of *Fgfr3* variants in the 5637 bladder cancer cells using RT-qPCR. T2A-Puro primers specifically detect *Fgfr3* transcripts derived from lentiviral constructs, *Fgfr3* E4-E5 primers (spanning the exon junction of exons 4 and 5 of *Fgfr3*) detect both endogenous transcripts and those derived from lentiviral constructs. Log2-transformed expression of each cDNA segment is normalized to *Usp1* housekeeping gene expression. Data are represented as mean of *n* = 3 technical replica per group of 1 experiment. **C**, Western blots showing FGFR3 protein expression in the 5637 cells stably expressing GFP, *Fgfr3*^{FL}, *Fgfr3*^{ΔE18}, *Fgfr3*^{FL}-*Tacc3*, *Fgfr3*^{ΔE18}-*Tacc3*, *Fgfr3*^{ΔE18}-*Add1*, *Fgfr3*^{ΔE18}-*Tnip2*, *Fgfr3*^{ΔE18}-*Jakmip1*, or *Fgfr3*^{K644E}. **D**, Dose-response curves of 5637 cells (either untransduced parental or transduced with lentiviruses encoding the indicated *Fgfr3* variants), treated for 3 days with dacomitinib. Cell viability was determined using CellTiter-Blue. Data were normalized to untreated conditions within each cell line and are presented as mean ± SD of *n* = 2 independent replica from 1 representative experiment. Assays were performed at least 2 times. **E**, Representative well images of (**D**), obtained after fixing and staining the cells with crystal violet. **F**, Dose-response curves of 5637 cells as in (**D**), treated for 3 days with fexagratinib. **G**, Representative well images of (**F**). **H**, Kaplan-Meier curves showing mammary tumor-specific survival of female wild-type mice intraductally injected with lentiviruses encoding indicated *Fgfr3* variants. *Fgfr3*^{ΔE18}, *n* = 14; *Fgfr3*^{ΔE18}-*Tacc3*, *n* = 15; *Fgfr3*^{ΔE18}-*Add1*, *n* = 14; *Fgfr3*^{ΔE18}-*Tnip2*, *n* = 14; *Fgfr3*^{ΔE18}-*Jakmip1*, *n* = 12 mice. The *Fgfr3*^{ΔE18} and the *Fgfr3*^{ΔE18}-*Tacc3* curves are duplicated from Supplementary Fig. S5A. *P* values were calculated using log-rank (Mantel-Cox) tests. In (**A**, **C**), β-Actin was run as sample processing control, blots were stained with Ponceau S to ensure equal loading of total protein, and molecular weights of protein marker bands are indicated in kDa.



Supplementary Figure S7. Tumor growth response, target engagement, and adaptive resistance following fexagratinib treatment in *Fgfr3^{ΔE18}-Tacc3*-driven mouse mammary tumors.

A, B, Individual growth curves of indicated tumor donors transplanted into the mammary fat pad of female syngeneic wild-type (**A**) or *Wap-Cre;Cdh1^{fl/fl}* (**B**) mice and treated daily orally with vehicle or 12.5 mg/kg fexagratinib using an intermittent dosing regimen. **C,** Drug concentration in plasma of vehicle- and fexagratinib-treated mice. Each dot represents an individual mouse; bars indicate mean $\pm$ SEM. Veh, vehicle; Fexa, fexagratinib. **D,** Fexagratinib concentration in the tumor tissue of the same animals. Statistical comparisons in (**C, D**) were performed using an unpaired two-tailed t-test. **E,** Correlation between plasma and tumor drug concentrations in the two donor-derived mouse cohorts shown in (**C, D**). Each dot/triangle represents a paired measurement from an individual mouse. **F,** Western blots showing expression and phosphorylation of indicated proteins in mammary tumors from wild-type and *Wap-Cre;Cdh1^{fl/fl}* donors treated for 2 or 5 days with vehicle or 12.5 mg/kg/day fexagratinib. **G,** Quantifications of indicated phosphoprotein band intensities based on Western blots in (**F**). Data from 2- and 5-day treatment groups in (**C-G**) were combined for clarity, as no significant differences were observed in FGFR3 signaling or fexagratinib concentrations in plasma and tumor between the two time points. **H,** Western blots showing expression and phosphorylation of indicated proteins in vehicle and fexagratinib-treated outgrown endpoint mammary tumors from a wild-type donor. **I,** Quantifications of indicated phosphoprotein band intensities based on Western blots in (**H**). **J,** Western blots showing expression and phosphorylation of indicated proteins in vehicle and fexagratinib-treated outgrown endpoint mammary tumors from a *Wap-Cre;Cdh1^{fl/fl}* donor. **K,** Quantifications of indicated phosphoprotein band intensities based on Western blots in (**J**). Ponceau S staining is shown as a total protein loading control in (**F, H** and **J**). In (**G, I, K**), phospho band intensities were normalized to their corresponding total protein band intensities and β -actin bands. Values were then normalized to vehicle-treated controls and variance-scaled across samples.

6

CHAPTER 6

MYC IS A CLINICALLY SIGNIFICANT DRIVER OF MTOR INHIBITOR RESISTANCE IN BREAST CANCER

Jinhyuk Bhin*, **Julia Yemelyanenko***, Xue Chao, Sjoerd Klarenbeek, Mark Opdam, Yuval Malka, Liesbeth Hoekman, Dinja Kruger, Onno Bleijerveld, Chiara S. Brambillasca, Justin Sprengers, Bjørn Siteur, Stefano Annunziato, Matthijs J. van Haren, Nathaniel I. Martin, Marieke van de Ven, Dennis Peters, Reuven Agami, Sabine C. Linn, Epie Boven, Maarten Altelaar, Jos Jonkers#, Daniel Zingg# & Lodewyk F.A. Wessels#

*equal contribution

#co-corresponding author

Journal of Experimental Medicine, 2023 Nov 6;220(11):e20211743
<https://doi.org/10.1084/jem.20211743>

ABSTRACT

Targeting the PI3K–AKT–mTOR pathway is a promising therapeutic strategy for breast cancer treatment. However, low response rates and development of resistance to PI3K–AKT–mTOR inhibitors remain major clinical challenges. Here, we show that MYC activation drives resistance to mTOR inhibitors (mTORi) in breast cancer. Multiomic profiling of mouse invasive lobular carcinoma (ILC) tumors revealed recurrent *Myc* amplifications in tumors that acquired resistance to the mTORi AZD8055. MYC activation was associated with biological processes linked to mTORi response and counteracted mTORi-induced translation inhibition by promoting translation of ribosomal proteins. In vitro and in vivo induction of MYC conferred mTORi resistance in mouse and human breast cancer models. Conversely, AZD8055-resistant ILC cells depended on MYC, as demonstrated by the synergistic effects of mTORi and MYCi combination treatment. Notably, MYC status was significantly associated with poor response to everolimus therapy in metastatic breast cancer patients. Thus, MYC is a clinically relevant driver of mTORi resistance that may stratify breast cancer patients for mTOR-targeted therapies.

INTRODUCTION

The PI3K–AKT–mTOR signaling pathway is central to multiple cellular processes and is frequently dysregulated in human cancer. In breast cancer, the PI3K–AKT–mTOR pathway is often activated by genomic abnormalities, most commonly by *PIK3CA* hotspot mutations, *PTEN* copy number loss, or activation of upstream signaling cues derived from receptor tyrosine kinases [1,2]. High prevalence of PI3K–AKT–mTOR pathway activation in breast cancer has guided clinical trials evaluating small-molecule compounds targeting this pathway [3]. This has resulted in the approval of two drugs for breast cancer, everolimus [4] and alpelisib [5], targeting mTOR and PIK3CA, respectively.

Although PI3K–AKT–mTOR signaling blockade is a promising therapeutic strategy for breast cancer, predicting patient response using biomarkers has remained challenging, which has compromised the effectiveness of these targeted therapies [6,7]. For the everolimus–exemestane combination treatment in estrogen receptor–positive (ER+) breast cancer, the clinical benefit rate, defined as the proportion of patients with a complete response (CR), partial response (PR), or stable disease (SD) for at least 16 wk, was about 51%, and the overall response rate (CR + PR) was about 16% [4]. This highlights the need for biomarkers to guide patient stratification for this treatment. Moreover, the acquisition of resistance to inhibitors targeting PI3K or mTOR is a major clinical obstacle, potentially induced by multiple types of resistance mechanisms [8,9]. *PIK3CA* and *MTOR* gatekeeper and non-gatekeeper mutations can give rise to resistance [10–13]. Activation of independent proliferation pathways [14,15] and alterations of downstream targets, such as eIF4E and 4E-BP1 [16–18], may represent alternative strategies for cancers to acquire resistance to PI3K and mTOR inhibition. However, most of these findings have been identified in in vitro models and/or by the use of preclinical mTOR inhibitors (mTORi); thus the in vivo and clinical relevance of these findings for breast cancer has remained elusive.

Oncogenomic studies have demonstrated frequent activation of PI3K–AKT–mTOR signaling in the invasive lobular carcinoma (ILC) histological subtype of breast cancers, suggesting this pathway to be a particularly promising therapeutic target for ILC [1,19,20]. Conceivably, we previously observed strong activation of and dependence on PI3K–AKT–mTOR signaling in the transplantable *K14-Cre;Cdh1^{F/F};Trp53^{F/F}* (KEP) mouse model of ILC [21,22], thus representing an optimal in vivo model system to study mechanisms of resistance during mTOR blockade. Using the transplantable KEP model, we here performed multiomic analyses of ILC tumors obtained during long-term mTORi treatment to study in vivo mechanisms of acquired resistance to mTOR inhibition in breast cancer. We uncovered that MYC activation is a hallmark of mTORi resistance, coordinating cell-intrinsic and -extrinsic processes involved in response to mTOR inhibition, including the compensation of suppressed protein translation by mTORi. Orthotopic modeling of mouse ILC tumors and human breast cancer–derived xenografts confirmed that MYC is an in vivo and clinically

relevant driver of mTORi resistance. We also showed that MYC status is significantly associated with clinical response to the mTORi everolimus in breast cancer patients, thus establishing MYC as a clinically significant driver of mTORi resistance.

RESULTS

Molecular profiling of AZD8055-resistant KEP tumors

We previously established an *in vivo* experimental setting to investigate mTORi resistance in the transplantable KEP mouse model of ILC [22–24]. In this setting, immunocompetent mice bearing orthotopically transplanted KEP tumors were enrolled into an intervention study using the mTORi AZD8055, and mammary tumor samples were harvested from vehicle-treated control tumors, sensitive tumors after 5 d of treatment, and resistant tumors that progressed under long-term treatment (Fig. 1 A). Tumors from the three cohorts were subjected to multiomic analyses, that is, low-coverage whole genome sequencing (LC-WGS), RNA sequencing (RNA-seq), reverse phase protein array (RPPA) profiling, mass spectrometry (MS)–based expression proteomics, and MS-based phosphoproteomics (Fig. S1 A). For MS, we quantified 7,003 proteins and 5,141 phosphosites on average using the Tandem Mass Tag (TMT) isobaric labeling approach (Fig. S1, B and C). Together, we established a high-quality multiomic data collection to assess genomic and (phospho) proteomic changes during mTORi sensitivity and resistance.

Consistent with previous RPPA analysis [23], MS-phosphoproteomic data corroborated sustained inhibition of mTOR downstream targets in the resistant tumors, indicating that the acquired resistance to AZD8055 was not driven by the recovery of mTOR activity (Fig. 1 B). Phosphosite-specific signature analysis using PTMSigDB [25] demonstrated that both sensitive and resistant tumors had low enrichment of the PI3K–AKT signature and high enrichment of the rapamycin signature as compared with control tumors, consistently supporting sustained suppression of the PI3K–AKT–mTOR signaling in the resistant tumors. Moreover, suppression of cell cycle kinase signatures (e.g., CDK1, CDK2, and CDK7) in sensitive tumors and re-enrichment of these signatures in resistant tumors indicated that resistant tumors were actively proliferating (Fig. 1 C). The only mTOR downstream target that regained phosphorylation in resistant tumors was EIF4B (Fig. S1 D). However, overexpression of EIF4B and its phospho-mimicking variants (S406D, S422D) in the KEP tumor–derived cell line KEP2.E3 [23] did not confer resistance to AZD8055 *in vitro* (Fig. S1, E and F). Likewise, overexpression of MELK, which can act as an alternative upstream activator to EIF4B [26] and which was reactivated in resistant tumors as compared with sensitive tumors, had no effect on AZD8055 sensitivity in KEP2.E3 (Fig. S1, D–F). This suggested that recovery of EIF4B and MELK phosphorylation represented a marker of proliferative tumor progression rather than a driver of resistance to AZD8055.

***Myc* is recurrently amplified in AZD8055-resistant KEP tumors**

Next, we analyzed genome-wide copy number profiles of control, sensitive, and resistant KEP tumors using LC-WGS. Interestingly, *Myc* was focally amplified in the majority of the resistant tumors (10/13), while control and sensitive tumors harbored normal *Myc* copy numbers (Fig. 1 D). The three remaining resistant tumors contained a focal amplification of chromosome 2 (Chr-2; 101,638,292–105,126,510; Fig. 1 D). The amplified region covered 35 protein-coding genes including translation initiation factors *Eif3m* and *Eif5* that are involved in processes such as cell adhesion, proteolysis, and immune responses (Fig. 1 D and Fig. S1, G and H). RNA-seq revealed that *Myc* expression was significantly upregulated in the resistant tumors as compared with control and sensitive tumors (Fig. 1 E), and *Myc* copy number amplification was highly correlated with *Myc* upregulation (Fig. 1 F). We next assessed the protein level of MYC using RPPA data, demonstrating a higher abundance of MYC in the resistant tumors as compared with control and sensitive tumors (Fig. 1 G). Finally, we quantified *Myc* expression in tumor cells using immunohistochemistry (IHC), which consistently revealed MYC upregulation in tumor cells from resistant samples versus tumor cells from control or sensitive samples (Fig. 1, H and I). Together, these data demonstrated that prolonged inhibition of mTOR fosters the genomic evolution of tumor cells resulting in the amplification and upregulation of MYC.

MYC governs the transcriptional landscape of AZD8055-resistant tumors

To explore the biological processes underlying resistance to prolonged mTOR inhibition and MYC upregulation, we performed gene set variation analysis (GSVA; [27]) for 50 MSigDB hallmark gene sets [28] using the RNA-seq profiles (Fig. 2 A). GSVA displayed the estimated activities of the gene sets across the tumor groups (Fig. S2, A and B). Compared with control and sensitive tumors, MYC_TARGETS_V1 and V2 displayed significantly higher scores in the resistant tumors (Fig. 2 B). This indicated functional activation of MYC in AZD8055-resistant tumors. We next performed transcription factor (TF) target enrichment analysis using the TF target information collected from previously published genome-wide chromatin immunoprecipitation (ChIP) assays and motif-based prediction, as deposited in the EnrichR resource [29]. Upregulated genes in resistant tumors showed the strongest enrichment for genes containing binding sites for MYC as well as MYC-binding motifs (Fig. 2, C and D). Finally, MS-expression proteomics data similarly revealed the activation of MYC targets in resistant tumors as compared with control and sensitive tumors (Fig. 2, E and F). Taken together, these data show that during long-term AZD8055 treatment, MYC governs a transcriptional program potentially driving resistance to mTORi.

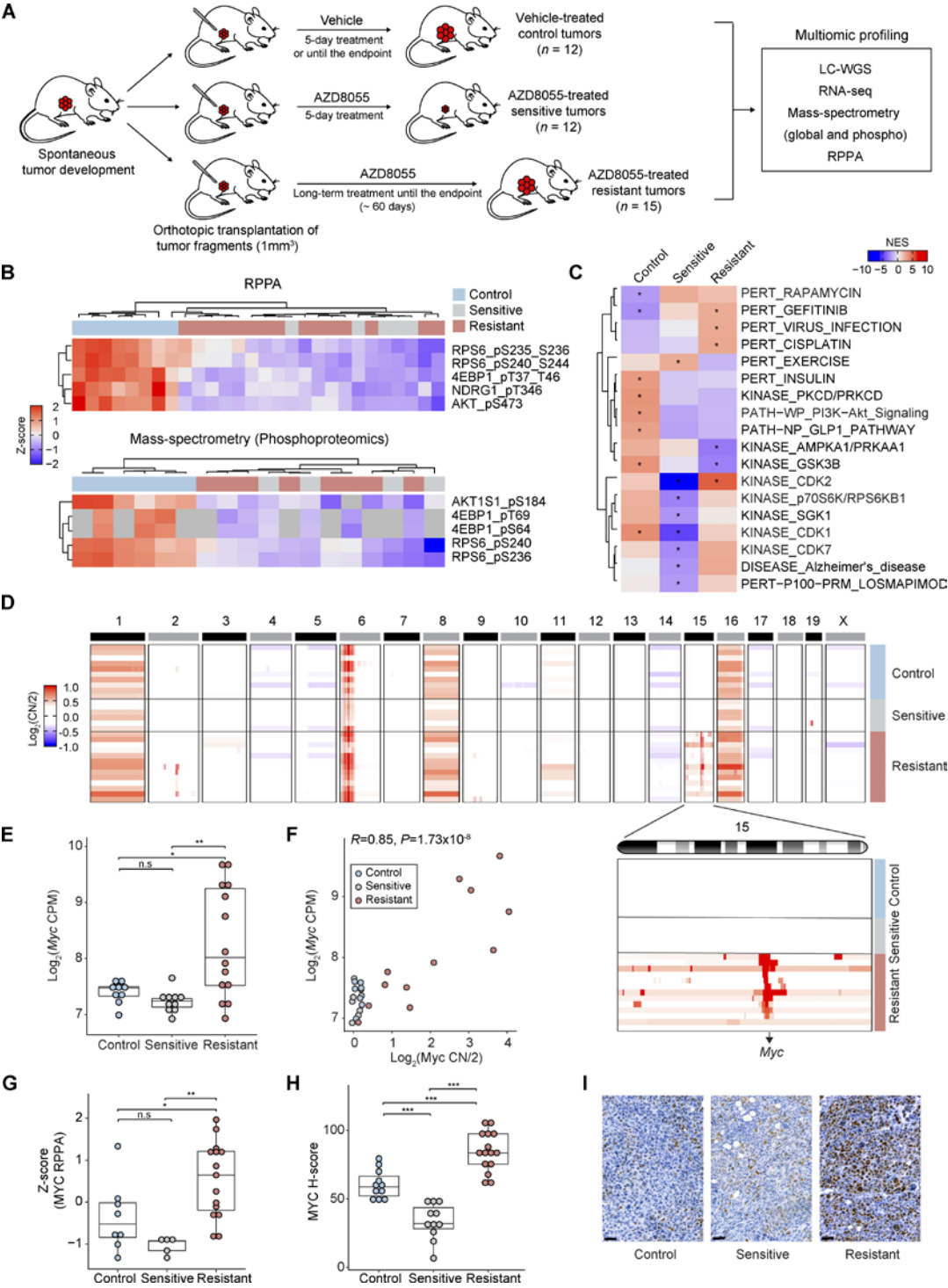


Figure 1. Upregulation of MYC in AZD8055-resistant KEP tumors.

(A) Scheme representing the experimental setup to generate vehicle-treated control and AZD8055-treated sensitive and resistant KEP mammary tumors used for the multiomic data analyses. Endpoint was reached when mammary tumor volume was $>1,500 \text{ mm}^3$. (B) mTOR downstream phosphosites measured by RPPA (8 control, 5 sensitive, and 15 resistant KEP tumors) and MS-based phosphoproteomics (6 control, 4 sensitive, and 8 resistant KEP tumors). Intensities from each assay were standardized using Z-scores. Gray cells in MS data indicate missing values. (C) Phosphosite-specific signature analysis based on the PTMSigDB database [25]. The average intensities for each KEP tumor group were used to identify representative signatures. *FDR < 0.05 . (D) Copy number (CN) ratio ($\log_2 \text{CN}/2$) of 10 control, 6 sensitive, and 13 resistant KEP tumors for all chromosomes and chromosome 15 (inset) highlighting focal MYC amplifications in resistant tumors. (E) Myc mRNA expression ($\log_2 \text{CPM}$) across 9 control, 10 sensitive, and 14 resistant KEP tumors. (F) Pearson's correlation analysis of Myc copy number ratio ($\log_2 \text{CN}/2$) versus Myc mRNA expression ($\log_2 \text{CPM}$). P value was calculated with two-tailed t -transformations of Pearson's correlation coefficient. (G) Z-score standardized MYC protein levels across 8 control, 5 sensitive, and 15 resistant KEP tumors as measured by RPPA. (H) H-score quantifications of MYC staining intensities across 13 control, 11 sensitive, and 15 resistant tumors. (I) Representative MYC IHC stainings of control, sensitive, and resistant KEP tumors in H. Scale bars, 50 μm . In E–H, data are represented as median $\pm$ interquartile range (IQR; box) and quartiles $\pm 1.5 \times \text{IQR}$ (whiskers), and one-way ANOVA and Tukey's post-hoc test were performed to compute adjusted P values (*P < 0.05 ; **P < 0.01 ; ***P < 0.001 ; n.s., not significant). Source data are available for this figure: SourceData F1.

MYC directs biological processes associated with mTORi response

To gain insight into the resistance-associated transcriptional programs mediated by MYC, we further explored the GSVA results (Fig. 2 A). Apart from terms directly connected to MYC activity, we also observed dynamic changes of immune-related pathways in the process of acquiring AZD8055 resistance (Fig. S2, A and B). The interferon α/γ response pathways were significantly downregulated in resistant tumors (Fig. 3 A), indicating a potential role of MYC in suppressing interferon signaling [30–33]. TF target enrichment analysis demonstrated that downregulated genes in resistant tumors were significantly enriched by the targets of immune-associated transcription regulators such as IRF8 and SMRT, supporting the suppression of immune response in resistant tumors (Fig. 2, C and D). Likewise, other immune processes such as IL6–JAK–STAT3 signaling, IL2–STAT5 signaling, inflammatory response, and complement were similarly triggered by short-term treatment of AZD8055 but became suppressed during resistance acquisition (Fig. 3 A and Fig. S2 C). Moreover, immune infiltration as assessed by Estimation of Stromal and Immune cells in Malignant Tumours using Expression data (ESTIMATE; [34]) was predicted to be enriched in sensitive tumors, while resistant tumors showed an immune exclusion phenotype (Fig. 3 B). We experimentally validated the relationship between MYC and immune cell infiltration using IHC. Quantification of MYC intensities in tumor cells and densities of immune cell types in consecutive slides revealed a significant negative correlation between MYC expression and infiltration of CD3⁺, CD4⁺, FOXP3⁺, pSTAT1⁺, and MHCII⁺ immune cells (Fig. 3 C). These immune cell types were particularly enriched in tumors sensitive to AZD8055 but became excluded in AZD8055-resistant tumors (Fig. 3 D and Fig. S2 D). Our results suggested that MYC-driven immune exclusion contributes to AZD8055 resistance, in line with the previous finding that maintenance of AZD8055 sensitivity in KEP tumors partly depends on the adaptive immune system [23].

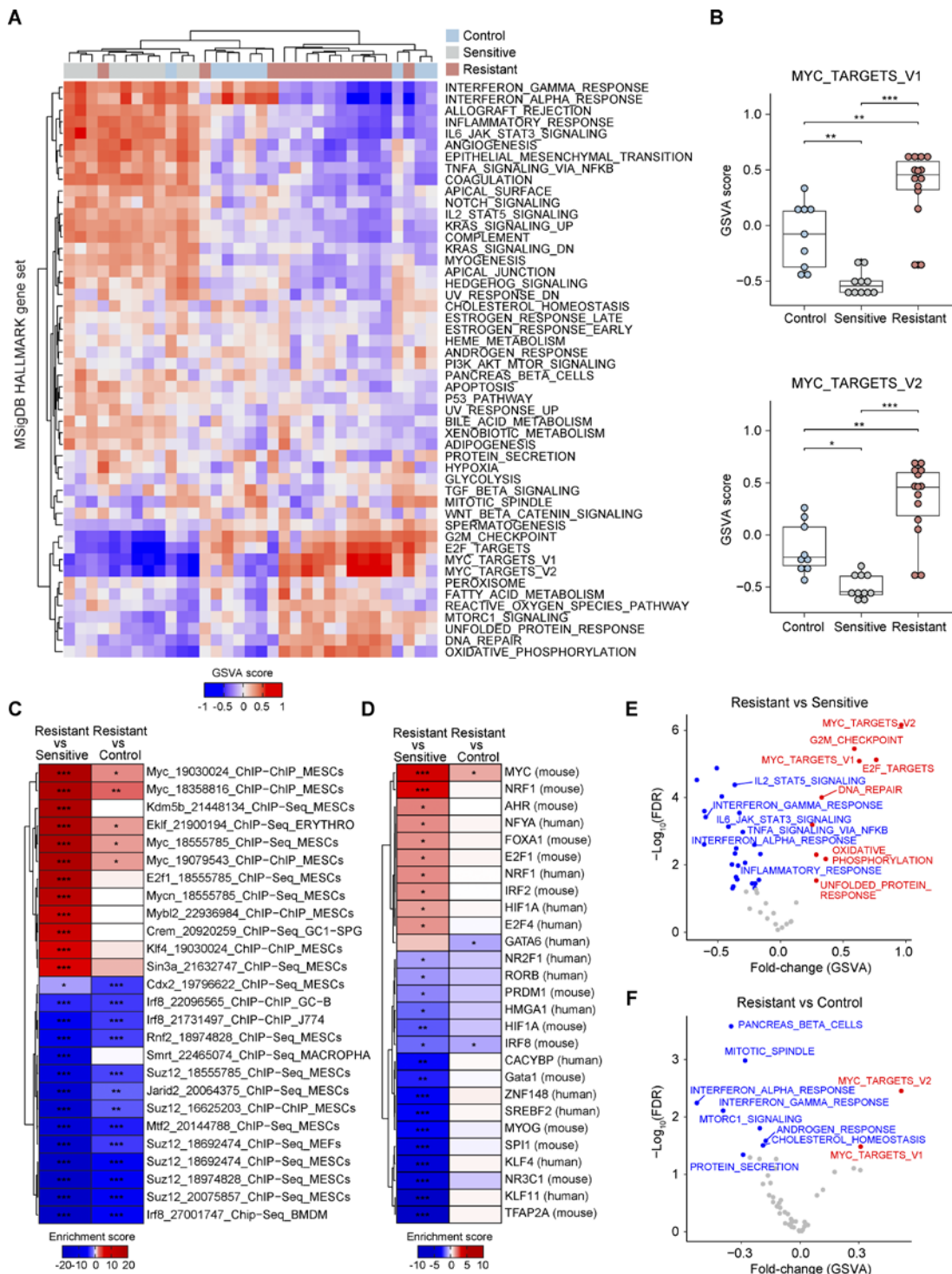


Figure 2. Functional MYC activity in AZD8055-resistant tumors.

(A) GSEA analysis for 50 MSigDB Hallmark gene sets across RNA-seq profiles derived from 9 control, 10 sensitive, and 14 resistant KEP tumors. Hierarchical clustering (Euclidean distance and complete-linkage clustering) was performed for GSEA scores. **(B)** GSEA scores for the two MSigDB Hallmark MYC target gene sets. Data are represented as median $\pm$ IQR (box) and quartiles $\pm 1.5 \times$ IQR (whiskers), and one-way ANOVA and Tukey's post-hoc test were performed to compute adjusted P values (* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$). **(C and D)** TF target and motif enrichment analyses for upregulated and downregulated genes in resistant vs. sensitive and resistant vs. control comparisons (RNA-seq profiles derived from 9 control, 10 sensitive, and 14 resistant KEP tumors). TF targets and motifs were obtained from ChEA2016 (C) and TRANSFAC_and_JASPAR_PWMs (D), all collected in the EnrichR database [29]. TF enrichment scores were defined as $-\log_{10}(\text{FDR})$ for upregulated genes in each comparison and as the additive inverse of $-\log_{10}(\text{FDR})$ for downregulated genes in each comparison and computed by Fisher's exact test followed by Benjamini & Hochberg correction. Row labels in C, "TF name", "Pubmed ID", "experimental assay", "cell line". *FDR < 0.05 ; **FDR < 0.01 ; ***FDR < 0.001 . **(E and F)** GSEA analysis for resistant versus sensitive (E) and resistant versus control (F) tumors using MS-based expression proteomic data (6 control, 4 sensitive, and 8 resistant KEP tumors). GSEA scores were compared using Student's *t* test. Red and blue dots, significantly (FDR < 0.05) up- and downregulated in resistant tumors.

Besides immune-related processes, GSEA also revealed oxidative phosphorylation, the reactive oxygen species pathway, DNA repair, and fatty acid metabolism to be specifically upregulated in resistant tumors (Fig. 3 E). GSEA using Gene Ontology (GO) gene sets also showed upregulation of glutamine metabolism in resistant tumors (Fig. 3 F and Fig. S2 E). Enhanced mitochondrial energy metabolism and glutaminolysis were previously connected to mTORi resistance [35,36]. The glutaminolysis rate-limiting enzymes *Gls* and *Gls2* were upregulated in the resistant tumors, indicating enhanced glutamine-fueled oxidative phosphorylation in resistant tumors (Fig. S2 F). MYC can promote glutamine metabolism through transcriptional regulation of genes involved in glutaminolysis and glutamine uptake [37,38]. Especially, *Gls2* expression was significantly correlated with *Myc* expression ($P = 7.6 \times 10^{-3}$) in resistant tumors and *Gls2* was upregulated in KEP cell lines upon *Myc* overexpression, suggesting that MYC stimulated the expression of *Gls2* (Fig. S2 G).

MYC can also act as a master regulator of translation by directly regulating the transcription of the genes involved in translation initiation and elongation, transfer RNA (tRNA) synthetases, and ribosomal proteins [39,40]. GSEA using GO gene sets demonstrated the upregulation of genes involved in protein translation initiation and elongation as well as ribosome biogenesis in resistant tumors, and the scores of these terms significantly correlated with *Myc* expression (Fig. 3, G and H; and Fig. S2 E). Intriguingly, gene sets connected to protein translation initiation were strongly diminished in AZD8055-sensitive tumors over controls (Fig. S2 E). Thus, MYC appeared to foster AZD8055 resistance by counteracting mTORi-mediated translation inhibition. Taken together, MYC regulated a range of tumor cell-intrinsic and -extrinsic biological processes associated with mTORi response and thus governed mTORi resistance.

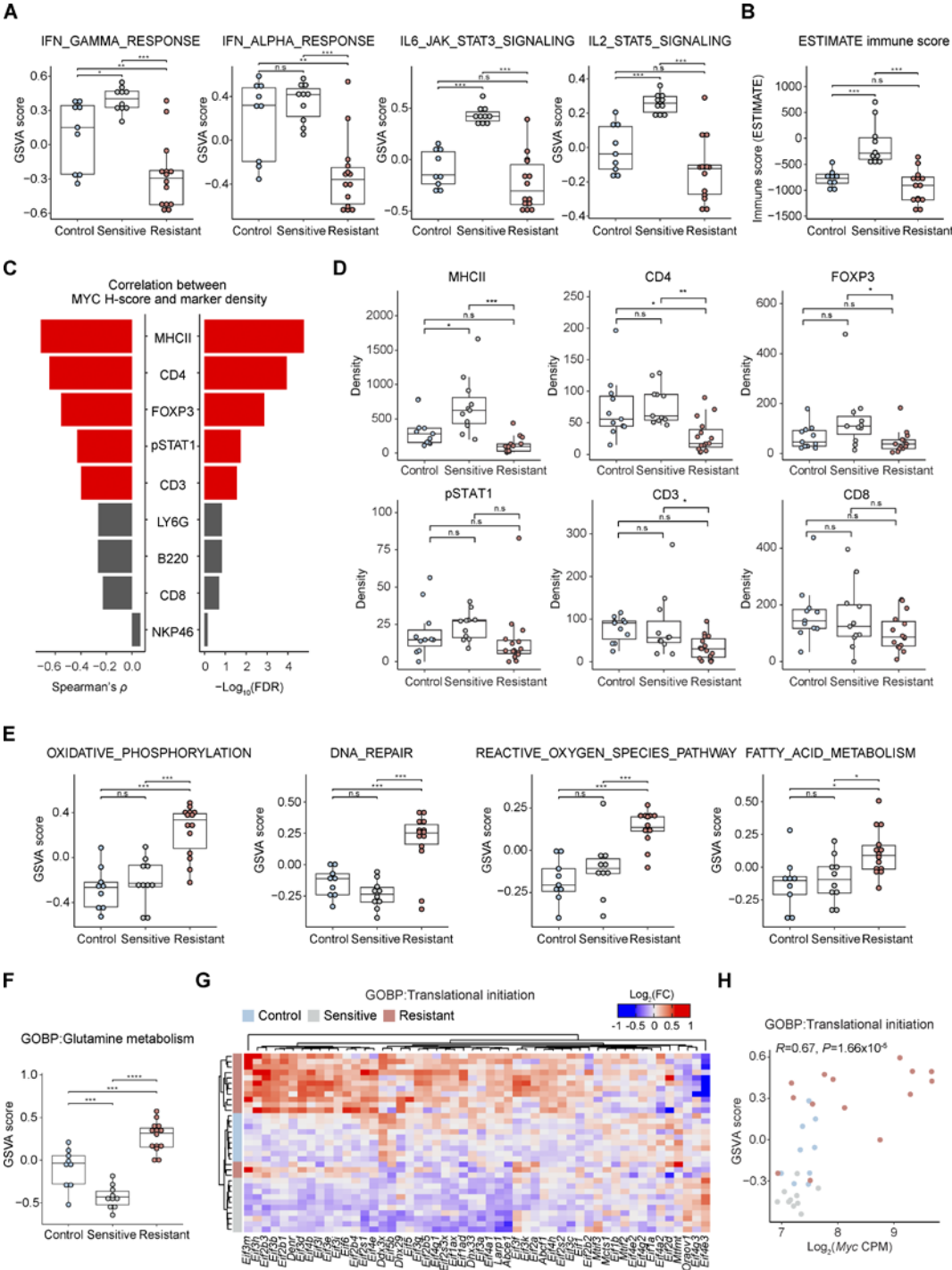


Figure 3. Biological processes associated with AZD8055 efficacy.

(A) GSVA scores for immune-related MSigDB Hallmark gene sets using KEP tumor RNA-seq profiles. (B) Immune infiltration scores based on RNA-seq data inferred by ESTIMATE [34]. (C) Correlations between MYC protein H-scores and immune cell densities based on quantified IHC for MYC and indicated immune markers. Spearman's ρ and $-\log_{10}(\text{FDR})$ are depicted as bar plots and red bars indicate significant correlations ($\text{FDR} < 0.05$). P values were calculated with two-tailed t -transformations of Spearman's correlation coefficient. (D) Immune cell densities based on quantified IHC for indicated immune markers. (E and F) GSVA scores for MSigDB Hallmark gene sets (E) and GO glutamine metabolism (F) upregulated in resistant tumors. (G) Heatmap displaying mRNA expression of translation initiation genes obtained from the GO Biological Processes (GOBP) translational initiation gene set. Hierarchical clustering (Euclidean distance and complete-linkage clustering) was performed on $\log_2$ -fold-change ($\log_2\text{FC}$) of CPM values centered on the median expression of control tumors. (H) Pearson's correlation analysis of *Myc* mRNA expression ($\log_2\text{CPM}$) vs. GOBP translation initiation gene set scores. P value was calculated with two-tailed t -transformations of Pearson's correlation coefficient. In A, B, and E–H, 9 control, 10 sensitive, and 14 resistant KEP tumors, and in C and D, 13 control, 11 sensitive, and 15 resistant KEP tumors were analyzed. In A, B, and D–F, data are represented as median $\pm$ IQR (box) and quartiles $\pm 1.5 \times$ IQR (whiskers), and one-way ANOVA and Tukey's post-hoc tests were performed to compute adjusted P values (* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; n.s., not significant).

MYC confers resistance to mTOR blockade in vitro

To assess whether MYC functioned as a cell-intrinsic mediator of mTORi resistance, we leveraged pharmacogenomic datasets derived from a compendium of human cancer cell lines [41–43]. We evaluated the correlation between *MYC* copy number status and mTORi response. Cancer cell lines harboring *MYC* amplification versus lines with normal *MYC* copy numbers displayed significantly elevated area under the dose–response curve (AUC) values for multiple mTORi (Fig. 4 A). An independent dataset derived from 30 human breast cancer cell lines [44] validated the negative association between *MYC* amplification and mTORi response (Fig. 4 B). Of note, MYC protein levels and AZD8055 response values also significantly correlated across human breast cancer cell lines (Fig. 4 C). We examined whether, depending on their MYC status, cells would also differentially respond to genetic disruption of the mTOR complex. Therefore, we made use of the Cancer Dependency Map [45] by focusing on the genes encoding the mTOR components MTOR, RPTOR, and RICTOR. This revealed that *MTOR*, and to a lesser extent *RPTOR* and *RICTOR*, are indispensable for the survival of the majority of cancer cell lines, regardless of *MYC* amplification status (Fig. S3 A). We functionally evaluated the essentiality of mTOR complex components in the mouse KEP tumor–derived KEP1.23 and KEP2.E3 cell lines [23] overexpressing either *GFP* or murine *Myc*. CRISPR/Cas9-based editing of *Mtor*, *Rptor*, or *Rictor* followed by Tracking of Indels by Decomposition (TIDE) analysis [46] revealed a continuous loss of frame-shift indels independent of *Myc* overexpression (Fig. S3 B). Thus, despite *Myc* amplification or experimental overexpression, mTOR components remained essential for cell survival, yet tolerance to mTORi appeared to be elevated.

To functionally test whether MYC reduced the sensitivity to mTORi, we subjected *Myc*-overexpressing KEP cell lines and the corresponding *GFP*-expressing controls to AZD8055, which inhibited mTOR signaling regardless of *Myc* overexpression (Fig. 5 A). Importantly, drug–response assays showed a three to fivefold increased AZD8055

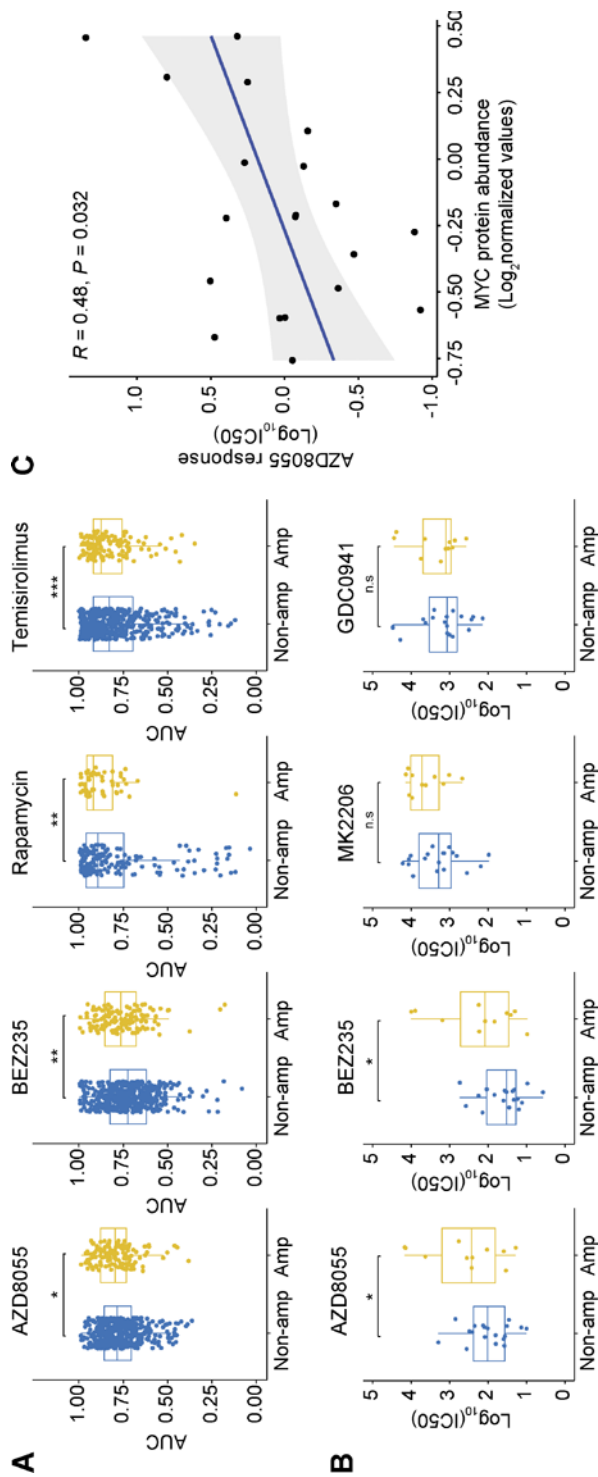


Figure 4. Association between MYC status and mTORi response in human cancer cell lines.

(A and B) Comparisons of mTORi responses of cancer cell lines with and without MYC amplifications. In A, AUC values and MYC amplification status were obtained from the cBioPortal database [41]. A total of 1,010 cell lines were analyzed (MYC amplification, $n = 209$; MYC non-amplified, $n = 801$). In B, IC50 values and MYC amplification status were obtained from a previously published breast cancer cell line dataset [44]. Data are represented as median $\pm$ IQR (box) and quartiles $\pm 1.5 \times$ IQR (whiskers) and P values were computed with one-tailed Student's t tests (* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; n.s, not significant; $P = 0.04, 0.04, 0.06$, and 0.25 for AZD8055, BEZ235, MK2206, and GDC0941, respectively). A total of 30 cell lines were analyzed (MYC amplification, $n = 11$; MYC non-amplified, $n = 19$). **(C)** Correlation between AZD8055 response and MYC protein abundance in 20 human breast cancer cell lines. P value was calculated with two-tailed t -transformation of Pearson's correlation coefficient. IC50 values were obtained from the GDSC database [43] and MYC protein levels, as measured by MS, were obtained from the CCLE database [42]. Source data are available for this figure: SourceData F4.

half-maximal inhibitory concentration (IC50) for KEP cells overexpressing *Myc* versus control cells (Fig. 5 B). Proliferation assays showed that mTOR inhibition efficiently hindered growth of control KEP cells, whereas *Myc*-overexpressing cells were less affected by AZD8055 (Fig. S3 C). Next, we subjected KEP cell lines to drug response assays using the clinically approved mTORi everolimus and the PI3K inhibitors (PI3Ki) alpelisib and buparlisib. Intriguingly, our results showed that MYC strongly induces resistance to everolimus, whereas the responses to alpelisib and buparlisib were only marginally affected by *Myc* overexpression (Fig. 5 C and Fig. S3 D). Finally, we overexpressed human *MYC* in the human breast cancer cell lines SUM52PE, MDA-MB-468, MCF7, and T47D (Fig. 5 D and Fig. S4 A). The mTORi AZD8055 and everolimus markedly suppressed mTOR signaling independently of MYC (Fig. 5 D). Consistent with the observations in KEP cell lines, colony formation and drug-response assays using SUM52PE and MDA-MB-468 demonstrated that *MYC* overexpression strongly reduced sensitivities to AZD8055 and everolimus (Fig. 5, E–H; and Fig. S4, B and C). In MCF7 and T47D cells, MYC also reduced mTORi sensitivity but to a lesser extent than in SUM52PE and MDA-MB-468 (Fig. S4, D–G). Noteworthy, in none of the four cell lines *MYC* overexpression affected sensitivity to the PI3Ki alpelisib and buparlisib (Fig. S4, H and I). Taken together, our results showed that MYC drives mTORi resistance in cell culture models of mouse and human breast cancer.

MYC confers resistance to mTOR blockade in vivo

Our in vitro findings prompted us to address whether MYC was also a tumor cell-intrinsic driver of mTORi resistance in vivo. Firstly, we generated a somatically engineered mouse model allowing us to induce MYC activity in situ in established ILC tumors. To this end, we somatically delivered lentiviruses containing *Myc^{ERT2}-P2A-Cre* to mammary glands of *Cdh1^{F/F};Trp53^{F/F}* (EP) female mice by using intraductal injections [47,48]. This led to the formation of EP-*Myc^{ERT2}* mammary tumors that express an MYC-ERT2 fusion protein, which translocates to the nucleus upon tamoxifen (TAM) binding. Histological examination confirmed emerging tumors to express MYC and to be solid ILCs, consistent with the predominant tumor type emerging from mammary glands of *K14-Cre;Cdh1^{F/F};Trp53^{F/F}* female mice [22–24]. We orthotopically transplanted EP-*Myc^{ERT2}* tumors into the mammary fat pads of immunocompetent female mice. When allografts became palpable, recipient mice were allocated to either normal or TAM-containing food pellets to activate nuclear translocation of MYC-ERT2, which is a well-established method to promote nuclear MYC activity in vivo [33]. Simultaneously, mice were subjected to daily vehicle versus AZD8055 treatments by oral gavage. AZD8055-mediated mTOR blockade efficiently delayed tumor growth resulting in prolonged tumor-specific survival (Fig. 6, A and B), in agreement with the observed tumoristasis in AZD8055-treated KEP tumors [23]. Strikingly, TAM-mediated MYC-ERT2 activation strongly accelerated the growth of EP-*Myc^{ERT2}* tumors treated with AZD8055, resulting in a shortened latency of tumor-specific survival comparable with vehicle-treated controls (Fig. 6, A and B). Second, we orthotopically transplanted human MDA-MB-468 breast cancer cells expressing either *Akaluc* or *MYC-P2A-Akaluc* into the mammary glands of immunocompromised NOD-*Prkdcscid*-*IL2rgTm1/Rj* (NXG) female

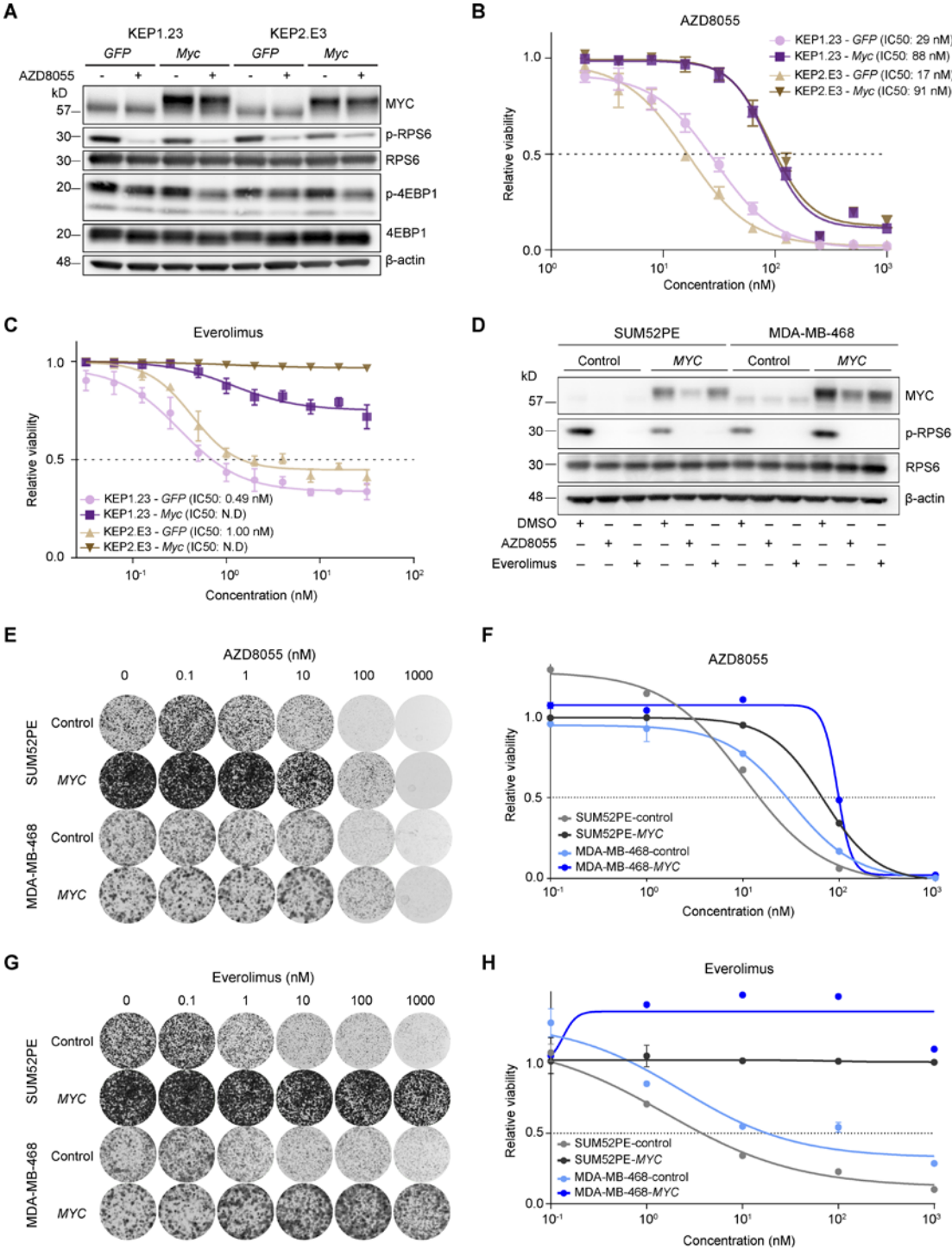


Figure 5. MYC drives AZD8055 resistance in vitro.

(A) Western blots for MYC and mTOR signaling markers on KEP1.23 and KEP2.E3 cell lines transduced with *GFP* or *Myc*-overexpressing lentiviral vectors and treated for 24 h with vehicle or 25 nM AZD8055. Data represent one experiment. **(B and C)** Dose–response curves of KEP cell lines expressing *GFP* or *Myc*, treated with AZD8055 (B) or everolimus (C) for 3 d and measured by SRB colorimetric assay. Data are represented as mean $\pm$ standard deviation of five replicas per group of one representative experiment of $n \geq 3$ independent experiments. N.D., not determined. **(D)** Western blots for MYC and (p-)RPS6 on serum-starved SUM52PE and MDA-MB-468 cell lines transduced with control or *MYC*-overexpressing lentiviral vectors and treated for 6 h with either vehicle DMSO, 200 nM AZD8055, or 10 nM everolimus. Data represent one experiment of two independent experiments. **(E–H)** Representative images (E and G) and corresponding quantification using CellTiter-Blue reagent (F and H) of long-term colony formation assays with SUM52PE and MDA-MB-468 cells transduced with control or *MYC*-overexpressing lentiviral vectors and treated with AZD8055 (E and F) or everolimus (G and H). Cells were treated with indicated drug doses for 10 d, and values were normalized to untreated conditions within each cell line. Data in F and H are represented as mean $\pm$ standard deviation of three technical replicas per group across one representative experiment of two independent experiments. Source data are available for this figure: SourceData F5.

mice via intraductal injection. Intraductal injection of patient-derived breast cancer cells represents a methodology that well-preserved breast cancer-specific pathology and tumor growth kinetics, and that comprehends superior tumor-take rates over fat pad or subcutaneous transplantations [49,50]. When the mammary glands displayed palpable tumors, mice were allocated to vehicle, AZD8055, or everolimus treatment arms. Daily administration of AZD8055 or everolimus via oral gavage suppressed tumor growth, as evidently shown by bioluminescence using the luciferin analog Akalumine (Fig. 6 C), and resulted in significantly prolonged tumor-specific survival of mTORi-treated mice as compared to animals treated with vehicle (Fig. 6 D). Importantly, *MYC* overexpression strongly reduced the efficacy of AZD8055 and everolimus, resulting in shortened mammary tumor-specific survival compared with the mTORi-treated cohort harboring MDA-MB-468 without *MYC* overexpression (Fig. 6, C and D). Collectively, our data demonstrated that MYC drives resistance to clinically approved mTOR-targeted therapies in mouse models of murine and human breast cancer.

MYC counteracts mTORi-mediated translation inhibition

Previous work showed that inhibition of mTOR suppresses protein translation [51–54]. Consistently, we found that mTOR blockade in KEP tumors suppressed transcriptional programs related to protein translation initiation (Fig. 3, G and H; and Fig. S2 E). In contrast, *Myc* copy number gains in AZD8055-resistant KEP tumors correlated with the upregulation of genes involved in protein translation, as revealed by gene set analysis of RNA-seq data (Fig. 3, G and H; and Fig. S2 E). We functionally addressed whether MYC directly counteracted mTORi-dependent ablation of protein translation. We first measured global translation in control and *Myc*-overexpressing KEP1.23 cells using a [³⁵S]-methionine incorporation assay. Both AZD8055 and everolimus treatment suppressed protein translation by 20–30%, whereas global translation in *Myc*-overexpressing cells was unimpaired by mTORi (Fig. 7 A). In line with this observation, evaluation of RNA-seq data derived from KEP1.23 cells showed that AZD8055 treatment resulted in the downregulation of genes related to translation and

ribosome processes, whereas *Myc* overexpression strongly induced these processes and considerably counteracted the effect of mTOR blockade (Fig. S5 A).

To elucidate the underlying mechanism of MYC-mediated translation regulation, we performed an integrative analysis of genome-wide ribosome footprints using ribosome profiling (Ribo-seq) and transcriptomes using RNA-seq. This allowed us to calculate translation efficiencies (TEs) of expressed genes and to identify genes with increased or decreased TEs across two experimental conditions, that is, the changes in ribosome footprints were higher or lower as compared with the changes in gene expression between the two conditions. We obtained gene sets that showed TE changes in KEP1.23 cells treated with AZD8055, overexpressing *Myc*, or both versus treatment-naïve control KEP cells (Fig. 7 B). Blockade of mTOR strongly suppressed TEs of genes related to the ribosome (Fig. 7 C). In contrast, overexpression of MYC-induced TEs of ribosomal genes, which significantly mitigated the effect of mTORi on ribosomal gene TEs (Fig. 7, C and D). Most ribosomal gene transcripts showed a stark decrease in ribosome occupancy in AZD8055-treated versus control cells, and *Myc* overexpression counteracted this phenotype by promoting the translation of ribosomal genes (Fig. 7 E). TEs of some ribosomal genes that are highly suppressed by AZD8055 were not restored by MYC overexpression. In line with this observation, we found a significant correlation in TE changes induced by AZD8055 between control KEP cells and *Myc*-overexpressing KEP cells (Pearson's $R = 0.7$, $P = 7.7 \times 10^{-12}$). Of note, MYC did not affect TEs of genes encoding translation machinery factors other than ribosomal proteins (Fig. 7 F). This suggested that MYC overcomes mTORi-mediated translation inhibition by recovering the translation of ribosomal gene transcripts.

To further dissect the connection of MYC to translational regulation, we performed motif enrichment analysis for RNA-binding proteins among the genes that showed increased TEs during *Myc* overexpression (Fig. 7 E). The 5'-UTR regions of these gene transcripts were enriched for motifs of known RNA-binding proteins including LARP1, SRSF1, and ZCRH14 (Fig. S5 B). Interestingly, *Larp1*, *Srsf1*, and *Zc3h14* transcripts were significantly upregulated in KEP cells overexpressing *Myc* as well as in AZD8055-resistant tumors containing focal *Myc* amplifications (Fig. S5, C and D), and the three genes were potential MYC transcriptional targets according to the ChIP Atlas database ([55]; Fig. S5 E). Taken together, these data implied that MYC can exert mTORi resistance through the transcriptional induction of RNA-binding proteins, which in turn fosters the translation of ribosomal proteins to restore protein synthesis during mTOR blockade.

AZD8055-resistant tumors depend on MYC activity

Next, we functionally dissected whether the acquired resistance to AZD8055 in KEP tumors depended on the observed *Myc* amplifications. To this end, we established primary cell cultures from KEP tumors that had simultaneously been used for the multiomics analyses. We grew cells from three vehicle-treated control tumors and from two AZD8055-resistant

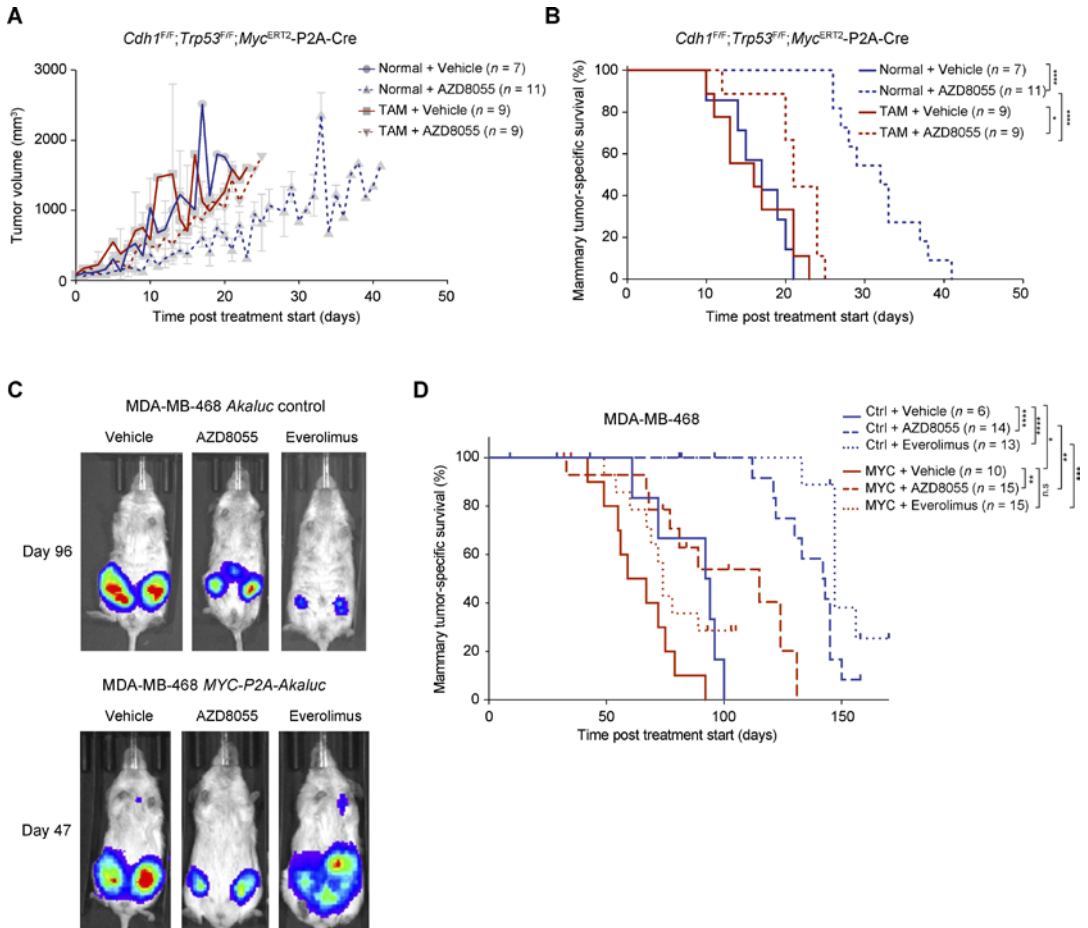


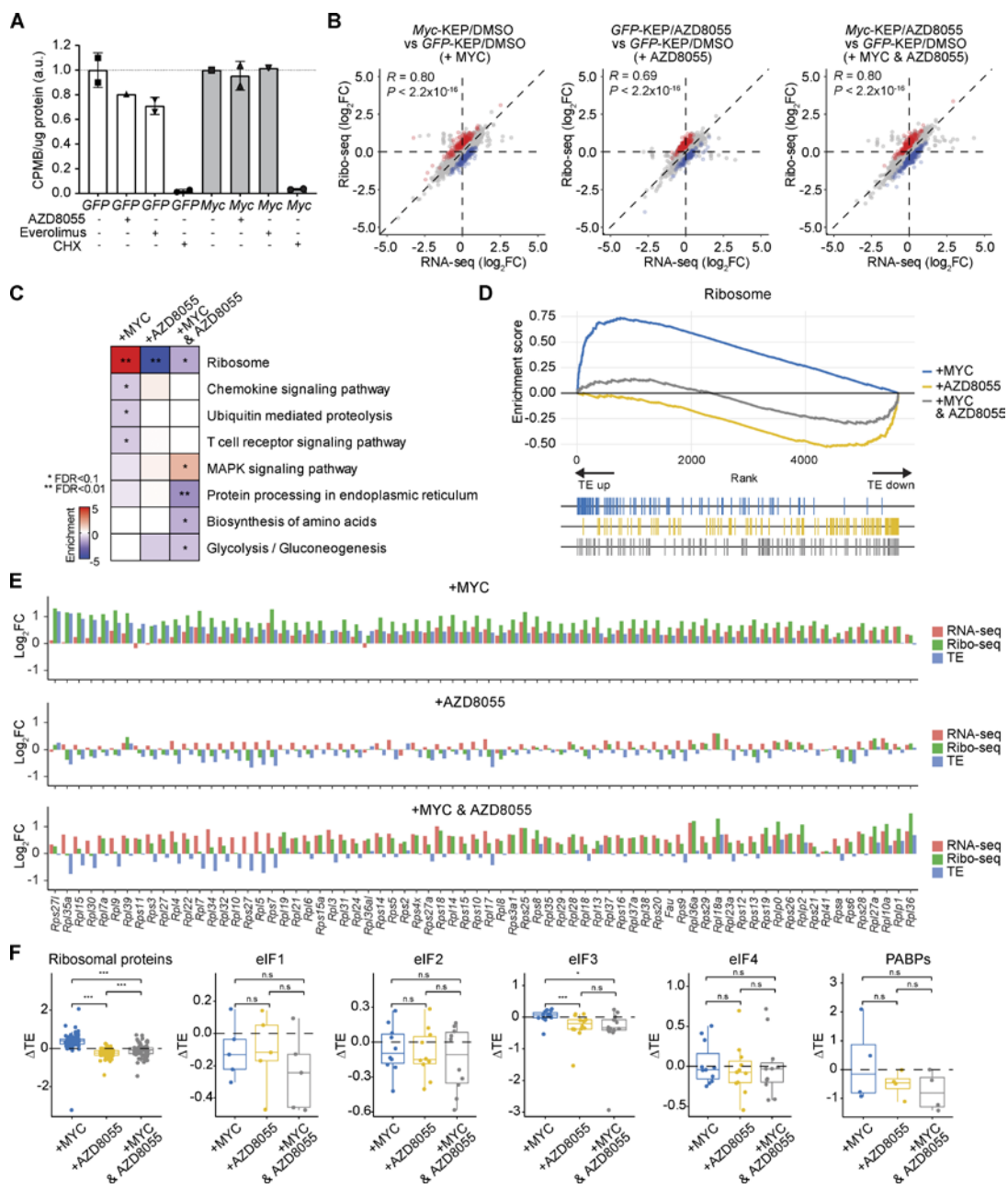
Figure 6. MYC drives AZD8055 resistance in vivo.

(A and B) Mammary tumor growth curves (A) and Kaplan-Meier curves showing tumor-specific survival (B) of female syngeneic mice bearing mammary fat pad transplants derived from *Cdh1^{FF};Trp53^{FF};MyceRT2-P2A-Cre* tumor donors. Mice were subjected ad libitum to a normal food pellet versus TAM-containing food pellet diet (TAM-induced nuclear translocation of MYC-ERT2) and treated daily orally with vehicle or 20 mg/kg AZD8055. For mammary tumor growth curves in A, data are represented as mean $\pm$ SEM of indicated numbers of replica per group. **(C)** Representative *in vivo* bioluminescence images of immunocompromised female NXG mice intraductally injected with MDA-MB-468 cells transduced with *Akaluc* or MYC-P2A-*Akaluc* lentiviral vectors and treated daily orally with 20 mg/kg AZD8055, 5 mg/kg everolimus, or either of the corresponding vehicles. Akaluciferase activity is shown following Akalumine-HCl administration 96 (*Akaluc*) or 47 (MYC-P2A-*Akaluc*) d after mTORi treatment start. **(D)** Kaplan-Meier curves showing mammary tumor-specific survival of female immunocompromised mice intraductally injected with MDA-MB-468 cells expressing *Akaluc* or MYC-P2A-*Akaluc* and treated daily with vehicles or indicated mTORi. In B and D, P values were calculated with log rank (Mantel-Cox) test (*P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001; n.s., not significant). n represents the number of mice per cohort included in the analysis. Mice treated with AZD8055-vehicle or everolimus-vehicle were grouped into single vehicle cohorts.

tumors both harboring focal *Myc* amplifications. Cells from control KEP tumors were sensitive to AZD8055 (Fig. 8, A and B). Cells from AZD8055-resistant KEP tumors displayed target inhibition (Fig. 8 A) and maintained resistance to AZD8055 in vitro, as evidenced by an average sixfold IC₅₀ increase in AZD8055-resistant cells versus sensitive control cells (Fig. 8 B) as well as the maintained proliferation of resistant cells during treatment with AZD8055 (Fig. S5 F). We subjected the AZD8055-sensitive and -resistant cell cultures to drug–response assays using the MYC inhibitor (MYCi) KJ-Pyr-9 [56]. The cell cultures derived from AZD8055-resistant and *Myc*-amplified KEP tumors were more vulnerable to KJ-Pyr-9 as compared with the control cultures (Fig. S5 G). We next assessed whether MYC blockade would resensitize the mTORi-resistant and *Myc*-amplified cells to AZD8055 inhibition. Dose-responses to AZD8055 were measured in the absence or presence of 8 μ M KJ-Pyr-9, a dose that roughly corresponded to the KJ-Pyr-9 IC₄₀ observed for AZD8055-resistant cell cultures (Fig. S5 G). MYC inhibition reduced baseline cell viability of both control and AZD8055-resistant cells. Using Bliss independence modeling, which assumes independent actions of two combined drugs [57], we predicted the additive effect when combining AZD8055 and KJ-Pyr-9 (“Bliss independence” in Fig. 8 C and Fig. S5 H). Comparison of the observed effect of combination treatment with the Bliss independence prediction showed that MYC blockade had no effect on AZD8055 sensitivity of control KEP cells (5 nM of average Δ IC₅₀ between Bliss model prediction and observed response; Fig. 8 C and Fig. S5 H). In contrast, KJ-Pyr-9 treatment synergistically resensitized mTORi-resistant KEP cells to AZD8055 inhibition (average Δ IC₅₀ of 127 nM; Fig. 8 C and Fig. S5 H). Consequently, upon MYC blockade, the AZD8055 IC₅₀ curves of *Myc*-amplified resistant KEP cultures decreased to a sensitivity level comparable with control KEP cultures (Fig. 8, B and D). Together,

Figure 7. MYC enhances the translation efficiency of ribosomal genes.

(A) [³⁵S]-methionine incorporation assay performed with KEP1.23 cells overexpressing *GFP* or *Myc* and treated for 2 h with vehicle, 50 nM AZD8055, or 10 nM everolimus prior to a 1 h pulse with [³⁵S]-labeled methionine. 100 μ g/ml cycloheximide (CHX) was added as a positive control to stop protein translation. Incorporation of [³⁵S]-methionine was quantified by scintillation counting and normalized to total protein. Data are represented as mean $\pm$ standard deviation of $n = 2$ technical replica per group (*GFP*-DMSO, *GFP*-everolimus, *GFP*-CHX, *Myc*-AZD8055, and *Myc*-CHX) or $n = 1$ replica (*GFP*-AZD8055, *Myc*-DMSO, and *Myc*-everolimus) of one experiment and normalized to their corresponding DMSO treated condition. **(B)** Correlations of log₂(FC) values between Ribo-seq reads and RNA-seq reads in indicated comparisons. Blue and red dots indicate genes showing enhanced or suppressed TEs, respectively. **(C)** Functional enrichment analysis using Kyoto Encyclopedia of Genes and Genomes pathways for genes with significantly altered TEs upon *Myc* overexpression, AZD8055 treatment, or the both, each compared to *GFP*-DMSO. Fisher’s exact test was performed followed by FDR correction. *FDR < 0.05, **FDR < 0.01. **(D)** Gene set enrichment plot showing changes in TEs upon *Myc* overexpression, AZD8055 treatment, or the both, each compared to *GFP*-DMSO. **(E)** Log₂FC values for all quantified ribosomal genes computed from RNA-seq reads (red), Ribo-seq reads (green), and TE (blue) for each comparison. **(F)** Log₂FC values of TE (Δ TE) in each condition when compared to *GFP*-DMSO for different protein synthesis factors. Data are represented as median $\pm$ IQR (box) and quartiles $\pm 1.5 \times$ IQR (whiskers), and one-way ANOVA and Tukey’s post-hoc test were performed to compute adjusted P values (*P < 0.05; ***P < 0.001; n.s., not significant).



these data showed that MYC can confer resistance to AZD8055, and MYC blockade might represent a therapeutic strategy to overcome acquired resistance to mTORi.

MYC status is associated with poor response to mTORi-based therapy

To examine the clinical relevance of our findings, we evaluated the association between *MYC* amplification status and mTORi response in tumor samples derived from patients with cancer. We first analyzed WGS data with associated clinical data from the

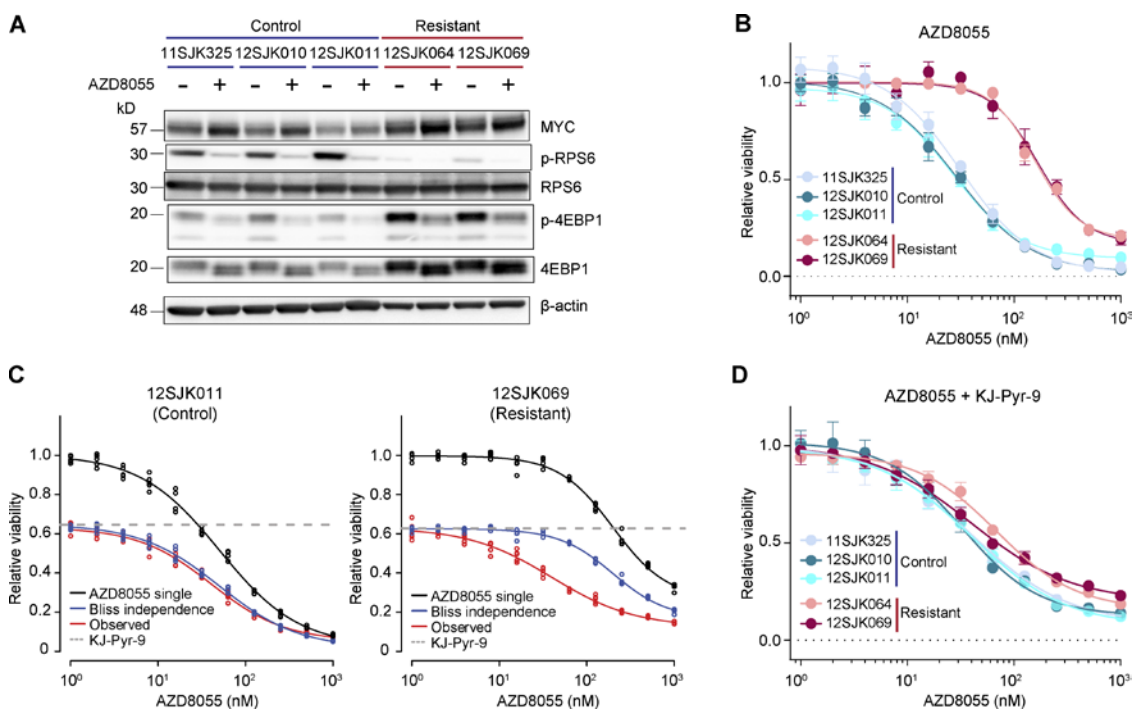


Figure 8. Synergistic effect of mTOR and MYC inhibition in AZD8055-resistant cell cultures.

(A) Western blots for MYC and mTOR signaling markers on control and AZD8055-resistant, *Myc*-amplified KEP tumor-derived primary cell cultures treated with vehicle, or 90 nM AZD8055 for 24 h. Data represent one experiment of three independent experiments. **(B)** Dose-response curves of control and resistant KEP cultures, treated with AZD8055 for 3 d, and measured by SRB colorimetric assay. Data are represented as mean $\pm$ standard deviation of five replicas per group of five independent experiments. **(C)** Dose-response curves of control (12SJK011) and resistant KEP (12SJK069) cultures, treated with an AZD8055 range and 8 μ M KJ-Pyr-9 MYC for 3 d, and measured by SRB colorimetric assay. Bliss independence model is shown to display the independent effect of KJ-Pyr-9 and AZD8055. The dotted line represents the basal effect of 8 μ M KJ-Pyr-9 and the solid lines represents dose response curves for AZD8055 single treatment (black), Bliss independence (blue), and observed effect of combination treatment (red). Data are represented as mean $\pm$ standard deviation of five replicas per group of one experiment. **(D)** Dose-response curves of control and resistant KEP cultures, treated with 8 μ M of MYC KJ-Pyr-9 in addition to a range of AZD8055 concentrations for 3 d, and measured by SRB colorimetric assay. Data are represented as mean $\pm$ standard deviation of five replicas per group of one experiment. Source data are available for this figure: SourceData F8.

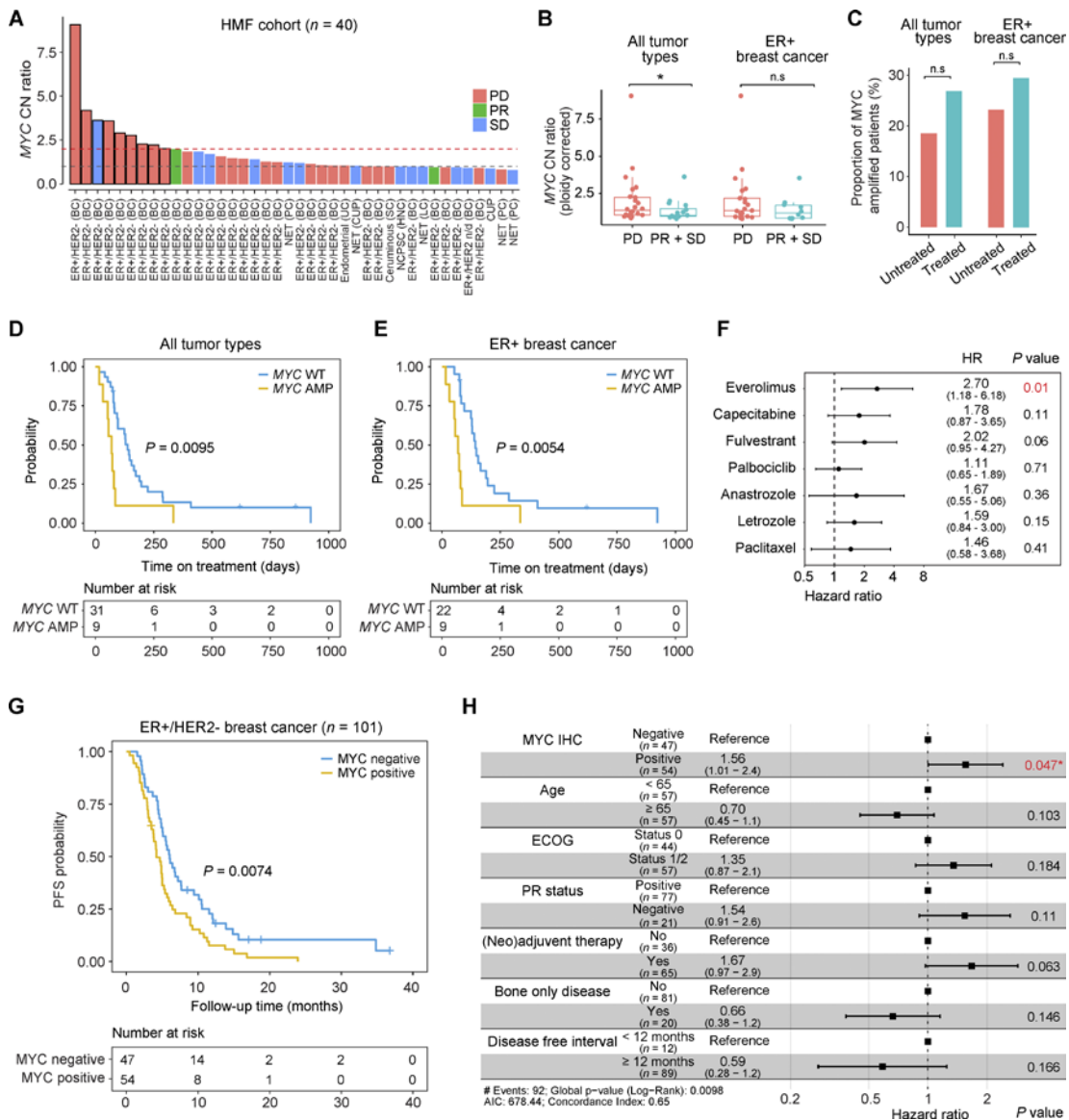
metastatic cancer cohort of the Hartwig Medical Foundation (HMF; [58]). This collection contained data for 40 cancer patients who received the mTORi everolimus as a single agent or in combination with other drugs. Analysis of genomic profiles from pretreatment biopsies revealed higher *MYC* copy number levels in non-responders (progressive disease) as compared with responders (SD or PR, as defined by response evaluation criteria in solid tumors [RECIST] criteria; Fig. 9, A and B). When focusing on ER+ breast cancer only (30 ER+/HER2– [human epidermal growth factor receptor 2], 1 ER+/HER2 unknown), we also observed a trend toward *MYC* amplification enrichment in non-responding patients (Fig. 9 B). We also observed a trend toward a higher prevalence of *MYC* amplification in the post-mTORi-treatment cohort versus the pretreatment cohort, although the difference did not reach statistical significance (Fig. 9 C). This suggested that *MYC* amplification-driven resistance to mTOR-targeted therapies may also occur in human cancer patients. Because of incomplete survival information for patients receiving everolimus, we used treatment duration as a proxy for progression-free survival [59,60]. Patients with *MYC*-amplified tumors received everolimus for a significantly shorter period of time, indicating that everolimus was less effective in these patients (Fig. 9, D and E). Of note, univariate Cox-regression analysis of ER+/HER2– breast cancer samples from the HMF data showed that *MYC* amplification is significantly associated with everolimus-based treatment duration (HR = 2.70, $P = 0.01$; Fig. 9 F).

Finally, we analyzed 101 ER+/HER2– metastatic breast cancer patients who participated in the everolimus biomarker study (EudraCT number 2013-004120-11) and received everolimus plus exemestane as a standard of care treatment [61]. Primary tumor tissues were stained for *MYC* by IHC, and the percentage of stained tumor cells was quantified to group patients into *MYC*-positive and *MYC*-negative samples. The *MYC*-positive group displayed significantly shortened progression-free survival as compared with the *MYC*-negative group (Fig. 9 G). Importantly, a multivariate Cox proportional hazard model including clinicopathological characteristics confirmed that, among the variants tested, *MYC* positivity was the only determinant significantly associated with worse progression-free survival (HR = 1.56, $P = 0.047$; Fig. 9 H).

Together, these results suggested that *MYC* is a clinically significant driver of mTORi resistance and that *MYC* status may potentially stratify patients that may benefit from mTOR-targeted therapy.

DISCUSSION

The PI3K–AKT–mTOR signaling pathway is frequently activated in breast cancer and has therefore been considered an attractive drug target [3]. However, lack of knowledge about the factors determining response to drugs targeting PI3K–AKT–mTOR signaling including



mTORi often hampers the optimal choice of treatment [6,7]. In this study, we demonstrated MYC activation to be a central mechanism of mTORi resistance in breast cancer. Resistance to mTORi has previously been investigated in several cancer types, and MYC has been reported as a resistance mechanism [62–70]. These studies were mostly based on in vitro cell culture experiments [62–64,69]. To recapitulate mTORi resistance acquisition occurring in breast cancer patients in a more appropriate way, we analyzed mammary tumors isolated

Figure 9. MYC status is a poor prognostic factor for everolimus-based therapy.

(A) MYC CN status of tumors from the HMF cohort ($n = 40$) prior to mTORi treatment. MYC CN status was normalized to sample ploidy (SP), and samples with MYC CN $> 2 \times SP$ were defined as MYC-amplified tumors (Priestley et al., 2019). Bar plots show MYC CN status of samples with clinical response to mTORi treatment. PD, progressive disease. Red dotted line denotes MYC amplification cutoff. The gray dotted line denotes normal MYC CN status (CN ratio = 1). BC, breast cancer; PC, pancreatic cancer; UC, uterine cancer; SC, skin cancer; HNC, head and neck cancer; LC, lung cancer; NET, neuroendocrine tumor; NCPSC, nasal cavity and paranasal sinus cancer; CUP, cancer of unknown primary; n/d, not determined. **(B)** Comparisons of MYC CN status of responders (PR + SD) vs. non-responders (PD) in all tumor types (total, $n = 40$; PR + SD, $n = 16$; PD, $n = 24$) and in ER+ breast cancer (total, $n = 31$; PR + SD, $n = 9$; PD, $n = 22$), specifically. Data are represented as median $\pm$ IQR (box) and quartiles $\pm 1.5 \times$ IQR (whiskers) and Wilcoxon rank-sum test was performed to compute P values (* $P < 0.05$; n.s, not significant; all tumor types, $P = 0.04$; ER+ breast cancer, $P = 0.13$). **(C)** Percentages of tumors with MYC amplifications in the pre and post mTORi-treated HMF cohorts (all tumor types, $n = 121$; untreated, $n = 54$; treated, $n = 67$; ER+ breast cancer, $n = 104$; untreated, $n = 43$; treated, $n = 61$). Proportion test was performed to compare the percentages between groups (n.s, not significant; all tumor types, $P = 0.19$; ER+ breast cancer, 0.31). **(D and E)** Kaplan–Meier curves showing everolimus treatment duration for patients with MYC WT or MYC-amplified cancers (D; $n = 40$) or specifically ER+ breast cancers (E; $n = 31$) using data from the HMF cohort. P values were calculated with log-rank (Mantel–Cox) test. **(F)** Univariate cox regression analysis to evaluate the association between individual therapeutics used in ER+/HER2– breast cancer and treatment duration from HMF cohort. Data are represented as estimated hazard ratio (HR) $\pm$ 95% confidence interval (CI). **(G)** Kaplan–Meier curve showing progression-free survival (PFS) of ER+/HER2– breast cancer patients ($n = 101$) from the everolimus biomarker study (EudraCT number 2013-004120-11) during everolimus/exemestane treatment stratified for MYC positivity using IHC. P values were calculated with log rank (Mantel–Cox) test. **(H)** Multivariate Cox regression analysis on breast cancer specimens from G using indicated clinicopathological variables. For each variable, data are represented as estimated hazard ratio $\pm$ 95% CI. ECOG, Eastern Cooperative Oncology Group performance status; PR, progesterone receptor status of the primary tumor of study participants.

from immunocompetent mice that were enrolled into a long-term mTORi intervention, we made use of clinically approved mTOR-targeted therapies, and we engaged in mouse modeling using tumor samples of mouse and human origin to functionally pinpoint MYC as an in vivo driver of mTORi resistance. Notably, we also made use of clinical data, which highlighted the clinical potential of MYC to predict mTORi response and to stratify patients that are less likely to benefit from mTORi therapy.

Although MYC was the most prominent hallmark of AZD8055-resistant tumors, copy number analysis also revealed a focal amplification of Chr-2 in three resistant tumors (Fig. 1 D) that was mutually exclusive to the *Myc* amplification. The Chr-2 region covered 35 protein-coding genes, two of which encoded the translation initiation factors EIF3M and EIF5 (Fig. S1 G). We demonstrated that MYC drives mTORi resistance by restoring protein synthesis during mTOR blockade. Conceivably, genes encoded by the Chr-2 amplicon including *Eif3m* and *Eif5* that foster protein translation might thus represent alternative resistance factors to overcome mTOR blockade beyond MYC.

We identified MYC as an acquired resistance factor during mTOR inhibition. Correspondingly, our analyses of human cancer cell line and patient data demonstrated that MYC amplification and high MYC expression prior to mTORi treatment are associated with poor mTORi response. This suggests that MYC-driven tumors are likely intrinsically resistant to mTOR blockade. Across human cancer, MYC is frequently amplified and MYC can mitigate antitumor effects of many cancer therapeutics [71]. This stresses the need

to identify the Achilles' heel(s) of MYC-driven cancers. Targeting MYC itself has remained challenging, partly because of the lack of an enzymatic pocket and the predominant nuclear localization of the protein, both of which make the use of small molecule inhibitors or blocking antibodies cumbersome [72]. However, recent cancer dependency map studies utilizing CRISPR/Cas9 and small hairpin RNA-based genetic screens may represent valuable resources to identify novel therapeutic opportunities to target MYC-dependent cancer cells [45,73].

At present, targeted therapies of the PI3K–AKT–mTOR pathway are not first-line in breast cancer. Instead, these therapeutics have been used in ER+/HER2– postmenopausal breast cancer patients who are refractory to endocrine therapies—acquired genetic alterations activating the PI3K–AKT–mTOR pathway frequently drive resistance to endocrine therapies [74]. Our study suggests that MYC drives primary resistance to mTORi and might thus serve as a predictive biomarker for response to second-line PI3K–AKT–mTOR-targeted therapies. Our analysis of the HMF cohort indeed revealed that 32% (10/31) of ER+/HER2– breast cancer patients showed an objective response to the everolimus–exemestane combination treatment regime (Fig. 9 A). It will be interesting to retrospectively evaluate MYC status using clinicogenomic data from the BOLERO-2 trial or other clinical trials that evaluated everolimus-based therapies for ER+/HER2– breast cancer in different first-line settings [75,76]. Considering the potential adversary role of MYC activity in the efficacy of a diverse set of cancer therapeutics, MYC status should also be evaluated in the context of other anticancer therapies. To conclude, our findings suggest that MYC is an important mediator of mTORi response and its status should be considered to select patients who may benefit from mTOR-targeted therapies.

MATERIALS AND METHODS

Lentiviral vectors and virus production

The SIN.LV.SF-T2A-Puro, SIN.LV.SF-GFP-T2A-Puro, and SIN.LV.SF-P2A-Cre lentiviral vectors were previously described [48]. Mouse *Eif4b* (NM_145625.3), mouse *Melk* (NM_010790.2), and human *MYC* (NM_001354870.1) cDNAs were isolated from custom-synthesized gBlocks Gene Fragments (Integrated DNA Technologies). The mouse *Myc* (NM_010849.4) and *Myc*^{ERT2} cDNAs were isolated from the *Myc* cDNA clone 8861953 (Source BioScience) and the *Frt-invCAG-Myc*^{ERT2}–*IRES-Luc* vector [33], respectively. The *Akaluc* cDNA was isolated from a previously described vector [77] by using BamHI and AgeI restriction enzymes (#R0136, #R3552; New England BioLabs). cDNAs were amplified using Q5 High-Fidelity DNA Polymerase (#M0491; NEB) and inserted into the SIN.LV.SF-T2A-Puro and SIN.LV.SF-P2A-Cre vectors, resulting in SIN.LV.SF-Eif4b-T2A-Puro, SIN.LV.SF-Melk-T2A-Puro, SIN.LV.SF-Myc-T2A-Puro, and SIN.LV.SF- *Myc*^{ERT2}–P2A-Cre. Human *MYC* together with the P2A linker sequence and *Akaluc* were amplified and assembled into the SIN.LV.SF-T2A-Puro

backbone by using the In-Fusion HD Cloning Plus kit (#638911; Takara Bio) according to the manufacturer's recommendations, resulting in SIN.LV.SF-Akaluc-T2A-Puro and SIN.LV.SF-MYC-P2A-Akaluc-T2A-Puro vectors. Point mutations in *Eif4b* and *Melk* were introduced using the QuikChange Lightning Site-Directed Mutagenesis Kit (#210519; Agilent) and primers designed with QuikChange Primer Design (Agilent). All constructs were verified by Sanger sequencing. Lentiviral stocks were produced by transient co-transfection of four plasmids in HEK 293T cells as previously described [78]. Viral titers were determined using the qPCR Lentivirus Titer Kit (#LV900; Applied Biological Materials).

Cell culturing and lentiviral transduction

HEK 293T cells (#CRL-3216; ATCC) were cultured in Iscove's Modified Dulbecco's Medium (#31980030; Thermo Fisher Scientific) containing 10% fetal bovine serum (FBS; #S-FBS-EU-015; Serana) and 100 IU/ml penicillin and streptomycin (Pen Strep, #15070; Thermo Fisher Scientific). MDA-MB-468 (#HTB-132; ATCC), MCF7 (#HTB-22; ATCC), T47D (#HTB-133; ATCC), and SUM52PE (#HUMANSUM-0003018; BioIVT) cells were cultured in Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F-12, #10565018; Thermo Fisher Scientific) containing 10% FBS, 100 IU/ml Pen Strep, and 10 μ M Y-27632 dihydrochloride (#M1817; AbMole). Rac-11P cells [79] were cultured in DMEM medium (#31966-021; Thermo Fisher Scientific) supplemented with 10% FBS and 100 IU/ml Pen Strep. No reauthentication of the abovementioned cell lines was carried out for this study beyond the reauthentications done by the providers. Mouse tumor-derived primary cells as well as the two previously established KEP1.23 and KEP2.E3 cell lines [21] were cultured in DMEM/F-12 containing 10% FBS, 100 IU/ml Pen Strep, 5 ng/ml epidermal growth factor (#E4127; Sigma-Aldrich), 5 ng/ml insulin (#10516; Sigma-Aldrich), and 10 μ M Y-27632 dihydrochloride. All cell lines were cultured in standard incubators at 37°C with 5% CO₂ and routinely tested for mycoplasma contamination using the MycoAlert Mycoplasma Detection Kit (#LT07-218; Lonza).

Mouse KEP1.23 and KEP2.E3 cell lines were transduced with lentiviral supernatants at equal multiplicity of infections (MOIs). MDA-MB-468, SUM52PE, MCF7, and T47D human breast cancer cell lines were transduced with lentiviral supernatants using a 1–50 MOI range, and resulting experimental control and MYC-expressing cell line pairs were selected based on equal RNA expression levels of the corresponding transgenes, as determined by reverse transcription quantitative PCR. Cell transductions were performed in the presence of 8 μ g/ml Polybrene (#H9268; Sigma-Aldrich) for 24 h. Transduced cells were selected with 2 μ g/ml puromycin (#A11138; Thermo Fisher Scientific) for at least 5 d, and subsequently cultured with 1 μ g/ml puromycin to ensure continued transgene expression.

Isolation of primary tumor cells

50–100 mm³ of cryopreserved mammary tumor pieces were thawed on ice and minced using scalpel blades. Minced tumor samples were washed twice with DPBS (#14190144; Thermo Fisher Scientific), digested with 2 mg/ml collagenase type IV (#17104019; Thermo

Fisher Scientific), and 4 $\mu\text{g/ml}$ DNase I (#DN25, Sigma-Aldrich) in DMEM/F-12 for 1 h at 37°C, and the cell suspension was passed through a 70- μm cell strainer. Initially, cells were cultured in complete medium supplemented with 5 μM Nutlin-3a (#SML0580; Sigma-Aldrich) to select *Trp53*-deficient cells, which was verified by genotyping PCR as previously described [24]. Cells derived from AZD8055-resistant KEP tumors were continuously cultured with 90 nM AZD8055 to maintain the resistance phenotype.

Cell proliferation assay

400 KEP1.23 or KEP2.E3 cells or 500 primary control or resistant tumor cells were seeded per well in 96-well flat-bottomed plates. After 24 h, cells were subjected to 25 nM AZD8055 (KEP1.23 and KEP2.E3) or 90 nM AZD8055 (primary control and resistant tumor cultures) treatment. Real-time proliferation of the cells was monitored by the Incucyte ZOOM Live-Cell Analysis System (Essen BioScience) over 7 d, and the IncuCyte™ ZOOM 2015A Control software was used to visualize and analyze the results.

Drug-response assays

400 KEP1.23 or KEP2.E3 cells expressing *GFP* or *Myc*; 1,000 KEP2.E3 cells expressing *GFP*, *Eif4b*, or *Melk*; 700 primary control or resistant tumor cells; 4,000 MDA-MB-468 cells; or 5,000 SUM52PE, MCF7, or T47D cells were seeded per well in 96-well flat-bottomed plates. After 24 h, cells were subjected to either the mTORi AZD8055 (0.1 nM–1 μM range; AstraZeneca) or everolimus (#HY-10218; MedChem Express; 0.03–32 nM range), the PI3K α inhibitor alpelisib (#HY-15244; MedChem Express; 80 nM–80 μM range), the pan-class I PI3Ki buparlisib (#HY-70063; MedChem Express; 10 nM–10 μM range), or the MYCi KJ-Pyr-9 (#HY-19735; MedChemExpress; 800 nM–80 μM range), each for 3 d. To measure cell viability upon combined mTOR and MYC inhibition, cells subjected to AZD8055 were additionally treated with 8 μM KJ-Pyr-9. Cell viability was assayed using Sulforhodamine B (SRB) staining. Cells were fixed with 50% ice-cold trichloroacetic acid in H_2O for 2 h at 4°C, washed five times with tap water, and air-dried. Cells were then stained for 30 min with 0.4% SRB–1% acetic acid in H_2O at room temperature and washed five times with 1% acetic acid– H_2O . Bound SRB was dissolved with 10 mM Tris– H_2O and optical density was measured at 540 nm with an Infinite 200 PRO plate reader (Tecan). For KEP2.E3 cells expressing *Eif4b* or *Melk*, cell viability was assayed using CellTiter-Blue reagent (#G808A; Promega) for 4.5 h and an Infinite M Plex plate reader (Tecan). Long-term colony formation assays were performed in 6-well plates precoated with laminin by using Rac-11P cells as previously described [80]. Cells were seeded at 5,000 cells per well and treated the following day at the indicated concentrations. After 10 d of treatment, cell viability was quantified by incubating the cells with Cell-Titer Blue reagent. Cells were then fixed with 4% formaldehyde and stained with 0.1% (wt/vol) crystal violet. Plates were imaged using GelCount (Oxford Optronix). Drug-response curves were modeled using [Inhibitor] versus response with variable slope (four parameters) and least squares regression in Prism (version 8, GraphPad Software). The Bliss independence model was applied to evaluate synergy between AZD8055 and KJ-Pyr-9 [81].

CRISPR/Cas9-mediated genome editing

For CRISPR/Cas9-mediated genome editing of *Mtor* (NM_020009.2), *Rictor* (NM_030168.3), and *Rptor* (NM_028898.2), two independent single-guide (sg)RNAs per gene were selected from the Brie Mouse CRISPR Knockout Pooled Library [82]. Sequences of non-targeting and *Mtor*-, *Rptor*-, and *Rictor*-targeting sgRNAs were as follows: non-targeting, 5'-TGATTGGGG GTCGTTCGCCA-3'; *Mtor*-sg1, 5'-TGCAGTCTGGCTAACCACGT-3'; *Mtor*-sg2, 5'-GCTGATGCACGT GAATACGG-3'; *Rptor*-sg1, 5'-GATGTGTCAAGGATTCGTTG-3'; *Rptor*-sg2, 5'-TGCAGGTCGTAT ATGGACAG-3'; *Rictor*-sg1, 5'-TGACATTCAGCAGAGCAACG-3'; *Rictor*-sg2, 5'-TACCTGGATCTA GCACGATG-3'. sgRNAs were cloned into the lentiCRISPRv2-Cas9-Blast backbone (#98293; Addgene) containing *Cas9* and a blasticidin resistance cassette. DNA oligos were annealed with T4 Polynucleotide Kinase (#M0201; New England BioLabs). The lentiCRISPRv2-Blast backbone was digested with the BsmBI-v2 restriction enzyme (#R0739; New England BioLabs) for 2 h at 55°C, and the correct band was isolated from a 1% agarose gel using the Isolate II PCR and Gel Kit (#BIO-52060; Bioline). Oligos were ligated into the linearized and digested backbone using the Quick Ligation Kit (#M2200; New England BioLabs) for 10 min at 23°C and transformed and expanded in *Stb13* bacteria. All constructs were verified by Sanger sequencing.

20,000 KEP1.23 GFP or MYC-overexpressing cells were seeded per well into 6-well plates 24 h prior to transfection with lentiCRISPRv2-Cas9-Blast plasmids containing control or sgRNAs of interest using Xfect (#631317; Takara) according to the manufacturer's instructions. 24 h later, cells were trypsinized and replated with fresh media containing 10 µM blasticidin. After 2 d of blasticidin selection, a cell pellet from the first time point (t1) was collected and frozen, with a split of the cells continuing in culture without blasticidin. Cell pellets were subsequently collected at days 4 (t2) and 6 (t3) of initial blasticidin selection start, respectively. Cell pellets were lysed with DirectPCR Lysis Reagent (#302-C; Viagen), PCR-amplified, and sent for Sanger sequencing using primers designed with SnapGene (version 6.1.2; Dotmatics) and compatible with TIDE analysis [46]. Sequencing results were used to estimate the spectrum and frequency of small insertions and deletions (indels) generated in the different samples and time points following the TIDE analysis methodology [46].

Protein isolation and Western blotting

Cells were lysed in radioimmunospecipitation buffer (50 mM Tris-HCl, pH 8, 100 mM NaCl, 1 mM EDTA, 1% NP40, 0.5% sodium deoxycholate, and 0.1% sodium dodecyl sulfate) containing Halt Protease and Phosphatase Inhibitor Cocktail (#78440; Thermo Fisher Scientific). Lysates were cleared by centrifugation at 4°C and protein concentrations were determined with the BCA Protein Assay Kit (#23227; Thermo Fisher Scientific) using an Infinite 200 PRO plate reader. BlueEye Prestained Protein Marker (#PS-104; Jena Bioscience) and 20 µg of protein were separated on NuPAGE 4-12% Bis-Tris Mini Protein Gels (#NP0323; Thermo Fisher Scientific) using NuPAGE MOPS SDS Running Buffer (#NP0001; Thermo

Fisher Scientific), and transferred overnight at 4°C onto Nitrocellulose Membranes (#88018; Thermo Fisher Scientific) in transfer buffer (25 mM Tris, 2 M Glycine, 20% methanol in demineralized water). Membranes were stained with Ponceau S (#ab270042; Abcam) and imaged with Fusion FX (Vilber), blocked in 5% bovine serum albumin (BSA; #A8022; Sigma-Aldrich) in PBS-T (0.05% Tween-20), and incubated in 5% BSA in PBS-T overnight at 4°C with the following primary antibodies: β -actin (1:10,000, #A5441; Sigma-Aldrich), C-MYC (1:1,000, #ab32072, Abcam), and phospho (T37/T46)-4E-BP1 (#2855), 4E-BP1 (#9644), phospho (S406)-eIF4B (#8151), phospho (S422)-eIF4B (#3591), eIF4B (#3592), phospho (S235/S236)-S6 ribosomal protein (#2211), and S6 ribosomal protein (all 1:1,000, #2217; all Cell Signaling Technology). Membranes were washed with PBS-T and incubated with anti-rabbit HRP (1:10,000, #P0448; DAKO) or anti-mouse HRP (1:10,000, #G-21040; Thermo Fisher Scientific) secondary antibodies in 5% BSA in PBS-T for 1 h at room temperature. Membranes were then washed in PBS-T, developed using SuperSignal West Pico PLUS Chemiluminescent Substrate (#34580; Thermo Fisher Scientific), and imaged with Fusion FX (Vilber).

Mouse models

To generate mouse ILC tumors deficient for *Cdh1* and *Trp53* and simultaneously expressing *Myc^{ERT2}*, 6-wk-old FVB/NCrl *Cdh1^{F/F};Trp53^{F/F}* female mice, genotyped as previously described [24], were intraductally injected with a lentivirus encoding *Myc^{ERT2}*-P2A-Cre using a previously established methodology [83]. Briefly, 20 μ l of high-titer lentivirus was injected into the third and fourth mammary glands by using a 34G needle. Lentiviral titers ranging from 2×10^8 to 2×10^9 TU/ml were used. Mice were weekly monitored for tumor development and sacrificed when reaching the humane endpoint.

To allograft tumors, cryopreserved 1 mm³ tumor fragments derived from the *Cdh1^{F/F};Trp53^{F/F};Myc^{ERT2}*-P2A-Cre (EP-*Myc^{ERT2}*) model were orthotopically transplanted into the right mammary fat pad of 8-wk-old syngeneic FVB/NCrl female mice (Janvier labs). Mice were twice weekly weighed and monitored for mammary tumor development and, as soon as tumors reached a volume of 62.5 mm³ (5 $\times$ 5 mm, measured in two dimensions using a caliper), mice were randomly allocated to vehicle versus AZD8055 mTORi treatment arms and normal food pellets versus TAM 400-citrate pellets (#TD55125; Envigo) to induce nuclear translocation of MYC-ERT2. The treatments were performed daily via oral gavage using vehicle (0.5% hydroxypropylmethylcellulose and 0.1% TWEEN 80 in demineralized water) or 20 mg/kg AZD8055 while food pellets were provided ad libitum. Mice were sacrificed 1 h after the last AZD8055 dosing when reaching the humane endpoint. AZD8055 intervention studies using cryopreserved KEP ILC pieces were previously described [21] using the same AZD8055 regimen as described above for EP-*Myc^{ERT2}* allografts. Vehicle-treated control tumors were harvested after 5 d or when they reached a volume of 1,500 mm³. Sensitive tumors were harvested after 5 d of AZD8055 treatment. Resistant tumors were obtained when mice reached the humane endpoint as stated above.

For the human breast cancer patient-derived xenograft model, 25,000 MDA-MB-468 cells overexpressing *Akaluc* or *MYC-P2A-Akaluc* were injected into the fourth mammary glands of 6-wk-old immunodeficient NXG (Janvier labs) female mice, as previously described [49,50]. Animals were weighed and monitored for mammary tumor development twice weekly and, as soon as tumors reached a volume of 62.5 mm³ as per the above-mentioned formula, mice were randomly allocated to AZD8055, everolimus, or the corresponding vehicle treatment arms. Mice treated with AZD8055-vehicle or everolimus-vehicle showed no obvious differences in tumor growth kinetics and were thus grouped into single-vehicle control cohorts for MDA-MB-468-*Akaluc* and MDA-MB-468-*MYC-P2A-Akaluc*-bearing mice, respectively. AZD8055 daily treatment regimen was 20 mg/kg AZD8055 in 0.5% hydroxypropylmethylcellulose and 0.1% TWEEN 80 in demineralized water as described for the EP-*Myc*^{ERT2} allografts. Everolimus was administered daily via oral gavage at 5 mg/kg as a microemulsion solution diluted with 5% glucose in demineralized water. Mice were sacrificed 1 h after the last mTORi dosing when reaching the humane endpoint. For a subset of mice within each treatment group, tumor development was additionally monitored once a week by bioluminescence using an IVIS Spectrum instrument (PerkinElmer). Animals were anesthetized prior to the administration of 0.05 mg/kg of Akalumine-HCl substrate (courtesy of Leiden University) diluted in water. The bioluminescence signal was normalized to radiance (photons/second/centimeter²/steradian), and the region of interest tool was used to determine the signal of individual tumors. IVIS was stopped after vehicle groups reached the humane endpoint.

For all mouse experiments, endpoint was reached when the total mammary tumor burden was a volume of either 1,500 mm³ for a single tumor or 2,000 mm³ for cumulative tumors or the mice suffered from clinical signs of distress, such as respiratory distress, ascites, distended abdomen, rapid weight loss, and/or severe anemia, caused by primary tumor burden or metastatic disease. The maximal permitted disease endpoints were not exceeded in any of the experiments. The sample size was determined using G*Power software (version 3.1; [84]) and was large enough to measure the effect size. Tumor measurements and post-mortem analyses were performed in a blinded fashion. The mouse colony was housed in a certified animal facility with a 12-h light/dark cycle in a temperature-controlled room. Mice were kept in individually ventilated cages, and food and water were provided ad libitum. All animal experiments were conducted under the Central Animal Testing Committee licenses number 9 and 24 (CCD 9, AVD3010020172464, appendix 2; CCD 24, AVD30100202011584, appendix 1), approved by the Animal Ethics Committee of the Netherlands Cancer Institute, and performed in accordance with institutional, national, and European guidelines for animal care and use.

Histology and IHC of mouse tumors

Tissues were formalin-fixed and paraffin-embedded (FFPE) by routine procedures and sectioned for H&E histochemical staining and IHC stainings, which were performed as previously described [85] using primary antibodies against E-cadherin (1:200, #3195; Cell Signaling Technology), C-MYC (1:100, #ab32072; Abcam), MHC-II (1:100, #NBP2-21789; Novusbio), CD4 (1:1,000, #14-9766-82; eBioscience), FOXP3 (1:200, #14-5773; eBioscience), phospho STAT1 (1:800, #9167; Cell Signaling Technology), CD3 (1:600, #RM-9107-S1; Thermo Fisher Scientific), LY6-G (1:150, #551459; BD Pharmingen), B-220 (1:4,000, #557390; BD Pharmingen), CD8 (1:2,000, #14-0808-82; eBioscience), or NKp46 (1:400, #AF2225; R&D Systems). For E-cadherin, C-MYC, MHC-II, CD4, phospho STAT1, CD3, CD8, and NKp46 IHC stainings, antigen retrieval was done with Tris-EDTA at pH 9; for FOXP3 and B-220 IHC stainings, it was done with Citrate Buffer (#CBB999; Scytek); and for LY6-G, it was done with 20 µg/ml Proteinase-K (#P6556; Sigma-Aldrich). Sections were incubated with primary antibodies overnight at 4°C. Primary antibodies against E-cadherin, C-MYC, CD3, and phospho STAT1 were labeled with the EnVision+ HRP Labelled Polymer Anti-Rabbit System (K4003; Dako); primary antibodies against MHC-II, B-220, CD8, LY6-G, and FOXP3 were labeled with Goat-anti-Rat-Bio (1:150, #3052-08; Biotech) and subsequently streptavidin/HRP (1:200, #P0397; Dako); and the primary antibody against NKp46 was labeled with Rabbit-anti-Goat-Bio (1:100, #E0466; Dako) and subsequently streptavidin/HRP. Secondary antibodies were then visualized with the Liquid DAB+ Substrate Chromogen System (#K3468; Dako) and counterstained with hematoxylin. The antibodies used were independently validated by certified pathologists by evaluation of IHC results in positive and negative biological control FFPE tissues to ensure specificity and sensitivity. In addition, negative technical controls for each primary antibody were performed. EP-Myc^{ERT2} mammary tumors were classified as solid ILCs using H&E and E-Cadherin slides, and according to the international consensus of mammary pathology [86]. Slides were digitally processed using a PANNORAMIC 1000 whole slide scanner (3DHISTECH), captured with CaseViewer software (version 3.2.2.1, 3DHISTECH), and quantified using HALO Image Analysis Platform (version 3.5.3577; Indica Labs). MYC IHC stains were HALO-quantified by calculating a histo (H)-score for each tumor. H-score was defined as follows: $H\text{-score} = 1 \times (\% \text{ tumor cells with weak staining intensity}) + 2 \times (\% \text{ tumor cells with moderate staining intensity}) + 3 \times (\% \text{ tumor cells with strong staining intensity})$, resulting in a score between 0 and 300. IHC stains for immune markers were HALO-quantified by counting marker-positive cells per 0.25 mm² tumor area to calculate immune cell densities per area. Per tumor, ten 0.25-mm² regions were randomly selected for quantification. IHC stains for MYC and immune markers were performed on consecutive slides and stains quantified using matched regions. All slides and quantifications thereof were reviewed by a comparative pathologist (X. Chao) in a blinded manner.

LC-WGS data generation and analysis

Genomic DNA from KEP control, sensitive, and resistant tumors was isolated using the ISOLATE II Genomic DNA Kit (#BIO-52066; Bioline) according to the manufacturer's guidelines. LC-WGS for copy number analysis was performed using double-stranded DNA (dsDNA) and quantified with the Qubit dsDNAHS Assay Kit (#Q32851; Invitrogen). 2 mg of dsDNA were fragmented by Covaris shearing and purified using 1.8× Agencourt AMPure XP PCR Purification beads according to the manufacturer's protocol (#A63881; Beckman Coulter). Next, sheared DNA was quantified on a BioAnalyzer system with the DNA7500 assay kit (#5067-1506; Agilent Technologies). Library preparation for Illumina sequencing was carried out with 1 mg of DNA and KAPA HTP Library Preparation Kit (#KK8234; KAPA Biosystems). To obtain a sufficient yield for sequencing, four to six PCR cycles were performed during the library enrichment step. Prepared libraries were cleaned up using 1× AMPure XP beads and analyzed on a BioAnalyzer system using the DNA7500 chips to determine the molarity. Finally, up to 11 uniquely indexed samples were pooled (equimolar pooling) in a final concentration of 10 nM and sequenced on an Illumina HiSeq2500 machine in one lane of a single read 65 bp run, according to the manufacturer's instructions. Sequencing reads were mapped to the human reference genome GRCh38 using Burrows-Wheeler Alignment (BWA-MEM, version 11.0.7.5a; [87]). Reads with mapping quality lower than 37 were excluded. The resulting alignments were analyzed with QDNAseq (version 1.14.0) using sequence mappability and guanine-cytosine content correction and a bin size of 20,000 bp to generate segmented copy number values [88].

RNA-seq data generation and analysis

Total RNA extraction was performed for frozen tissue samples from KEP control, sensitive, and resistant tumors, as well as for KEP cell lines transduced with lentivirus containing *GFP* or *Myc* and treated with DMSO or 50 nM AZD8055 using the RNeasy Mini Kit (#74104; Qiagen) according to the manufacturer's guidelines. Quality and quantity of RNA were assessed using the 2100 Bioanalyzer system and a Nano chip (Agilent). RNA samples with RNA integrity number >8 were subjected to polyA-stranded library preparation using the TruSeq RNA Library Prep Kit v2 (#RS-122-2001/2; Illumina) according to the manufacturer's instructions, quality-checked with the 2100 Bioanalyzer system using a 7,500 chip, and pooled equimolar into a 10-nM sequencing stock solution. Libraries from KEP tumors were sequenced with 100 bp paired-end reads using the HiSeq 2500 System with V4 chemistry and libraries from KEP cell lines were sequenced with 51 bp paired-end reads using NovaSeq 6000. Sequencing reads were mapped to the mouse reference genome Ensembl GRCm38 using STAR (version 2.7.2; [89]). For gene expression analysis, read counts were quantified by featureCounts (version 1.6.2; [90]). Genes with count per million (CPM) values higher than one in at least 10% of the total number of samples were considered expressed and used for downstream analysis. Read counts for expressed genes were normalized by the Trimmed Mean of M-value method using edgeR (version 3.26.6; [91]). Differential expression analysis was performed using limma-voom (version

3.52.4; [92]). For TE analysis, read counts were quantified by HTSeq (version 1.99.2; [93]) for the transcript that represents each gene. The representative transcript for a gene was defined as the one that shares most exons with other transcripts in the gene, using CGAT (version 0.6.5; [94]).

Ribo-seq data generation and analysis

1.4 million GFP or MYC-overexpressing KEP1.23 cells were seeded in 10-cm dishes and after 24 h subjected to DMSO or 50 nM AZD8055 treatment, which represented an IC80 for GFP cells and an IC30 for MYC cells. Ribosome-protected fragments sequencing libraries were constructed as previously described [95]. In brief, 24 h after treatment, cells were pretreated with cycloheximide (100 µg/ml) for 1 min, followed by detergent lysis and ribosome footprinting by RNase I digestion. Library constructions were generated from 26 to 34 nt footprint fragments and sequenced with 75-bp single-reads using the NextSeq HIGH system (Illumina). Sequencing reads were first trimmed to remove adaptor sequences in reads using the Cutadapt (version 3.5; [96]) and resulting reads shorter than 20 bp were discarded. Reads derived from rRNAs and tRNAs were filtered by alignment to sets of mouse rRNA and tRNA references using bowtie2 (version 2.4.5; [97]). The remaining reads were aligned to the mouse reference genome Mouse_M15_CTAT_lib_Nov012017 using the Tophat2 (version 2.1.1; [98]). Only primary alignments with mapping qualities of 10 or greater were used for the quantification of a representative transcript for each gene using HTSeq (version 1.99.2; [93]). The transcript that shared most exons with other transcripts in the gene was selected as a representative transcript using CGAT (version 0.6.5; [94]).

Analysis of TE and RNA-binding motif

TE was calculated by dividing the number of ribosome footprint reads by the number of mRNA reads. DESeq2 (version 1.30.1) was used to normalize RNA-seq and Ribo-seq read counts for representative transcripts and subsequently calculate TE based on a design formula that took into account the sample groups (GFP-vehicle, GFP-AZD8055, MYC-vehicle, MYC-AZD8055) and the data type (RNA-seq, Ribo-seq): counts ~ group + group*data type. The interaction terms were used to compute TEs in each group, and the changes in TE were calculated by comparing the coefficient of the interaction term between groups. RNA-binding motif enrichment analysis was performed using the analysis of motif enrichment (AME; version 5.4.1; [99]) embedded in the MEME Suite [100]. 5'-UTR sequences of transcripts that were upregulated by MYC were against previously reported RNA binding motifs [101,102].

[³⁵S]-methionine incorporation assay

400,000 GFP or MYC-overexpressing KEP1.23 cells were seeded into 6-well plates and after 24 h subjected to media containing DMSO, 50 nM AZD8055, or 10 nM everolimus. Cells were incubated for 2 h at 37°C. For the last 20 min, complete media was substituted by DMEM methionine-free media (#21013024; Thermo Fisher Scientific). Subsequently, cells

were incubated with 7 μ l [35 S]-methionine label (#NEG772007MC; PerkinElmer) for 1 h at 37°C. As positive control, 100 μ g/ml cycloheximide was added with [35 S]-methionine to block protein synthesis. Subsequent cell harvesting, cell lysis, and protein precipitation steps were previously described [103]. Scintillation was measured using a liquid scintillation counter (PerkinElmer) and the activity was normalized by total protein content, as determined by Bradford Assay (#5000006; Bio-Rad).

MS data generation and analysis

Peptides were labeled with TMT 10plex isobaric tags (Pierce) according to the manufacturer's protocol, including an internal standard consisting of equal aliquots of all digests to be analyzed. For global proteome analysis, fractions were concatenated to 22 fractions and 11 fractions; for phosphoproteome analysis, fractions were concatenated to five fractions. For phosphoproteomic analysis, phosphorylated peptides were enriched from concatenated fractions using a High-Select Fe-NTA Phosphopeptide Enrichment Kit (Thermo Fisher Scientific), according to the manufacturer's instructions. Prior to MS data analysis, the peptides used for proteome analysis were reconstituted in 2% formic acid. Peptide mixtures were analyzed by nanoLC-MS/MS on an Orbitrap Fusion Tribrid mass spectrometer equipped with a Proxeon nLC1000 system (Thermo Fisher Scientific; [104]). Global proteome data were analyzed by Proteome Discoverer (version 2.2.0.388) and phosphoproteome data were analyzed by MaxQuant (version 1.5.8.3). MS/MS data were searched against the *Mus musculus* Swissprot database (17,397 entries, release 2018_02) using Mascot. False discovery rates (FDRs) for peptide and protein identification were set to 1%, and as additional filter Mascot Ions score >20 was set. The TMT dataset was quantified using Most Confident Centroid with the "reporter ions quantifier" node. For both global and phosphoproteome data, the intensity ratio ($\log_2$ scale) divided by the intensities of internal control was processed by quantile normalization followed by batch correction using Combat [105].

GSVA

For RNA-seq and MS-based global proteomics data, the GSVA [27] was performed based on MSigDB Hallmark [28] and GO gene sets [106]. $\log_2$ scale of CPM for RNA-seq and normalized abundance ratio for global proteomics data were used to perform GSVA with the default parameters (kcdf = "Gaussian," mx.differ = "TRUE"). Gene sets smaller than 10 were excluded from the GSVA analysis.

Phosphosite-specific signature analysis

Phosphosites identified from MS-phosphoproteomics analysis of KEP tumors were first mapped to the orthologous phosphosites in human proteins using the PhosphoSitePlus database (<https://www.phosphosite.org>). The phosphosites without corresponding human orthologous sites were excluded. Then, normalized TMT intensities for filtered phosphosites were averaged for each KEP tumor group to obtain average intensities for

control, sensitive, and resistant tumors. Finally, post-translational modification (PTM) signature enrichment analysis was performed based on the PTMSigDB database to identify significantly enriched upstream signals for each tumor group [25].

Analysis of public pharmacological dataset

AUC values for mTORi and copy number status of *MYC* for CCLE human cancer cell lines were obtained from the cBioPortal database [41]. Data for the mTORi AZD8055, rapamycin, and temsirolimus and the PI3Ki BEZ235, JW-7-52-1, omipalisib, OSI-027, and QL-VIII-58 were examined to associate *MYC* amplification status with PI3Ki/mTORi response. AUC values were compared between cell lines with or without *MYC* amplifications by one-tailed Student's *t* test. A dataset of 30 breast cancer cell lines including response data for the mTORi AZD8055, BEZ235, MK2206, and GDC0941 was also used to examine *MYC* amplification versus PI3Ki/mTORi response association [18]. Copy number estimates ($\log_2$ ratio) obtained from DNA sequencing were used to define *MYC* amplification ($\log_2$ ratio >1). $\log_{10}(\text{IC}_{50})$ values were compared between cell lines with and without *MYC* amplification and *MYC* amplification by one-tailed Student's *t* test. To measure the correlation between *MYC* protein and AZD8055 response in human breast cancer cell lines, *MYC* protein abundances measured by MS were obtained from the Cancer Cell Line Encyclopedia (CCLE) database [42] and AZD8055 response was obtained from the Genomics of Drug Sensitivity in Cancer (GDSC) database [43]. The HCC70 cell line was excluded from this correlation analysis due to an apparent discrepancy between MS data, which showed very high protein abundance, and the RNA-seq and RPPA data, which indicated low RNA and protein levels, respectively.

Analysis of gene dependency dataset

Gene dependency data (CRISPR, Public 22Q1) was obtained from the DepMap database [45] to evaluate the essentiality of the mTOR complex components MTOR, RPTOR, and RICTOR for cancer cell survival. *MYC* copy number status from Cell Model Passports [107] was used to compare gene deletion effects in cancer cell lines containing *MYC* amplifications versus lines without *MYC* amplifications.

Analysis of HMF dataset

Processed WGS data on metastatic tumors before or after receiving everolimus and clinical information were obtained from HMF under data-sharing agreement DR-184. The data can be obtained through standardized procedures and request forms online (<https://www.hartwigmedicalfoundation.nl/en/>). The WGS data was analyzed based on the HMF bioinformatics pipeline (<https://github.com/hartwigmedical/pipeline>). In addition, clinical information on those patients was obtained, including treatment response and treatment duration. Treatment response was measured using the RECIST criteria. Copy number values larger than two times the sample ploidy were defined as amplification according to the earlier HMF WGS study [58]. The samples with low purity (<35%) and no available treatment

response and treatment duration information were excluded from downstream analysis. In total, 40 patients including 31 breast cancer patients were analyzed in this study.

IHC of MYC in ER+/HER2– breast cancer samples

FFPE blocks from primary tumor of the patients participating in the everolimus biomarker study (EudraCT number 2013-004120-11) were collected as previously described [61]. The study was approved by the Independent Ethics Committee of Amsterdam University Medical Centers and Institutional Review Board at each participating site and was performed in compliance with Good Clinical Practices, the Declaration of Helsinki, and carried out in keeping with applicable local law(s) and regulation(s). Also, all patients signed informed consent before enrolment, which included, among others, recording the efficacy of treatment, if possible a non-osseous tumor biopsy at baseline, and the retrieval of archival primary tumor tissue. Tissue microarrays (TMAs) were constructed using three 0.6-mm cores taken from the blocks. Paraffin sections were cut at 3 μ m, heated at 75°C for 28 min, and deparaffinized in the instrument with EZ prep solution (Ventana Medical Systems). Heat-induced antigen retrieval was carried out using Cell Conditioning 1 (Ventana Medical Systems) for 64 min at 95°C. IHC staining was done on an automated BenchMark XT Slide Stainer (Ventana Medical Systems) using a C-MYC primary antibody (Y69, 6504612001; Roche Ventana) according to the manufacturer's protocol. Bound antibody was detected using the OptiView DAB Detection Kit (Ventana Medical Systems). Slides were counterstained with Hematoxylin II and Bluing Reagent (Ventana Medical Systems). The stained TMA slides were scanned and the percentage of positive tumor cells (0–100%) was scored using Slide Score (<https://www.slidescore.com>). For log-rank test and multivariate Cox-regression analysis, patients were divided into MYC-positive (>0%) and MYC-negative (0%) groups depending on the presence of MYC-positive tumor cells on the TMA slides. Details on the collection of clinicopathological information from ER+/HER2– metastatic breast cancer patients treated with everolimus and exemestane were previously described [61].

ONLINE SUPPLEMENTAL MATERIAL

The supplementary information provides an overview of multiomic molecular profiling of KEP tumors performed in this study (Fig. S1), differences in biological processes represented by different KEP tumor groups (Fig. S2), associations between MYC status and mTOR complex dependency, as well as the response to PI3K and mTOR inhibitors (Fig. S3), in vitro drug response assay for inhibitors of PI3K and mTOR in human breast cancer cell lines (Fig. S4), and RNA-binding proteins associated with MYC-driven translation regulation (Fig. S5). <https://rupress.org/jem/article/220/11/e20211743/276218/MYC-is-a-clinically-significant-driver-of-mTOR>

DATA AVAILABILITY

The KEP tumor sequencing and MS proteomics data underlying Figs. 1, 2, 3, S1, S2, and S5 are available in the European Nucleotide Archive under the accession number PRJEB46418 and Proteomics Identification Database under the accession number PXD041927, respectively. RNA- and Ribo-seq data for KEP cell lines underlying Figs. 7 and S5 are available in the European Nucleotide Archive under the accession number PRJEB61759. CCLE, GDSC, and cBioPortal pharmacogenomic data underlying Fig. 4 are available through the respective data portals [41–43]. Gene dependency data underlying Fig. S3 are available in the DepMap portal (<https://depmap.org/portal/>). WGS and clinical metadata underlying Fig. 9 were provided by the HMF under license number DR-185 and can be obtained through standardized data request procedures (<https://www.hartwigmedicalfoundation.nl/en>).

ACKNOWLEDGMENTS

We thank Catrin Lutz (Netherlands Cancer Institute, Amsterdam, Netherlands) for assisting the xeno-transplantation work of human breast cancer cell lines, Koen Schipper (The Institute of Cancer Research, London, UK) for providing the SIN.LV.SF-Akaluc-T2A-Puro lentiviral vector, Sarah Moser (Netherlands Cancer Institute, Amsterdam, Netherlands) for help with the CRISPR knock-out experiment, Daniel Vis (Netherlands Cancer Institute, Amsterdam, Netherlands) for sharing insights on the Bliss model, and Pierre-René Körner (Netherlands Cancer Institute, Amsterdam, Netherlands) for help with the Ribo-seq analysis. We also thank Anna Dopler and Joana da Silva (Netherlands Cancer Institute, Amsterdam, Netherlands) for their insights and help with the [³⁵S]-methionine radiolabeling assay. We thank the people from the preclinical intervention unit of the Mouse Clinic for Cancer and Ageing at the Netherlands Cancer Institute for their technical support performing the animal experiments. We express gratitude to our colleagues from the Wessels and Jonkers groups for insightful discussions. We thank the Netherlands Cancer Institute genomics core facility, the Research High-Performance Computing facility, the animal pathology facility, the core facility molecular pathology and biobanking, and the proteomics facility for their excellent support. This publication and the underlying study have been made possible partly on the basis of the data that the Hartwig Medical Foundation (Utrecht, Netherlands) and the Center of Personalised Cancer Treatment (Rotterdam, Netherlands) have made available to the study. Research at the Netherlands Cancer Institute is supported by institutional grants from the Dutch Cancer Society and the Dutch Ministry of Health, Welfare and Sport. This work is part of the Oncode Institute, which is partly financed by the Dutch Cancer Society.

This work was funded by the Dutch Cancer Society (KWF 2017-61169 to J. Bhin, L.F.A. Wessels, and J. Jonkers; KWF 2020-12894 to D. Zingg, L.F.A. Wessels, and J. Jonkers), the Swiss National Science Foundation (SNF P2ZHP3_175027 to D. Zingg), the Dutch NWO X-omics Initiative (to L. Hoekman and M. Altelaar), and a faculty research grant of Yonsei University College of Medicine (6-2023-0156 and 2023-32-0036). Open Access funding provided by the Netherlands Cancer Institute.

REFERENCES

1. Ciriello, G. *et al.* (2015) Comprehensive molecular portraits of invasive lobular breast cancer. *Cell* 163, 506–519
2. Mukohara, T. (2015) PI3K mutations in breast cancer: Prognostic and therapeutic implications. *Breast Cancer* 7, 111–123
3. Bahrami, A. *et al.* (2018) The therapeutic potential of PI3K/Akt/mTOR inhibitors in breast cancer: Rational and progress. *J. Cell. Biochem* 119, 213–222
4. Baselga, J. *et al.* (2012) Everolimus in postmenopausal hormone-receptor-positive advanced breast cancer. *N. Engl. J. Med* 366, 520–529
5. André, F. *et al.* (2019) Alpelisib for PIK3CA-Mutated, Hormone Receptor-Positive Advanced Breast Cancer. *New England Journal of Medicine* 380, 1929–1940
6. Lee, J.J. *et al.* (2015) PI3K/Akt/mTOR inhibitors in breast cancer. *Cancer Biol. Med* 12, 3941 2015 0089
7. Yi, Z. and Ma, F. (2017) Biomarkers of everolimus sensitivity in hormone receptor-positive breast cancer. *J. Breast Cancer* 20, 321–326
8. Brandão, M. *et al.* (2019) Biomarkers of response and resistance to PI3K inhibitors in estrogen receptor-positive breast cancer patients and combination therapies involving PI3K inhibitors. *Ann. Oncol* 30:x27–x42
9. Formisano, L. *et al.* (2020) Mechanisms of resistance to mTOR inhibitors. *Crit. Rev. Oncol. Hematol* 147
10. Zunder, E.R. *et al.* (2008) Discovery of drug-resistant and drug-sensitizing mutations in the oncogenic PI3K isoform p110 alpha. *Cancer Cell* 14, 180–192
11. Wu, T.-J. *et al.* (2015) Identification of a non-gatekeeper hot spot for drugresistant mutations in mTOR kinase. *Cell Rep* 11, 446–459
12. Lorenz, M.C. and Heitman, J. (1995) TOR mutations confer rapamycin resistance by preventing interaction with FKBP12-rapamycin. *J. Biol. Chem* 270, 27531–27537
13. Nakanishi, Y. *et al.* (2016) Activating mutations in PIK3CB confer resistance to PI3K inhibition and define a novel oncogenic role for p110β. *Cancer Res* 76, 1193–1203
14. Rozengurt, E. *et al.* (2014) Suppression of feedback loops mediated by PI3K/mTOR induces multiple overactivation of compensatory pathways: An unintended consequence leading to drug resistance. *Mol. Cancer Ther* 13, 2477–2488
15. O'Reilly, K.E. *et al.* (2006) mTOR inhibition induces upstream receptor tyrosine kinase signaling and activates Akt. *Cancer Res* 66, 1500–1508
16. Alain, T. *et al.* (2012) eIF4E/4E-BP ratio predicts the efficacy of mTOR targeted therapies. *Cancer Res* 72, 6468–6476
17. Cope, C.L. *et al.* (2014) Adaptation to mTOR kinase inhibitors by amplification of eIF4E to maintain cap-dependent translation. *J. Cell Sci* 127, 788–800
18. Jastrzebski, K. *et al.* (2018) Integrative Modeling Identifies Key Determinants of Inhibitor Sensitivity in Breast Cancer Cell Lines. *Cancer Research* 78, 4396–4410
19. Curtis, C. *et al.* (2012) The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* 486, 346–352
20. Michaut, M. *et al.* (2016) Integration of genomic, transcriptomic and proteomic data identifies two biologically distinct subtypes of invasive lobular breast cancer. *Sci. Rep* 6
21. Klarenbeek, S. *et al.* (2020) Response of metastatic mouse invasive lobular carcinoma to mTOR inhibition is partly mediated by the adaptive immune system. *Oncolimmunology* 9
22. Doornebal, C.W. *et al.* (2013) A preclinical mouse model of invasive lobular breast cancer metastasis. *Cancer Res* 73, 5472-11-4208

23. Klarenbeek, S. *et al.* (2020) Response of metastatic mouse invasive lobular carcinoma to mTOR inhibition is partly mediated by the adaptive immune system. *Oncol Immunology* 9
24. Derksen, P.W.B. *et al.* (2006) Somatic inactivation of E-cadherin and p53 in mice leads to metastatic lobular mammary carcinoma through induction of anoikis resistance and angiogenesis. *Cancer Cell* 10, 437–449
25. Krug, K. *et al.* (2019) A curated resource for phosphosite-specific signature analysis. *Mol. Cell. Proteomics* 18, 576–593
26. Wang, Y. *et al.* (2016) Mitotic MELK-eIF4B signaling controls protein synthesis and tumor cell survival. *Proc. Natl. Acad. Sci. USA* 113, 9810–9815
27. Hanzelmann, S. *et al.* (2013) GSVA: Gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 14
28. Liberzon, A. *et al.* (2015) The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* 1, 417–425
29. Kuleshov, M.V. *et al.* (2016) Enrichr: A comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* 44:W90–W97
30. Muthalagu, N. *et al.* (2020) Repression of the type I interferon pathway underlies MYC- and KRASdependent evasion of NK and B cells in pancreatic ductal adenocarcinoma. *Cancer Discov* 10, 19–0620
31. Swaminathan, S. *et al.* (2020) MYC functions as a switch for natural killer cell-mediated immune surveillance of lymphoid malignancies. *Nat. Commun* 11
32. Kim, T.W. *et al.* (2016) Transcriptional repression of IFN regulatory factor 7 by MYC is critical for type I IFN production in human plasmacytoid dendritic cells. *J. Immunol* 197, 3348–3359
33. Zimmerli, D. *et al.* (2022) MYC promotes immune-suppression in triple-negative breast cancer via inhibition of interferon signaling. *Nat. Commun* 13
34. Yoshihara, K. *et al.* (2013) Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun* 4, 2612
35. Momcilovic, M. *et al.* (2018) The GSK3 signaling axis regulates adaptive glutamine metabolism in lung squamous cell carcinoma. *Cancer Cell* 33, 905–921
36. Tanaka, K. *et al.* (2015) Compensatory glutamine metabolism promotes glioblastoma resistance to mTOR inhibitor treatment. *J. Clin. Invest* 125, 1591–1602
37. Wise, D.R. *et al.* (2008) Myc regulates a transcriptional program that stimulates mitochondrial glutaminolysis and leads to glutamine addiction. *Proc. Natl. Acad. Sci. USA* 105, 18782–18787
38. Bott, A.J. *et al.* (2015) Oncogenic myc induces expression of glutamine synthetase through promoter demethylation. *Cell Metab* 22, 1068–1077
39. Riggelen, J. *et al.* (2010) MYC as a regulator of ribosome biogenesis and protein synthesis. *Nat. Rev. Cancer* 10, 301–309
40. Schmidt, E.V. (2004) The role of c-myc in regulation of translation initiation. *Oncogene* 23, 3217–3221
41. Gao, J. *et al.* (2013) Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal* 6:pl1
42. Barretina, J. *et al.* (2012) The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 483, 603–607
43. Yang, W. *et al.* (2013) Genomics of drug sensitivity in cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res* 41:D955–D961
44. Jastrzebski, K. *et al.* (2018) Integrative modeling identifies key determinants of inhibitor sensitivity in breast cancer cell lines. *Cancer Res* 78, 2698
45. Tsherniak, A. *et al.* (2017) Defining a cancer dependency map. *Cell* 170, 564–576

46. Brinkman, E.K. *et al.* (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* 42:e168
47. Annunziato, S. *et al.* (2016) Modeling invasive lobular breast carcinoma by CRISPR/Cas9-mediated somatic genome editing of the mammary gland. *Genes Dev* 30, 1470–1480
48. Zingg, D. *et al.* (2022) Truncated FGFR2 is a clinically actionable oncogene in multiple cancers. *Nature* 608, 609–617
49. Valdez, K.E. *et al.* (2011) Human primary ductal carcinoma in situ (DCIS) subtype-specific pathology is preserved in a mouse intraductal (MIND) xenograft model. *J. Pathol* 225, 565–573
50. Hutten, S.J. *et al.* (2023) A living biobank of patient-derived ductal carcinoma in situ mouseintraductal xenografts identifies risk factors for invasive progression. *Cancer Cell* 41, 986–1002
51. Jefferies, H.B. *et al.* (1997) Rapamycin suppresses 59TOP mRNA translation through inhibition of p70s6k. *EMBO J* 16, 3693–3704
52. Yang, M. *et al.* (2022) The translational regulation in mTOR pathway. *Biomolecules* 12
53. Thoreen, C.C. *et al.* (2012) A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature* 485, 109–113
54. Philippe, L. *et al.* (2020) Global analysis of LARP1 translation targets reveals tunable and dynamic features of 59 TOP motifs. *Proc. Natl. Acad. Sci. USA* 117, 5319–5328
55. Oki, S. *et al.* (2018) ChIP-atlas: A data-mining suite powered by full integration of public ChIP-seq data. *EMBO Rep* 19, 46255
56. Hart, J.R. *et al.* (2014) Inhibitor of MYC identified in a Kröhnke pyridine library. *Proc. Natl. Acad. Sci. USA* 111, 12556–12561
57. Greco, W.R. *et al.* (1995) The search for synergy: a critical review from a response surface perspective. *Pharmacol Rev* 47, 331–385
58. Priestley, P. *et al.* (2019) Pancancer whole-genome analyses of metastatic solid tumours. *Nature* 575, 210–216
59. Bogemann, M. *et al.* (2020) Using data from a sickness fund claims database to assess the treatment patterns and healthcare resource utilization among patients with metastatic renal cell carcinoma in Germany. *Urol. Int* 104, 982–993
60. Hari, P. *et al.* (2018) Prolonged duration of therapy is associated with improved survival in patients treated for relapsed/refractory multiple myeloma in routine clinical care in the United States. *Clin. Lymphoma Myeloma Leuk* 18, 152–160
61. Kruger, D.T. *et al.* (2020) PI3K pathway protein analyses in metastatic breast cancer patients receiving standard everolimus and exemestane. *J. Cancer Res. Clin. Oncol* 146, 3013–3023
62. Allen-Petersen, B.L. *et al.* (2019) Activation of PP2A and inhibition of mTOR synergistically reduce MYC signaling and decrease tumor growth in pancreatic ductal adenocarcinoma. *Cancer Res* 79, 209–219
63. Ilic, N. *et al.* (2011) PI3K-targeted therapy can be evaded by gene amplification along the MYC-eukaryotic translation initiation factor 4E (eIF4E) axis. *Proc. Natl. Acad. Sci. USA* 108:E699–E708
64. Matsumoto, M. *et al.* (2015) Control of the MYC-eIF4E axis plus mTOR inhibitor treatment in small cell lung cancer. *BMC Cancer* 15
65. Muellner, M.K. *et al.* (2011) A chemical-genetic screen reveals a mechanism of resistance to PI3K inhibitors in cancer. *Nat. Chem. Biol* 7, 787–793
66. Muranen, T. *et al.* (2016) ERK and p38 MAPK activities determine sensitivity to PI3K/mTOR inhibition via regulation of MYC and YAP. *Cancer Res* 76, 1158 0008-5472-16–0155
67. Shepherd, C. *et al.* (2013) PI3K/mTOR inhibition upregulates NOTCHMYC signalling leading to an impaired cytotoxic response. *Leukemia* 27, 650–660
68. Song, X. *et al.* (2021) Synergistic targeting of CHK1 and mTOR in MYC-driven tumors. *Carcinogenesis* 42, 448–460

69. Tan, J. *et al.* (2013) PDK1 signaling toward PLK1-MYC activation confers oncogenic transformation, tumor-initiating cell activation, and resistance to mTOR-targeted therapy. *Cancer Discov* 3, 1156–1171
70. Dey, N. *et al.* (2015) MYC-xing it up with PIK3CA mutation and resistance to PI3K inhibitors: summit of two giants in breast cancers. *Am J Cancer Res* 5, 1–19
71. Fallah, Y. *et al.* (2017) MYC-driven pathways in breast cancer subtypes. *Biomolecules* 7
72. Wang, C. *et al.* (2021) Alternative approaches to target Myc for cancer treatment. *Signal Transduct. Target. Ther* 6
73. Meyers, R.M. *et al.* (2017) Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat. Genet* 49, 1779–1784
74. Nunnery, S.E. and Mayer, I.A. (2020) Targeting the PI3K/AKT/mTOR pathway in hormone-positive breast cancer. *Drugs* 80, 1685–1697
75. Rozenblit, M. *et al.* (2021) Patterns of treatment with everolimus exemestane in hormone receptor-positive HER2-negative metastatic breast cancer in the era of targeted therapy. *Breast Cancer Res* 23
76. Dhakal, A. *et al.* (2018) Outcome of everolimus based therapy in hormone receptor positive metastatic breast cancer patients after progression on palbociclib combination. *J Clin Oncol* 36, 1064–1064
77. Iwano, S. *et al.* (2018) Single-cell bioluminescence imaging of deep tissue in freely moving animals. *Science* 359, 935–939
78. Follenzi, A. *et al.* (2000) Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences. *Nat. Genet* 25, 217–222
79. Delwel, G.O. *et al.* (1993) Expression and function of the cytoplasmic variants of the integrin alpha 6 subunit in transfected K562 cells. Activation-dependent adhesion and interaction with isoforms of laminin. *J. Biol. Chem* 268, 25865–25875
80. Schipper, K. *et al.* (2019) Rebalancing of actomyosin contractility enables mammary tumor formation upon loss of E-cadherin. *Nature Communications* 10, 1–14
81. Greco, W.R. *et al.* (1995) The search for synergy: A critical review from a response surface perspective. *Pharmacol. Rev* 47, 331–385
82. Doench, J.G. *et al.* (2016) Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol* 34, 184–191
83. Annunziato, S. *et al.* (2016) Modeling invasive lobular breast carcinoma by CRISPR/Cas9-mediated somatic genome editing of the mammary gland. *Genes Dev.* 30, 1470–1480
84. Faul, F. *et al.* (2009) Statistical power analyses using G*power 3.1: Tests for correlation and regression analyses. *Behav. Res. Methods* 41, 1149–1160
85. Gogola, E. *et al.* (2018) Selective loss of PARG restores PARylation and counteracts PARP inhibitor-mediated synthetic lethality. *Cancer Cell* 33, 1078–1093
86. Cardiff, R.D. *et al.* (2000) The mammary pathology of genetically engineered mice: The consensus report and recommendations from the annapolis meeting. *Oncogene* 19, 968–988
87. Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760
88. Scheinin, I. *et al.* (2014) DNA copy number analysis of fresh and formalin-fixed specimens by shallow whole-genome sequencing with identification and exclusion of problematic regions in the genome assembly. *Genome Res* 24, 2022–2032
89. Dobin, A. *et al.* (2013) STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21
90. Liao, Y. *et al.* (2014) featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930
91. Robinson, M.D. and Oshlack, A. (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* 11:R25

92. Ritchie, M.E. *et al.* (2015) Limma powers differential expression analyses for RNAsequencing and microarray studies. *Nucleic Acids Res* 43:e47
93. Anders, S. *et al.* (2015) HTSeq—a python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169
94. Sims, D. *et al.* (2014) CGAT: Computational genomics analysis toolkit. *Bioinformatics* 30, 1290–1291
95. Ingolia, N.T. *et al.* (2012) The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat. Protoc* 7, 1534–1550
96. Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J* 17, 10–12
97. Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359
98. Kim, D. *et al.* (2013) TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14:R36
99. McLeay, R.C. and Bailey, T.L. (2010) Motif enrichment analysis: A unified framework and an evaluation on ChIP data. *BMC Bioinformatics* 11
100. Bailey, T.L. *et al.* (2009) MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Res* 37:W202–W208
101. Ray, D. *et al.* (2013) A compendium of RNA-binding motifs for decoding gene regulation. *Nature* 499, 172–177
102. Hong, S. *et al.* (2017) LARP1 functions as a molecular switch for mTORC1-mediated translation of an essential class of mRNAs. *Elife* 6, 25237
103. Silva, J. *et al.* (2022) Ribosome impairment regulates intestinal stem cell identity via ZAKα activation. *Nat Commun* 13, 4492
104. Ameziane, N. *et al.* (2015) A novel Fanconi anaemia subtype associated with a dominant-negative mutation in RAD51. *Nat. Commun* 6
105. Leek, J.T. *et al.* (2012) The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 28, 882–883
106. Carbon, S. *et al.* (2009) AmiGO: Online access to ontology and annotation data. *Bioinformatics* 25, 288–289
107. Meer, D. *et al.* (2019) Cell Model Passports—a hub for clinical, genetic and functional datasets of preclinical cancer models. *Nucleic Acids Res* 47:D923–D929

SUPPLEMENTARY FIGURES

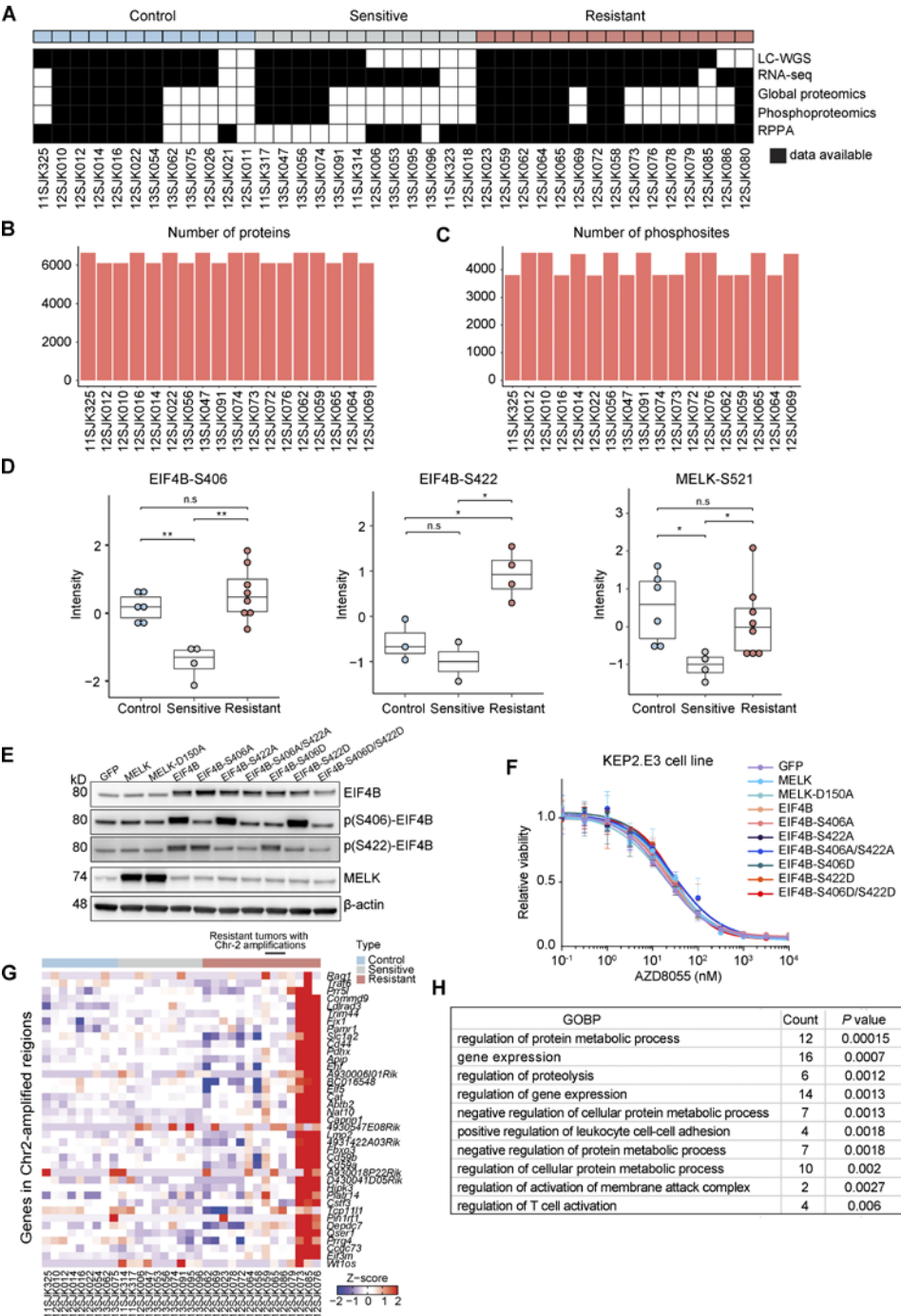


Figure S1. Multiomic molecular profiling of KEP tumors.

(A) Control ($n = 12$), sensitive ($n = 12$), and resistant ($n = 15$) KEP tumors were analyzed by LC-WGS, RNA-seq, MS-based expression proteomics, phosphoproteomics, and RPPA. **(B)** The number of proteins quantified by MS-based expression proteomics. **(C)** The number of phosphosites quantified by MS-based phosphoproteomics. **(D)** Phosphorylation levels of EIF4B and MELK were measured by MS-based phosphoproteomics across six control, four sensitive, and eight resistant KEP tumors. Phosphosites for some tumors were not available due to the limited dynamic range of MS. One-way ANOVA and Tukey's post-hoc tests were performed to compute adjusted P value (* $P < 0.05$; ** $P < 0.01$; n.s, not significant). **(E)** Western blots for EIF4B, MELK, and their phosphosites on KEP2.E3 cells transduced with lentiviral overexpression vectors for *GFP*, *Eif4b*, its phosphomimetic (*S406D* and *S422D*) and non-phosphorylatable mutants (*S406A* and *S422A*), *Melk*, or its non-phosphorylatable mutant (*S150A*). Data represent one experiment of two independent experiments. **(F)** Dose-response curves of KEP2.E3 cells expressing indicated *Eif4b* or *Melk* variants, treated with AZD8055 for 3 d, and assayed using CellTiter-Blue reagent. Data are represented as mean $\pm$ standard deviation of five replicas per group of two independent experiments. **(G)** Expression of genes present in Chr-2 amplicon identified in several resistant KEP tumors. $\log_2(\text{CPM})$ values were transformed to Z-scores. Data are shown for 9 control, 10 sensitive, and 14 resistant KEP tumors. **(H)** GOBPs significantly enriched for the Chr-2 amplicon genes.

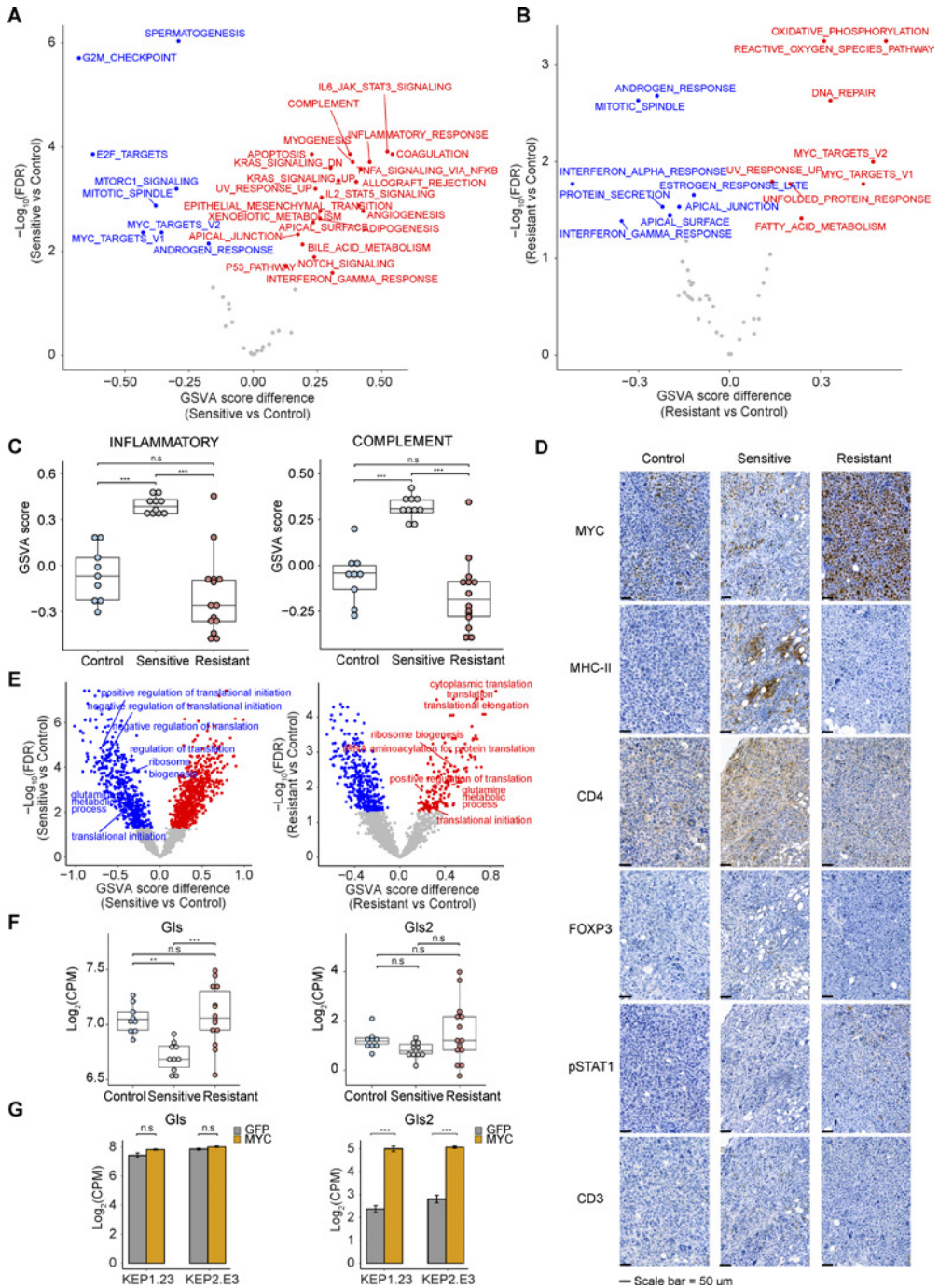


Figure S2. Biological processes differentially regulated in different KEP tumor groups.

(A and B) GSVA analysis for sensitive versus control (A) and resistant versus control (B) tumors using RNA-seq data. GSVA scores were compared using Student's *t* test. Red and blue dots, significantly (FDR < 0.05) up- and downregulated in resistant tumors. **(C)** GSVA scores for immune-related gene sets (INFLAMMATORY, COMPLEMENT) across 9 control, 10 sensitive, and 14 resistant KEP tumors. Data are represented as median $\pm$ IQR (box) and quartiles $\pm 1.5 \times$ IQR (whiskers), and one-way ANOVA and Tukey's post-hoc test were performed to compute adjusted P values. **(D)** Representative IHC staining of MYC and diverse immune cell markers across different KEP tumor groups. Scale bars, 50 μ m. **(E)** GSVA analysis for sensitive versus control tumors (left) and resistant versus control tumors (right) using RNA-seq data based on GOBP gene sets. GSVA scores were compared using the Student's *t* test. Red and blue dots, significantly (FDR < 0.05) up- and downregulated in sensitive tumors (left) and resistant tumors (right). Translation- and glutamine metabolism-associated gene sets are shown in the graph. **(F and G)** Expression ($\log_2(\text{CPM})$) of glutaminolysis enzymes *Gls* and *Gls2* across different KEP tumor groups (n = 9, 10, and 14 for control, sensitive, and resistant tumors, respectively; F) and KEP cell lines expressing *GFP* or *Myc* (two replica per group; G). In C and F, data are represented as median $\pm$ IQR (box) and quartiles $\pm 1.5 \times$ IQR (whiskers), and one-way ANOVA and Tukey's post-hoc test were performed to compute adjusted P values (**P < 0.01; ***P < 0.001; n.s, not significant).

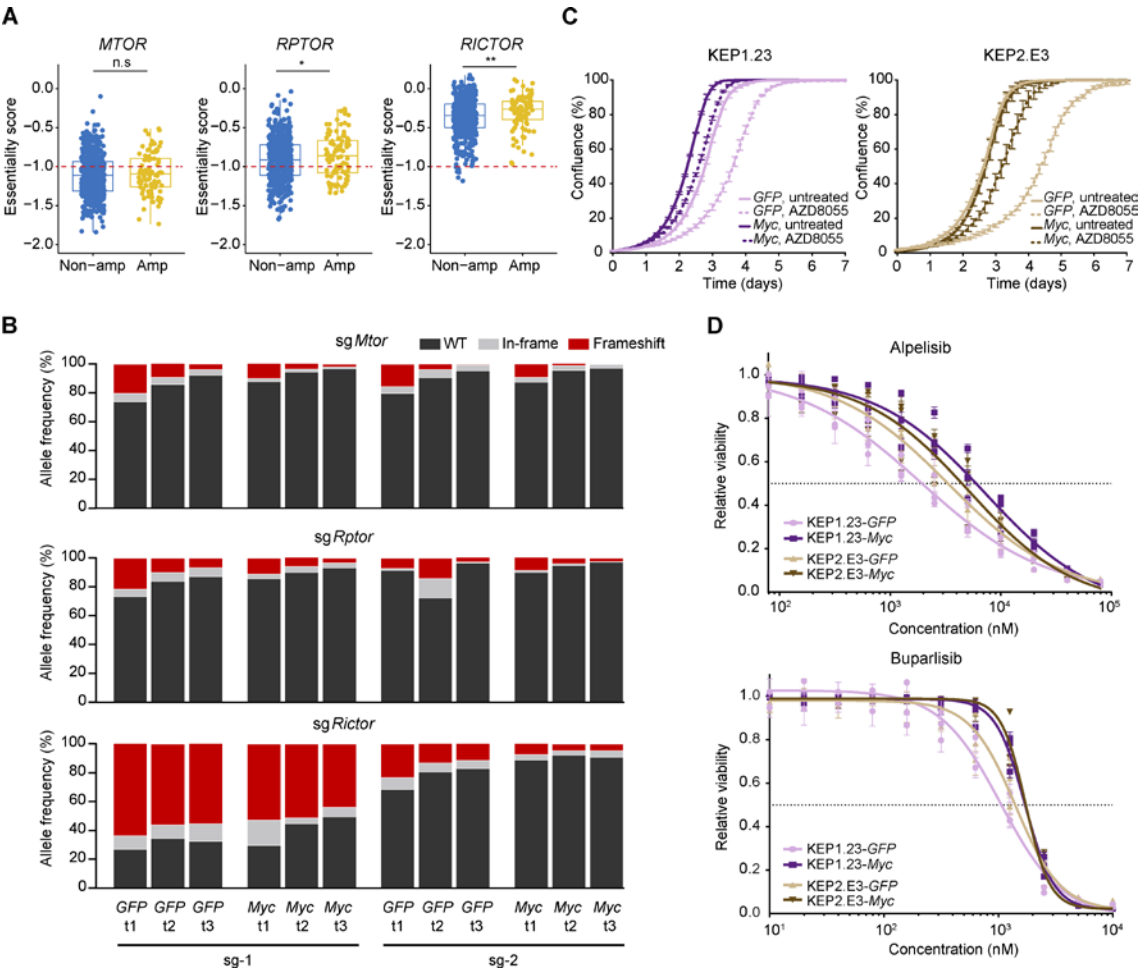


Figure S3. Associations of MYC status versus response to PI3K/mTOR inhibition.

(A) Comparison of *MTOR*, *RPTOR*, and *RICTOR* gene essentiality scores of cancer cell lines with and without *MYC* amplifications. Scores and copy number data were obtained from the DepMap (CRISPR, Public 22Q1) and the Cell Model Passports [45,107] database, respectively. Red dotted line (−1) indicates the average gene essentiality scores of known essential genes. Data are represented as median ± IQR (box) and quartiles ± 1.5 × IQR (whiskers) and P values were computed with one-tailed Student's *t* tests (**P* < 0.05; ***P* < 0.01; n.s., not significant). **(B)** KEP1.23 cells overexpressing *GFP* or *Myc* were transfected with lentiCRISPRv2-sgRNA-Cas9-Blast vectors containing the indicated sgRNAs targeting *Mtor*, *Rptor*, or *Rictor* and selected with blasticidin. At days 2 (t1), 4 (t2), and 6 (t3) after blasticidin selection, cells were lysed and allele frequencies for the indicated genes were determined using TIDE [46]. **(C)** Incucyte Live-Cell proliferation assay of KEP cell lines overexpressing *GFP* or *Myc* and treated with vehicle or 25 nM AZD8055 for 7 d. Data are represented as mean ± SEM of three replica per group of two independent experiments. **(D)** Dose–response curves of KEP cell lines overexpressing *GFP* or *MYC*, treated with alpelisib or buparlisib for 3 d, and measured by SRB colorimetric assay. Data are represented as mean ± standard deviation of five technical replicas per group of three independent experiments.

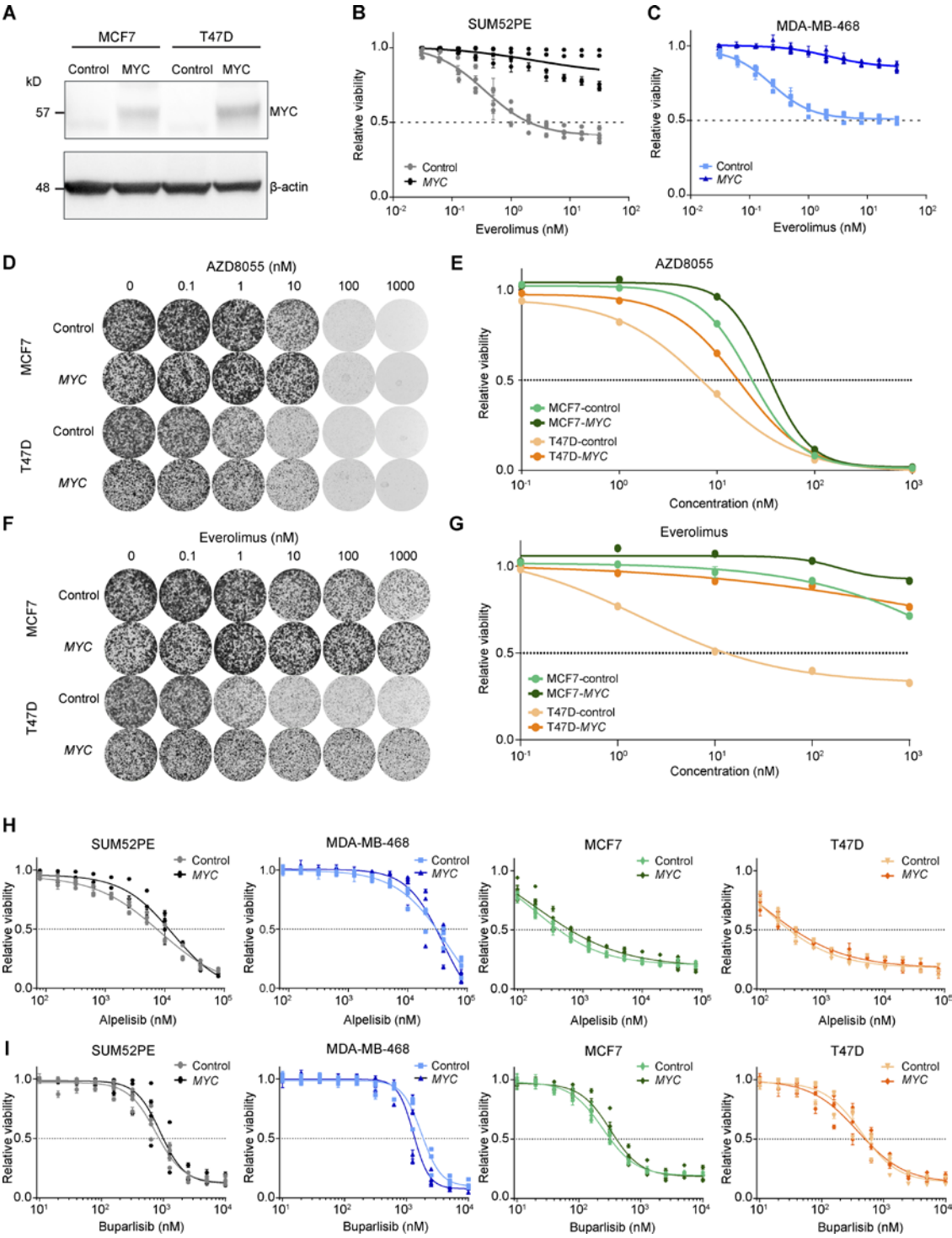


Figure S4. In vitro drug response assay for inhibitors of PI3K and mTOR in human breast cancer cell lines.

(A) Western blots for MYC and β -actin control on MCF7 and T47D cell lines transduced with control or MYC-overexpressing lentiviral vectors. Data represent one experiment. **(B and C)** Dose-response curves of SUM52PE (A) and MDA-MB-468 (B) cell lines transduced with control or MYC-overexpressing lentiviral vectors and treated with everolimus for 3 d. Cell viability was measured by SRB colorimetric assay. Data are represented as mean $\pm$ standard deviation of five technical replicas per group of $n \geq 3$ independent experiments. **(D–G)** Representative images (D and F) and corresponding quantification using CellTiter-Blue reagent (E and G) of long-term colony formation assays with MCF7 and T47D cells transduced with control or MYC-overexpressing lentiviral vectors and treated with AZD8055 (D and E) or everolimus (F and G). Cells were treated with indicated drug doses for 10 d, and values were normalized to untreated conditions within each cell line. Data in E and G are represented as mean $\pm$ standard deviation of three technical replicas per group of one representative experiment of two independent experiments. **(H and I)** Dose-response curves of T47D, MCF7, SUM52PE and MDA-MB-468 cell lines transduced with control or MYC-overexpressing lentiviral vectors, treated with alpelisib (H) or buparlisib (I) for 3 d, and measured by SRB colorimetric assay. Data are represented as mean $\pm$ standard deviation of five technical replica per group of $n \geq 3$ independent experiments.

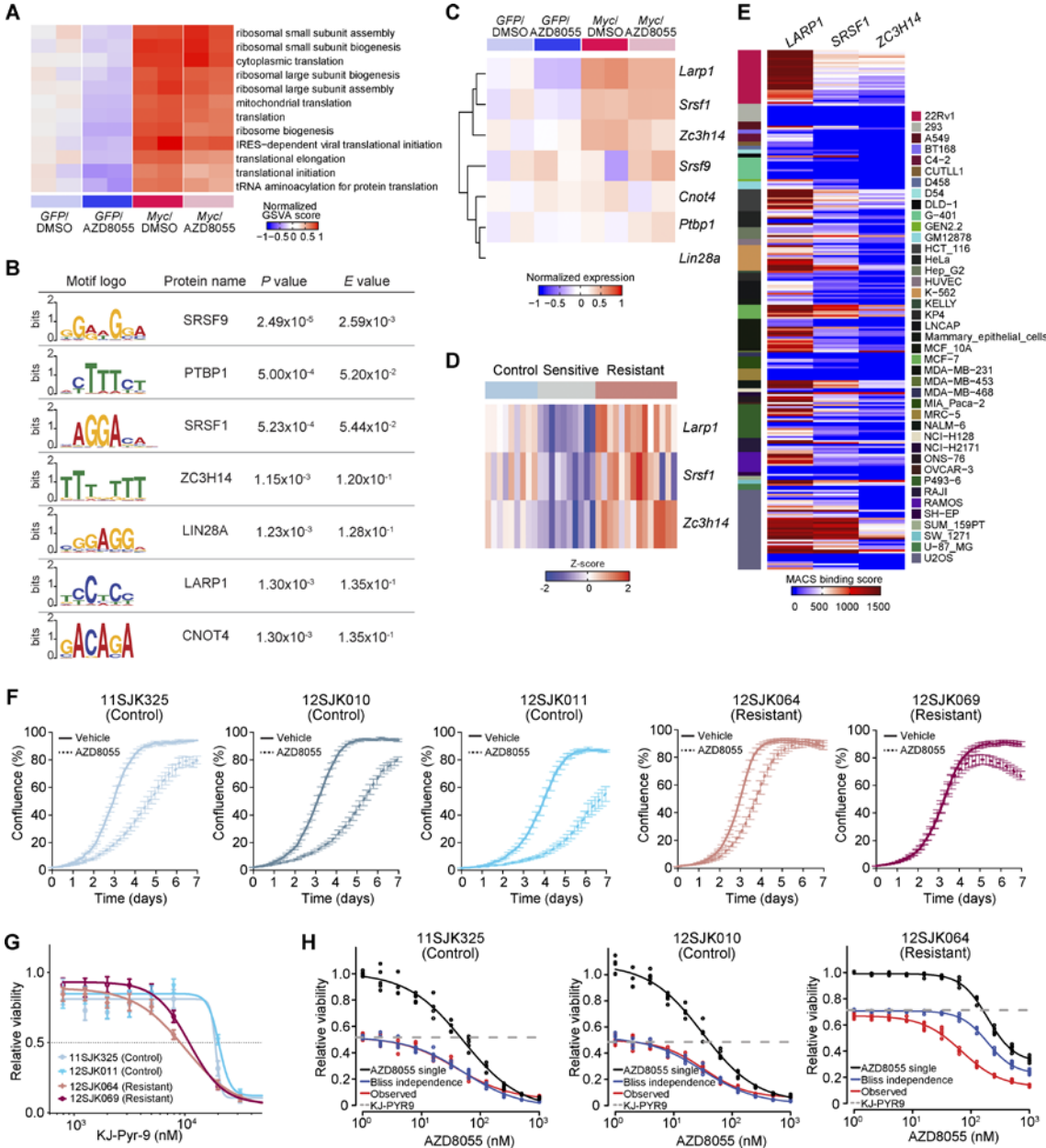


Figure S5. RNA-binding proteins associated with MYC-driven translation regulation.

(A) Normalized GSVA scores (centered on the mean of control *GFP*/DMSO samples) for translation and ribosome-associated processes in GOBPs computed based on the RNA-seq profiles of *GFP* and *Myc*-overexpressing KEP1.23 cells after DMSO or AZD8055 treatment (two replicas per condition). **(B)** Motifs of RNA-binding proteins that were significantly enriched in 5'-UTRs of transcripts in *Myc*-overexpressing KEP1.23 cells versus controls, as identified using the AME motif scanner [99]. Motif logos that were P value (Wilcoxon rank-sum test) adjusted by Bonferroni correction and E-value (adjusted P value multiplied by the number of motifs) were provided by AME. **(C)** Gene expression of RNA-binding proteins identified in A in *GFP* and *Myc*-overexpressing KEP1.23 cells after DMSO or AZD8055 treatment (two replicas per condition). Log₂(CPM) values were normalized by subtracting the mean values of *GFP*-DMSO samples. **(D)** Gene expression of the RNA-binding factors *Larp1*, *Srsf1*, and *Zc3h14* across 9 control, 10 sensitive, and 14 resistant KEP tumors. Log₂(CPM) values were transformed to Z-scores. **(E)** Model-based analysis of ChIP-seq (MACS) binding scores ($-10\log_{10}$ Q-value) in the promoter regions (± 5 KB from transcription start site) of *LARPI*, *SRSF1*, and *ZC3H14* obtained from the ChIP-Atlas database [55]. **(F)** Incucyte Live-Cell proliferation assay of control and resistant KEP cultures treated with vehicle or 90 nM AZD8055 for 7 d. Data are represented as mean $\pm$ SEM of five replicas per group of two independent experiments. **(G)** Dose-response curves of control and resistant KEP cultures, treated with KJ-Pyr-9 for 3 d, and measured by SRB colorimetric assay. Data are represented as mean $\pm$ standard deviation of five replicas per group of two independent experiments. **(H)** Dose-response curves of control and resistant KEP cultures treated with an AZD8055 range and 8 μ M KJ-Pyr-9 MYCi for 3 d, and measured by SRB colorimetric assay. Bliss independence model is shown to display an independent effect of KJ-Pyr-9 and AZD8055. Data are represented as mean $\pm$ standard deviation of five replicas per group of one experiment.



CHAPTER 7

GENERAL DISCUSSION

Precision cancer medicine requires precision data. To deliver the right drug(s) to the right patient at the right time, precision oncology must integrate several layers of information describing both the tumor and the individual patient to guide treatment selection. This includes genomic data (e.g., presence of somatic or germline mutations, copy-number changes, and structural rearrangements), other multiomic tumor profiling information (e.g., transcriptomics, epigenomics, proteomics, metabolomics, immune composition), presence of biomarkers predictive of response or resistance, longitudinal clinical characteristics (e.g., disease stage, histology, comorbidities, prior treatment), and complementary imaging and pathology data. Patient context such as demographics, lifestyle and environmental factors further influence treatment feasibility and effectiveness. Ultimately, the promise of precision oncology is realized when these diverse data streams are systematically integrated over time, enabling evidence-based therapeutic decisions that continuously learn from outcomes and refine care for current and future patients with similar characteristics.

Although precision oncology depends on the integration of multiple complementary data streams, genomic profiling remains the cornerstone of its current clinical implementation. The identification of somatic and germline alterations that drive tumorigenesis and create therapeutic vulnerabilities has enabled patient stratification and the development of targeted therapies that have led to meaningful improvements in clinical outcomes. In some cases, these advances have been transformative. Chronic myeloid leukemia (CML), once a fatal disease, has become a manageable chronic condition following the introduction of the tyrosine kinase inhibitor (TKI) imatinib (which targets the BCR-ABL1 fusion oncoprotein that drives CML) as first-line therapy in 2002. In solid tumors, outcome improvements have generally been more modest but nonetheless clinically significant, particularly among patients harboring actionable alterations in genes such as *HER2*, *EGFR*, *ALK*, *RET* and *NTRK*, who benefit from matched targeted therapies [1–5]. Together, these examples illustrate how linking specific genomic alterations to mechanism-based treatments can fundamentally alter the treatment of cancer. Consistent with this, biomarker-guided therapy is one of the strongest predictors of improved clinical outcomes in oncology [6–9].

Despite these successes, and as discussed in **Chapter 1**, the overall benefit of targeted therapies across cancer types remains limited. Indeed, all new anticancer agents approved by the European Medicines Agency (EMA) and the U.S. Food and Drug Administration (FDA) between 2003 and 2013 extended overall survival by an average of only 3.4 months [10]. The reasons underlying these limited gains are multifactorial [11–13]. However, a critical and increasingly evident challenge is the incomplete functional understanding of the numerous genetic alterations identified through routine tumor genomic profiling, many of which remain classified as variants of unknown significance (VUS). Determining whether such variants represent biologically relevant tumor-driving events or neutral alterations

with no impact on tumorigenesis or therapeutic response is essential, particularly in known cancer genes for which targeted therapies already exist. Addressing this knowledge gap is crucial to improving patient stratification and enabling more accurate and personalized treatment allocation.

1. DECODING ONCOGENIC SIGNALING AND THERAPY RESISTANCE: INSIGHTS FROM C-TERMINAL FGFR2/3 TRUNCATIONS AND MYC-DRIVEN MTOR INHIBITOR RESISTANCE

Genomic alterations in the *FGFR* family in cancer

As introduced in **Chapter 2** [14], genomic alterations affecting the *FGFR* family (*FGFR1*-*FGFR4*) are recurrent across many cancer types. In fact, up to 7.1% of all human cancers harbor alterations in the *FGFR* genes [15–19]. As tumor sequencing technologies and real-world genomic datasets expand, new *FGFR* alterations are continuously uncovered across multiple cancer types. However, the precise functional roles and therapeutic implications of many of these alterations remain unknown. It is crucial to define which *FGFR* alterations represent genuine oncogenic drivers with therapeutic relevance and which ones do not. While gene amplifications and hotspot mutations have long been recognized as oncogenic drivers, only recent oncogenomic and functional studies (including those presented in this thesis) have uncovered previously underappreciated *FGFR* rearrangements as a distinct and clinically relevant class of tumor-driving alterations that can dictate therapeutic response to *FGFR* inhibitors.

C-terminal truncation of *FGFR2*: an oncogenic driver and predictive biomarker

In **Chapter 3** [20], a Sleeping Beauty transposon-based insertional mutagenesis screen identified loss of exon 18 (E18) of *Fgfr2* (*Fgfr2*^{ΔE18}) – which encodes its C-terminal tail – as a driver alteration in mammary tumorigenesis in mice. Prompted by these findings, we systematically interrogated large-scale human oncogenomic datasets from the Hartwig Medical Foundation (HMF; >2,000 whole-genome sequencing (WGS) profiles with a subset of matched RNA-seq data) and Foundation Medicine (FMI; >249,000 targeted sequencing profiles) to determine the prevalence and spectrum of *FGFR2* alterations. Our analyses uncovered that human tumors express diverse *FGFR2*^{ΔE18} transcripts arising from multiple distinct alternative splicing events and genomic alterations. These include *FGFR2*-E1-E17 partial amplifications, canonical in-frame fusions with 3'-partner genes, non-canonical rearrangements, E18 splice-acceptor-site mutations and E18-truncating mutations. Notably, almost half of all alterations potentially perturbing E18 in the FMI cohort (44.6%, *n* = 610) were initially classified as VUS, with the remainder consisting of in-frame gene fusions.

We demonstrated that E18 truncation of *Fgfr2* constitutes a bona fide tumor driver alteration that is targetable by FGFR inhibitors by functionally interrogating *Fgfr2*^{ΔE18} variants using *in vitro* assays (i.e., cell outgrowth, downstream signaling, 3D colony formation, proteomics, phosphoproteomics and FGFR inhibitor sensitivity assays) together with somatic *in vivo* tumor modeling of mammary glands in mice. Importantly, we showed that this oncogenic activity is independent of fusion partners. To assess the clinical relevance of these findings, we re-analyzed the phase II FIGHT-202 clinical trial of the FGFR inhibitor pemigatinib in patients with advanced cholangiocarcinoma, which revealed that objective responses are largely restricted to tumors harboring fusions or rearrangements affecting the *FGFR*-I17/E18 hotspot region, with an objective response rate of 35.5% compared with patients harboring other or no *FGF/FGFR* alterations [21,22]. Further stratification showed that 80-100% of patients with *FGFR2* in-frame fusions or other non-canonical rearrangements achieved objective response or durable stable disease, irrespective of rearrangement type or *FGFR2* amplification status. These data suggest that *FGFR2* fusions and non-canonical rearrangements that involve E18 are predictive biomarkers for response to the FGFR inhibitor pemigatinib and, conceivably, for other FGFR-targeted therapies. Future studies in larger patient cohorts should confirm that alterations involving E18 reliably predict benefit from FGFR inhibitors across different cancer types.

Together, our findings establish C-terminal truncation of *FGFR2* as a recurrent, fusion partner-independent tumor driving alteration that is sensitive to FGFR inhibition. In doing so, we can re-classify many of the *FGFR2*-E18 alterations previously considered VUS as clinically actionable driver events, thereby expanding the population of patients who may benefit from FGFR-targeted therapies. Future clinical trials evaluating FGFR inhibitors should therefore include all patients with E18-truncating alterations, not only those with known hotspot mutations or known oncogenic E18-truncating fusions. More broadly, our work underscores the importance of functional characterization of somatic variants in cancer driver genes to guide patient selection for targeted therapies.

***FGFR2* amplification within a co-driver landscape**

One unexpected observation from our *in vivo* mammary tumorigenesis studies in **Chapter 3** [20] was that expression of full-length *Fgfr2* (*Fgfr2*^{FL}) – which mirrors *FGFR2* amplification (*FGFR2*^{amp}) in human cancers – showed surprisingly limited oncogenic competence. This is surprising given that *FGFR2* amplification has long been regarded as a tumor-driving event. Our findings revealed that its oncogenic potential depends on additional co-driver alterations that are tissue-specific. We demonstrated that *Fgfr2*^{FL} requires cooperative genetic co-alterations such as *Trp53* or *Pten* loss, or overexpression of *Myc*, *Ccnd1*, or *Fgf3* to promote tumorigenesis in the mammary gland of mice, consistent with the enrichment of these alterations in human breast cancers harboring *FGFR2*^{amp}. Such context dependence might also explain why *FGFR2*-amplified tumors show limited objective responses to FGFR inhibitors in clinical trials [23–25]. If tumor growth is sustained by additional oncogenic

drivers, inhibition of the FGFR signaling axis alone is unlikely to achieve durable control. In this setting, combinatorial therapeutic strategies are likely to yield more effective clinical outcomes. This is particularly relevant in advanced malignancies, which frequently harbor multiple driver events, with an average of 4-5 identifiable alterations in a single tumor [26,27].

In the context of *FGFR2^{amp}* tumors, combination therapy may not only improve response rates but also overcome resistance mechanisms that emerge under selective therapeutic pressure. For instance, it has been shown that *CCND1* amplification confers resistance to FGFR inhibitors [28], and early clinical data indicate that concurrent FGFR and CDK4/6 inhibition yields superior responses compared with monotherapy [29,30]. Moreover, FGFR inhibitors exhibit strong synergy with agents targeting PI3K [31–35] or EGFR/AXL [36]. Accordingly, ongoing clinical trials are evaluating FGFR-targeted therapies in combination with PI3K and CDK4/6 inhibitors, immune checkpoint inhibitors, chemotherapy and DNA damage response agents. Importantly, the potential additive or overlapping toxicities of these combinations must be carefully considered, as increased treatment-related adverse effects may limit dose intensity and long-term tolerability. One potential strategy to mitigate this challenge is the exploration of multiple low-dose combination regimens, which aim to maintain pathway suppression while reducing toxicity burdens [37]. Moving forward, a deeper mechanistic understanding of co-occurring driver alterations, resistance pathways, and treatment-associated toxicities will be essential to rationally design biomarker-guided combination therapies that maximize efficacy while maintaining an acceptable safety profile.

In contrast to *Fgfr2^{FL}*, *Fgfr2^{ΔE18}* acts as a potent oncogene capable of driving tumorigenesis independently of cooperating alterations. Consistent with this, *FGFR2^{ΔE18}* alterations in human cancers are largely mutually exclusive with *TP53* hotspot mutations, *PTEN* loss-of-function, and *MYC*, *CCND1*, or *FGF3/4/19* amplifications [20,38,39].

The oncogenicity of *FGFR2^{amp}* may be influenced not only by accompanying co-driver alterations but also by its capacity to generate *FGFR2^{ΔE18}* transcripts. Alternative splicing of *FGFR2* that generates shorter C-terminal isoforms is well documented and particularly linked to breast and gastric cancers [40–43]. Furthermore, cell lines that express the shorter C-terminal transcripts show more transforming potential [20,40,43].

FGFR3: a different story

FGFR2 and FGFR3 are two members of the FGFR family that share a highly conserved overall architecture and many key residues and functional domains, including the kinase catalytic core and docking sites for adaptor proteins [44]. We identified recurrent *FGFR3* breakpoints in both the FMI and HMF oncogenomic datasets that also result in C-terminal truncation. Given their structural and functional similarities, we reasoned that these *FGFR3* alterations

might mirror the oncogenic mechanism observed for *FGFR2* C-terminal truncations. We therefore hypothesized that loss of the *FGFR3* C-terminal tail (encoded by E18) would be sufficient to confer oncogenic activity. However, functional annotation of *FGFR3* alterations in **Chapter 5** [45] revealed a distinctly different biology. In contrast to *FGFR2*, C-terminal truncation alone is insufficient to drive transformation. Instead, *FGFR3* requires both C-terminal truncation and fusion to a partner gene that retains the expression of a functional dimerizing domain to effectively drive oncogenic signaling and tumorigenesis.

Nevertheless, not all *FGFR3* E18-truncating alterations occur in the context of a fusion with a dimerizing partner. In the FMI cohort analyzed in **Chapter 5** [45], E18-truncating mutations accounted for 12.3% (98/799) of all alterations predicted to generate a C-terminally truncated *FGFR3* lacking a fusion partner. Such an elevated percentage raises the question of whether C-terminally truncated *FGFR3* can confer oncogenic activity in combination with specific co-driver alterations. Our findings suggest that this might be the case. Tumors harboring *FGFR3*^{ΔE18} alterations devoid of a partner gene (i.e., E18-truncating mutations, intergenic or internal rearrangements, or partial amplifications) were enriched for alterations in *KRAS*, *ERBB2*, *MYC*, *PIK3CA* and *PTEN*, compared to tumors containing *FGFR3*^{ΔE18} fusions with a dimerizing partner. This supports a model in which, in the absence of a dimerizing partner, C-terminally truncated *FGFR3* may require additional tumor co-drivers to achieve full transforming potential.

C-terminal tail-mediated autoinhibition of *FGFR2* versus *FGFR3*

RTK signaling output is tightly controlled by autoinhibition [46]. Cells have evolved multiple regulatory mechanisms to prevent hyperactive FGFR signaling, and several studies have hinted at the importance of the C-terminal region in fine-tuning receptor activity [47,48]. Autoinhibitory functions have also been attributed to the C-terminal segments of other RTKs, highlighting a broader regulatory principle across the RTK superfamily. For example, the EGFR C-terminal tail stabilizes an unphosphorylated kinase domain dimer [49], whereas the MET C-tail limits kinase function by restricting ATP access to the active site [50], thereby suppressing receptor signaling. Consistent with this regulatory role, truncation of the C-terminal regions of EGFR, MET, HER2 and other RTKs promotes pro-tumorigenic activities [51–54].

In the FGFRs, phosphorylation of multiple tyrosine and serine residues within the C-terminal tail by downstream signaling components establishes negative feedback loops that promote receptor endocytosis and attenuate FGFR signaling [55–58]. Moreover, the distal portion of the C-terminal tail of FGFRs interacts with the inhibitory adaptor protein GRB2 [47,59–61], while its proximal part contains the clathrin-mediated endocytosis motif [47,62]. Together, these features position the FGFR C-terminal tail as a key negative regulatory module that restrains excessive receptor activation.

The divergence in the functional consequences of C-terminal truncation of *FGFR2* versus *FGFR3* raises the question of whether the C-terminal tails of these paralogs exert different autoinhibitory effects on their respective kinase domains. Loss of the *FGFR2* C-terminal tail represents a critical determinant of *FGFR2* oncogenicity (**Chapter 3** [20]) indicating that, under normal conditions, the intact tail functions to suppress excessive signaling. In **Chapter 4** [63], we systematically dissected the *FGFR2* C-terminal tail to identify the precise residues required for this suppressive function. Our findings revealed a previously unannotated phenylalanine-serine motif located in the proximal region of E18 to be essential for binding of the C-tail to the kinase domain and for restraining *FGFR2* kinase activity. Accordingly, we termed these two residues the ‘kinase domain-binding and suppression’ (KDBS) motif. Functional *in vitro* and *in vivo* efforts demonstrated that the C-terminal tail suppresses *FGFR2* oncogenicity through this KDBS motif and the concurrent regulation of endocytosis and the PI3K/ERK negative-feedback loop. Loss of either the entire C-tail or the KDBS motif activates PI3K-AKT-mTOR and STAT3 signaling and promotes transformation.

Notably, *FGFR3* is the only member of the *FGFR* family that does not fully conserve the KDBS motif. Instead of a phenylalanine-serine, *FGFR3* contains a phenylalanine-alanine sequence. This serine-to-alanine change could potentially diminish the capacity of the *FGFR3* C-terminal tail to exert autoinhibitory control over the kinase domain. The replacement of serine with alanine removes the polar hydroxyl group that could otherwise stabilize hydrogen bonds with the kinase domain, reduces favorable polar interactions that contribute to binding affinity, and may alter the local conformation or flexibility of the proximal C-terminal segment. In addition, the absence of a serine residue eliminates a potential phosphorylation site, thereby removing an extra layer of post-translational regulation. Collectively, these changes might weaken tail-kinase interactions and attenuate C-terminal tail-mediated suppression of receptor activity in *FGFR3*. As a result, *FGFR3* likely operates under comparatively lower baseline C-terminal-mediated autoinhibition than *FGFR2*, such that truncation or removal of the *FGFR3* C-terminal tail does not produce a substantial increase in signaling output. Without the release of a strong inhibitory constraint, *FGFR3*^{ΔE18} alone would be insufficient to drive oncogenic transformation and instead require an additional activating mechanism, such as fusion to a partner that promotes receptor dimerization, or other co-driver alterations (**Chapter 5** [45]).

Collectively, our findings from **Chapters 3-5** [20,45,63] underscore the importance of dissecting the biology of each *FGFR* paralog individually, rather than extrapolating functional or clinical implications of similar alterations across the receptor family. Our work demonstrates that even subtle structural differences between *FGFR2* and *FGFR3* can dictate distinct mechanisms of oncogenic activation. In addition, cellular and tissue context may further modulate the functional and therapeutic consequences of a given *FGFR* alteration.

Although C-terminal truncations of *FGFR2* and *FGFR3* occur across multiple cancer types, their distribution is not uniform, with certain tumor types showing relative enrichment. For instance, C-terminal truncating *FGFR2* alterations are more frequent in cholangiocarcinoma, gastroesophageal and breast cancers (**Chapter 3** [20]), whereas *FGFR3* alterations are more prevalent in bladder, glioblastoma and lung cancers (**Chapter 5** [45]). Together, these observations emphasize that the experimental models used to investigate *FGFR*-driven oncogenic mechanisms and therapeutic vulnerabilities should take tumor (sub)type into account, as the tissue of origin could shape the biological implications of *FGFR* alterations. More broadly, it remains to be investigated whether other RTKs beyond the *FGFR* paralogs contain phenylalanine-serine KDBS motifs and, if so, whether these motifs play a role in the autoinhibitory regulation of their kinase activity.

FGFR3*^{ΔE18} fusions beyond *TACC3

In **Chapter 5** [45], we observed a strong predominance of *TACC3* as the fusion partner in *FGFR3* E18-truncating (*FGFR3*^{ΔE18}) rearrangements, accounting for 85.3% of the 799 tumors analyzed in the FMI cohort. All other rearrangements were rare, with intergenic space rearrangements comprising the second largest category (3.3% of all cases) and partner genes other than *TACC3* occurring in fewer than 2% of cases. Although *FGFR3*^{ΔE18}-*TACC3* fusions were first reported as oncogenic drivers in glioblastoma in 2012 [64], to our knowledge we were the first to demonstrate their causal role in mammary tumorigenesis through *in vivo* somatic modeling in both wild-type and *Wap-Cre;Cdh1*^{F/F} mice, a model of invasive lobular carcinoma (ILC). Mechanistically, we showed that transformation depends strictly on *FGFR3* kinase activity, while the contribution of *TACC3* is largely structural: its coiled-coil domains provide the *FGFR3*^{ΔE18}-*TACC3* fusion protein with dimerization capacity necessary to promote constitutive receptor activation, whereas its canonical mitotic functions are dispensable. Consistent with this mechanism, tumors driven by *Fgfr3*^{ΔE18}-*Tacc3* remain sensitive to pharmacologic *FGFR* inhibition, with tumor growth significantly suppressed by the ATP-competitive *FGFR1-3* inhibitor fexagratinib [65], supporting *FGFR*-directed therapies as a rational treatment strategy for cancers harboring *FGFR3*^{ΔE18}-*TACC3* fusions. Importantly, expression of the truncated receptor alone is insufficient to drive tumor formation in the mouse mammary gland or lung, indicating that both C-terminal tail loss and a dimerizing partner are required for oncogenicity.

This behavior contrasts sharply with *FGFR2*, for which C-terminal truncation alone is sufficient to confer oncogenic potential and does not rely on a specific fusion partner (**Chapter 3** [20]). Indeed, more than a hundred distinct *FGFR2*^{ΔE18} fusion partners have been identified across tumor types, consistent with a mode of activation that is independent of fusion partner identity [19,20,66] (**Chapter 2** [14]). The striking predominance of *TACC3* among *FGFR3*^{ΔE18} rearrangements, together with the rarity of alternative partners, raises the possibility that some of these low-frequency fusions may represent passenger events rather than bona fide oncogenic drivers. With the exception of *FGFR3*^{ΔE18}-*BAIAP2L1*

[67], most of these variants remain functionally uncharacterized and their therapeutic relevance is therefore uncertain. To address this gap, in **Chapter 5** [45] we evaluated three previously untested, low-frequency *FGFR3^{ΔE18}* fusions identified in the FMI and HMF cohorts, namely *FGFR3^{ΔE18}-ADD1* (6 cases), *FGFR3^{ΔE18}-TNIP2* (11 cases) and *FGFR3^{ΔE18}-JAKMIP1* (4 cases). In all 21 cases, the fusion partners retained predicted self-interacting domains. All three *FGFR3* fusions proved oncogenic *in vitro*, activating downstream signaling in a kinase-dependent manner and conferring resistance to HER2/EGFR TKIs, yet remaining sensitive to pharmacologic FGFR blockade. This preserved susceptibility to FGFR inhibition underscores a potential therapeutic opportunity for patients harboring these rare *FGFR3* fusions. *In vivo*, however, their tumorigenic capacity diverged. Whereas *Fgfr3^{ΔE18}-Jakmip1* functioned as a robust tumor driver in the mouse mammary gland, *Fgfr3^{ΔE18}-Tnip2* produced an intermediate phenotype and *Fgfr3^{ΔE18}-Add1* failed to induce tumors. One possible explanation for the differences observed between our *in vitro* versus *in vivo* findings is tissue specificity, as *FGFR3^{ΔE18}-ADD1* and *FGFR3^{ΔE18}-TNIP2* fusions are mainly observed in urothelial carcinomas, whereas our functional modeling was performed in a breast cancer setting. Future research should investigate the *in vivo* oncogenicity of these fusions in the relevant tissue, i.e., urothelium, to more accurately define their oncogenic potential and their therapeutic relevance.

Taken together, our findings broaden the spectrum of actionable *FGFR3^{ΔE18}* rearrangements beyond *FGFR3^{ΔE18}-TACC3* fusions. Although each of these alternative fusions occurs at very low frequency, grouping them according to shared structural features – namely C-terminal truncation of *FGFR3* coupled with retention of a functional dimerization domain in the partner – rather than by the specific identity of the partner gene, substantially enlarges the pool of potentially targetable cases. This stratification suggests that more patients than previously appreciated might benefit from FGFR-targeted therapies (discussed more in detail in **Chapter 2** [14]). In clinical practice, however, the rarity of *FGFR3^{ΔE18}* gene fusions beyond *TACC3* renders conventional, histology-restricted trials largely impractical. Basket trials, which enroll patients based on the presence of a shared molecular alteration irrespective of tumor type, may provide a more feasible and efficient framework to evaluate therapeutic responses to FGFR-targeted therapies in this subgroup [68]. Such biomarker-driven strategies have already transformed precision oncology, leading to tissue-agnostic approvals of targeted therapies, including inhibitors directed against TRK and RET fusions, *BRAF^{V600E}*, and pembrolizumab for microsatellite instability-high (MSI-H) or tumor mutation burden-high (TMB-H) malignancies, following evidence of clinical benefit across diverse cancer types [3,69–74].

Overcoming treatment resistance

Resistance to targeted therapies remains one of the main limitations of precision oncology and a major cause of treatment failure and poor patient survival [9,10,75,76] (**Chapter 1**). Because tumors often harbor heterogeneous and complex genomic alterations, particularly

in advanced disease stages, monotherapies directed against a single driver rarely achieve durable disease control, since cancer cell plasticity and clonal diversity foster rapid selection and expansion of resistant clones. Resistance may be intrinsic, already present at diagnosis in the form of drug-resistant subclones, or it may emerge as a consequence of therapy. Acquired resistance is therapy-driven: although tumors may initially respond to treatment, they subsequently adapt under therapeutic pressure. This adaptation commonly involves target reactivation (e.g., on-target mutations or gene amplification), bypass signaling pathway activation, alterations in downstream effectors, or non-genetic mechanisms such as phenotypic switching, epigenetic remodeling, or remodeling of the tumor microenvironment [77]. Intratumoral heterogeneity further complicates therapeutic response, as distinct subclones may harbor different resistance mechanisms.

In the context of reversible ATP-competitive small-molecule FGFR inhibitors, acquired resistance most frequently involves secondary mutations in the gatekeeper or molecular brake residues within the FGFR kinase domain (FGFR1-FGFR3), although pathway reactivation via the upregulation or mutation of alternative RTKs or downstream signaling components is also commonly observed (**Chapter 2** [14]). Consistent with these mechanisms, our *in vivo* drug intervention study using the FGFR inhibitor fexagratinib revealed incomplete eradication of *Fgfr3^{ΔE18}-Tacc3*-driven tumors (**Chapter 5** [45]). A subset showed reactivation of FGFR3 downstream signaling during treatment outgrowth, whereas others appeared to develop FGFR3-independent resistance, potentially reflecting tumor cell plasticity or stromal adaptations [78–80].

Beyond FGFR-targeted therapy, alterations in the FGF/FGFR signaling axis have been implicated in resistance to chemotherapy, radiotherapy and other targeted agents, including CDK4/6 inhibitors, endocrine therapy and HER2-directed treatments [28,81–85]. Consistently, we demonstrated that expression of *Fgfr3^{ΔE18}-Tacc3* and the rare *Fgfr3^{ΔE18}-Add1*, *Fgfr3^{ΔE18}-Tnip2* and *Fgfr3^{ΔE18}-Jakmip1* fusions conferred marked resistance to the dual HER2/EGFR TKIs afatinib and dacomitinib in human bladder cancer cells, underscoring the capacity of oncogenic *FGFR3* rearrangements not only to act as primary drivers but also to reshape therapeutic vulnerabilities across different treatment modalities. Notably, these fusions retained sensitivity to pharmacologic FGFR inhibition, highlighting a potential opportunity for rational combination therapy. In fact, the use of rational combination strategies represents one of the most established approaches to overcome acquired resistance and improve clinical outcomes in precision oncology. This can be achieved through vertical pathway inhibition by targeting multiple nodes within the same signaling cascade to prevent downstream reactivation, as exemplified by combined BRAF and MEK inhibition in melanoma [86], or through horizontal blockade of compensatory bypass pathways, such as the combination of EGFR and MET inhibition in EGFR-mutant non-small cell lung cancer [87].

A context in which vertical pathway co-targeting may be particularly effective is the simultaneous inhibition of mTOR and MYC in breast cancer. MYC functions as a master regulator of multiple cellular processes, acting predominantly as a transcription factor that either directly or indirectly regulates the expression of thousands of genes [88]. It is deregulated in up to 70% of human cancers, and its activation contributes to the acquisition of most, if not all, hallmarks of cancer [89]. In **Chapter 6** [90], we used a transplantable mouse model of ILC, a breast cancer subtype characterized by frequent activation of the PI3K-AKT-mTOR signaling axis, to study the resistance mechanisms acquired during prolonged *in vivo* treatment with the mTOR inhibitor (mTORi) AZD8055. Multiomic analyses of matched therapy-naïve and AZD8055-resistant ILC tumors revealed that MYC activation represents a central feature of mTORi resistance. MYC coordinated both tumor cell-intrinsic and -extrinsic processes involved in response to mTOR inhibition, including the compensation of suppressed protein translation by mTORi. Orthotopic modeling of mouse ILC tumors and human breast cancer-derived xenografts further demonstrated that MYC is an *in vivo* driver of mTORi resistance. Importantly, while our experimental models established MYC as an acquired resistance mechanism, complementary analyses of human cancer cell lines and patient data showed that MYC amplification and elevated MYC expression prior to treatment correlated with poor response to mTORi. These observations indicate that MYC contributes not only to acquired resistance but also to intrinsic resistance. Collectively, this work highlights the potential of MYC not only as a therapeutic target but also as a predictive biomarker for patient stratification before treatment initiation.

From a therapeutic perspective, MYC co-targeting represents a rational strategy to overcome mTORi resistance and, potentially, resistance to other anticancer therapies, given that MYC amplification occurs across multiple cancer types and has been associated with reduced therapeutic sensitivity, underscoring its broad relevance in treatment resistance. Consistent with this rationale, in **Chapter 6** [90], pharmacologic MYC inhibition using the preclinical compound KJ-Pyr-9 restored sensitivity of Myc-amplified, mTOR inhibitor-resistant KEP tumor-derived cells to AZD8055, providing proof-of-concept for this combinatorial approach.

Despite its compelling rationale as a target, MYC has long been considered ‘undruggable’ due to its intrinsically disordered structure and lack of an enzymatic binding pocket, which complicate traditional small-molecule development. To date, no direct MYC inhibitors have been approved for clinical use. Nonetheless, several MYC-targeting compounds have entered early-phase clinical evaluation, and an expanding array of direct and indirect strategies is currently under active investigation [91,92]. Overall, there is growing optimism that effective pharmacologic targeting of this ‘undruggable’ target may be feasible after all [92]. The successful development of clinically effective MYC-directed therapies could have transformative implications across tumor types, including in combination strategies designed to prevent or overcome mTORi resistance.

Notably, rational combination therapy is only one of several approaches aimed at overcoming—or at least delaying—the emergence of therapeutic resistance. For instance, in the setting of resistance to reversible ATP-competitive FGFR inhibitors driven by kinase domain mutations in FGFR2, sequential treatment with covalent FGFR inhibitors has been shown to improve clinical outcomes [93]. These observations highlight the importance of developing more selective next-generation inhibitors as well as alternative FGFR-targeting modalities (**Chapter 2** [14]). In parallel, circulating tumor DNA (ctDNA) analysis offers a non-invasive strategy to detect emerging resistance mutations in real time, enabling dynamic treatment adjustments [94–96]. Furthermore, multiomic analysis of resistant tumors can uncover novel therapeutic vulnerabilities. Translating such discoveries into clinical benefit, however, requires sustained efforts in functional validation and therapeutic development.

Taken together, the findings presented in this thesis reinforce the notion that the central challenge of precision oncology lies not only in identifying actionable driver alterations, but also in anticipating resistance mechanisms that may pre-exist or arise during treatment. A deeper understanding of both intrinsic and acquired resistance is essential to extend the duration of therapeutic responses and, ultimately, to improve patient outcomes. Achieving this will require systematic functional validation in appropriate experimental models to distinguish causal resistance drivers from passenger events and to inform rational therapeutic strategies.

2. ADVANCING PRECISION ONCOLOGY BY INTEGRATING LARGE-SCALE GENOMICS, COMPUTATIONAL MODELING AND FUNCTIONAL ANNOTATION

The power of numbers and WGS

The identification of rare but therapeutically relevant alterations, such as the non-canonical *FGFR* rearrangements described in this thesis, underscores the need for large-scale sequencing efforts to reliably uncover low-frequency yet biologically meaningful variants that might otherwise remain overlooked. Rare driver mutations, resistance mechanisms, and complex structural variants often require extremely large cohorts—in some cases numbering in the hundreds of thousands of tumors—to reveal recurrent patterns indicative of functional relevance and to achieve statistical robustness. Major international initiatives, including The Cancer Genome Atlas (TCGA), the International Cancer Genome Consortium (ICGC), the Pan-Cancer Analysis of the Whole Genomes (PCAWG) and the 100,000 Genomes Project (100kGP), among others, have profoundly reshaped our understanding of cancer as a genetically driven disease [27,97–100]. In particular, the application of WGS within these programs has expanded the focus beyond protein-coding regions toward genome-wide

complexity, including structural variation, non-coding alterations, and mutational processes.

Routine clinical diagnostics continue to rely primarily on targeted gene panels or whole exome sequencing (WES), which efficiently detect single-nucleotide variants and small insertions or deletions within predefined cancer genes and capture most currently actionable biomarkers. However, these approaches provide limited resolution of structural rearrangements, copy-number alterations, mutational signatures, and alterations in non-coding regions that comprise 98% of the genome. In contrast, WGS enables unbiased detection across all variant classes, offering a more comprehensive molecular characterization in one single test [27,97,99–101]. Although its widespread clinical adoption remains limited by cost, analytical complexity and challenges in interpreting VUS, the use of WGS is steadily increasing in specific clinical contexts, particularly pediatric cancers, rare tumors, and cancers of unknown primary [102,103]. Of note, although short-read sequencing remains the predominant technology in clinical practice, it may fail to fully resolve complex structural rearrangements, including many *FGFR* fusions and other C-terminal truncations. The limited read length can hinder accurate reconstruction of breakpoint-spanning events and complex genomic alterations such as gene fusions. In contrast, the integration of long-read sequencing technologies could provide greater resolution of such rearrangements, allowing precise delineation of fusion breakpoints and overall structural architecture, and enabling more accurate variant annotation [104–106].

The added value of integrating genomic and transcriptomic profiling

The combined application of whole-genome and transcriptome sequencing (WGTS) has the potential to even further improve diagnostic precision by complementing unbiased DNA-level variant detection with RNA-based assessment of gene expression. This integrated approach has been shown to improve therapeutic stratification and facilitate identification of new biomarkers [103,107–113]. The added value of transcriptome profiling is particularly evident for gene rearrangements, for which WGS alone may not conclusively resolve reading-frame status or transcript structure. This was exemplified in our own analyses in **Chapters 3** [20] and **5** [45]. In **Chapter 5** [45], integration of matched RNA-seq with WGS data from the HMF cohort revealed that 85.7% (12/14) of cases initially classified as “frame-unknown” *FGFR3* E18-truncating rearrangements based on DNA sequencing alone expressed detectable in-frame rearrangement transcripts, underscoring the importance of integrating DNA- and RNA-based sequencing to reliably identify oncogenic rearrangements.

WGTS has the potential to ultimately replace many conventional molecular diagnostic approaches, although its widespread integration into routine oncology practice will require healthcare systems to address key operational challenges. These include

standardized procedures for sample collection and biobanking, robust analytical pipelines, clear frameworks for data interpretation and reporting, and aligned regulatory and reimbursement landscapes [103]. Ultimately, the implementation of either WGS or WGTS into routine clinical care will depend on demonstrable cost-effectiveness. Although sequencing costs have declined considerably over the past decade, comprehensive genomic and transcriptomic profiling remains expensive in clinical settings [103,109,114–116]. Nevertheless, broader adoption may prove economically justified in the long term if these approaches provide the most comprehensive analysis of tumor alterations resulting in improved diagnostic accuracy, reduced multiple redundant testing and more precise therapeutic stratification.

Knowledge resources and data infrastructure in precision oncology

The vast amounts of genomic data generated by next-generation sequencing (NGS) technologies are of limited value without robust computational and statistical methods that enable the extraction of meaningful biological insights from raw sequencing data. Over the past 25 years, numerous computational tools have been developed to annotate both the human genome in general [117] and cancer genomes in particular [118]. In oncology, large-scale consortia such as the ICGC/TCGA PCAWG project have driven the establishment of standardized bioinformatics pipelines that now serve as reference workflows for tumor genomic analyses [27,119]. In parallel, open-access repositories and visualization platforms, such as TCGA, the Genomic Data Commons (GDC), cBioPortal and the UCSC Genome Browser, have markedly accelerated translational cancer research by facilitating large-scale data sharing and integrative omics analyses [118]. Additionally, these resources underpin the development of expert-curated ‘knowledge bases’ that aggregate and annotate the clinical significance of somatic variants to support their interpretation in the clinical setting. Prominent examples include ClinVar [120], COSMIC [121], CIViC [122] and OncoKB [123,124].

The ultimate goal is to comprehensively integrate genomic alterations with functional evidence and real-world clinical data within unified frameworks that support outcome prediction. Such integration would not only facilitate more informed, tailored treatment decisions by oncologists but also contribute to a learning healthcare system in which insights derived from past and current patients continuously refine care for future patients [102]. One representative example of progress towards this vision is Project GENIE (Genomics Evidence Neoplasia Information Exchange), an international data-sharing initiative led by the American Association for Cancer Research (AACR) [125,126]. To date, this publicly accessible clinico-genomic registry has linked harmonized NGS data with demographic, diagnostic, treatment, and longitudinal outcome information from more than 220,000 patients treated across 20 leading cancer centers worldwide. GENIE has underscored the critical importance of national and international data sharing, collaborative research infrastructures, and secure, large-scale and unified frameworks that enable cross-border access to genomic and clinical patient data [127].

Artificial intelligence in precision oncology

As molecular and clinical datasets continue to expand in scale and complexity, more powerful analytical tools are required. Optimized therapy will likely depend on rational combination therapies that target tumors from different angles, further increasing the analytical demands on precision oncology. In this context, the incorporation of machine learning and artificial intelligence (AI) into diagnostic and decision-support workflows becomes increasingly necessary [128,129]. AI-based approaches are particularly well suited for handling large, multimodal data and identifying subtle patterns and interactions that may not be spotted by conventional methods or the expert human eye [130–134]. Consequently, these approaches hold considerable promise for generating clinically meaningful outcome predictions.

At the same time, methodological rigor is required to ensure generalizability, interpretability and clinical reliability. Algorithm optimization, external validation, and prospective evaluation remain indispensable prerequisites for clinical implementation [128]. Moreover, the integration of AI into precision oncology must proceed transparently and responsibly, with careful consideration of ethical challenges including data privacy and security, informed consent, management of incidental findings, responsible data sharing, and the prevention of genetic data misuse [118,129].

Functional validation as the gold standard

Despite major advances in computational modeling, machine learning and AI, experimental functional characterization remains the gold standard for establishing the biological and clinical relevance of genomic alterations. *In silico* prediction frameworks can provide important speed, scalability, prioritization, and hypothesis-generating capacity, but they cannot substitute for direct experimental evidence. For instance, Tangermann *et al.* [135], who performed saturation mutational scanning of all 11,520 possible kinase domain point mutations across *FGFR1-4* to systematically identify activating and FGFR inhibitor resistance-mediating variants, demonstrated that widely used *in silico* prediction tools such as REVEL, EVE and AlphaMissense failed to reliably classify activating gain-of-function mutations. Moreover, the algorithms could not distinguish between gain- and loss-of-function effects. Tumor recurrence in COSMIC was likewise insufficient as a surrogate for activation. In another example, Ng *et al.* [136] developed a moderate-throughput functional genomic platform and used it to annotate more than a 1,000 VUS—including gene amplifications, point mutations, indels and gene fusions—potentially doubling the number of driver mutations characterized in clinically actionable genes. When the authors compared their experimental annotations against 21 commonly used algorithms that predict mutation effects, their accuracy was variable (50–76%) and none fully recapitulated the experimental classification. Together, these findings further illustrate that computational predictions cannot substitute for direct functional validation to establish causality.

Some machine learning-based methodologies have been described to outperform selected experimental efforts, such as the *in silico* saturation mutagenesis framework boostDM and its capacity to score potential driver mutations in certain cancer genes such as *TP53* or *EGFR* based on patterns derived from tumor sequencing datasets [137]. However, the performance of boostDM is inherently dependent on the availability of extensive, gene-specific training data that are not available for many targets. For example, the boostDM framework provides coverage for *FGFR3* based on mutational data from bladder cancer, but offers no predictions for *FGFR1*, *FGFR2* or *FGFR4*. Thus, even advanced AI-driven approaches remain constrained by data availability and generalizability, and cannot yet replace systematic experimental interrogation.

In this thesis, computational analyses played an essential role in interrogating diverse large-scale datasets derived from cancer patients, mouse models and *in vitro* experimental systems to prioritize candidate alterations. However, rigorous functional *in vitro* and *in vivo* interrogation was required to conclusively demonstrate that (i) C-terminal truncation of *FGFR2* is a strong oncogenic driver in cancer (**Chapter 3** [20]); (ii) this oncogenic activity is mechanistically driven by loss of the KDBS motif (**Chapter 4** [63]); (iii) the oncogenic potential of *FGFR3* requires both C-terminal truncation and fusion to a partner with a conserved dimerizing capacity (**Chapter 5** [45]); and (iv) *MYC* activation drives resistance to mTOR inhibitors in breast cancer (**Chapter 6** [90]). Therefore, while computational approaches are indispensable for navigating molecular complexity at scale, experimental validation remains essential to distinguish between causation and mere correlation, and to define true drivers of oncogenicity and therapy resistance.

3. CONCLUSIONS

The work presented in this thesis underscores that systematic functional characterization is essential for translating genomic findings into clinically actionable biomarkers and enabling more precise patient selection for targeted therapies. By demonstrating the divergent oncogenic behavior of the *FGFR2* and *FGFR3* paralogs following C-terminal truncation, this work provides a strong rationale for expanding *FGFR*-targeted treatment strategies to patients harboring E18-truncating *FGFR2* alterations or *FGFR3* fusions involving dimerizing partner genes. In parallel, the identification of *MYC* as a clinically relevant driver of resistance to mTOR inhibition in breast cancer further refines patient stratification, highlighting high *MYC* expression as a potential biomarker to identify patients less likely to benefit from mTOR-targeted therapy. Together, these findings illustrate how functional interrogation can refine therapeutic stratification and support more informed therapeutic decisions in precision oncology.

Despite remarkable advances over the past two decades, precision oncology remains a developing field. The fact that driver alterations in approximately 5% of cancers remain unidentified [27] underscores that our understanding of tumor biology is still incomplete and that important discoveries remain ahead. Moving forward, precision oncology must extend beyond genomics alone, since genomic alterations represent only the first layer of tumor complexity. Integrated pan-omic approaches combining genomic, transcriptomic, proteomic, metabolomic and immunogenomic data will be essential to better define the full complexity of individual tumors. At the same time, rational biomarker-guided combination strategies should increasingly replace one-size-fits-all monotherapies, particularly as cancers adapt under therapeutic pressure and develop resistance.

Ultimately, translating molecular insight into sustained clinical benefit will require the coordinated integration of large-scale omics datasets, rigorous functional validation, prospectively collected real-world clinical outcome data, and advanced computational and AI-enabled analytical frameworks. Only through this truly multidimensional and translational ecosystem, coupled with continued innovation in drug discovery and development, can genomic insights be reliably transformed into tangible, durable improvements in patient outcomes.

REFERENCES

1. Piccart-Gebhart, M.J. *et al.* (2005) Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *N Engl J Med* 353, 1659–1672
2. Soria, J.-C. *et al.* (2018) Osimertinib in Untreated EGFR-Mutated Advanced Non-Small-Cell Lung Cancer. *N Engl J Med* 378, 113–125
3. Drilon, A. *et al.* (2018) Efficacy of Larotrectinib in TRK Fusion-Positive Cancers in Adults and Children. *N Engl J Med* 378, 731–739
4. Drilon, A. *et al.* (2020) Efficacy of Selpercatinib in RET Fusion-Positive Non-Small-Cell Lung Cancer. *N Engl J Med* 383, 813–824
5. Shaw, A.T. *et al.* (2020) First-Line Lorlatinib or Crizotinib in Advanced ALK-Positive Lung Cancer. *N Engl J Med* 383, 2018–2029
6. Jardim, D.L. *et al.* (2015) Impact of a Biomarker-Based Strategy on Oncology Drug Development: A Meta-analysis of Clinical Trials Leading to FDA Approval. *J Natl Cancer Inst* 107, djv253
7. Schwaederle, M. *et al.* (2015) Impact of Precision Medicine in Diverse Cancers: A Meta-Analysis of Phase II Clinical Trials. *J Clin Oncol* 33, 3817–3825
8. Schwaederle, M. *et al.* (2016) Precision Oncology: The UC San Diego Moores Cancer Center PREDICT Experience. *Mol Cancer Ther* 15, 743–752
9. Schwaederle, M. *et al.* (2016) Association of Biomarker-Based Treatment Strategies With Response Rates and Progression-Free Survival in Refractory Malignant Neoplasms: A Meta-analysis. *JAMA Oncol* 2, 1452–1459
10. Salas-Vega, S. *et al.* (2017) Assessment of Overall Survival, Quality of Life, and Safety Benefits Associated With New Cancer Medicines. *JAMA Oncol* 3, 382–390
11. Prasad, V. *et al.* (2015) The Strength of Association Between Surrogate End Points and Survival in Oncology: A Systematic Review of Trial-Level Meta-analyses. *JAMA Intern Med* 175, 1389–1398
12. Subbiah, V. and Kurzrock, R. (2018) Challenging Standard-of-Care Paradigms in the Precision Oncology Era. *Trends Cancer* 4, 101–109
13. Wahida, A. *et al.* (2023) The coming decade in precision oncology: six riddles. *Nat Rev Cancer* 23, 43–54
14. Yemelyanenko, J. *et al.* (2026) FGFR rearrangements: oncogenic drivers and therapeutic targets. *Trends in Cancer* DOI: 10.1016/j.trecan.2026.01.010
15. Helsten, T. *et al.* (2016) The FGFR landscape in cancer: Analysis of 4,853 tumors by next-generation sequencing. *Clinical Cancer Research* 22, 259–267
16. Krook, M.A. *et al.* (2020) Pan-cancer analysis of FGFR1-3 genomic alterations to reveal a complex molecular landscape. *JCO* 38, 3620–3620
17. Zuo, W. *et al.* (2020) Landscape of FGF/FGFR Alterations in 12,372 Chinese Cancer Patients. *Journal of Cancer* 11, 6695–6699
18. Sun, Y. *et al.* (2020) A comprehensive pan-cancer study of fibroblast growth factor receptor aberrations in Chinese cancer patients. *Annals of Translational Medicine* 8, 1290–1290
19. Murugesan, K. *et al.* (2022) Pan-tumor landscape of fibroblast growth factor receptor 1-4 genomic alterations. *ESMO Open* 7, 100641
20. Zingg, D. *et al.* (2022) Truncated FGFR2 is a clinically actionable oncogene in multiple cancers. *Nature* 608, 609–617
21. Abou-Alfa, G.K. *et al.* (2020) Pemigatinib for previously treated, locally advanced or metastatic cholangiocarcinoma: a multicentre, open-label, phase 2 study. *Lancet Oncol* 21, 671–684
22. Rodón, J. *et al.* (2024) Pemigatinib in previously treated solid tumors with activating FGFR1–FGFR3 alterations: phase 2 FIGHT-207 basket trial. *Nat Med* DOI: 10.1038/s41591-024-02934-7

23. Pearson, A. *et al.* (2016) High-Level Clonal FGFR Amplification and Response to FGFR Inhibition in a Translational Clinical Trial. *Cancer Discov* 6, 838–851
24. Van Cutsem, E. *et al.* (2017) A randomized, open-label study of the efficacy and safety of AZD4547 monotherapy versus paclitaxel for the treatment of advanced gastric adenocarcinoma with FGFR2 polysomy or gene amplification. *Ann Oncol* 28, 1316–1324
25. Chae, Y.K. *et al.* (2020) Phase II Study of AZD4547 in Patients With Tumors Harboring Aberrations in the FGFR Pathway: Results From the NCI-MATCH Trial (EAY131) Subprotocol W. *J Clin Oncol* 38, 2407–2417
26. Martincorena, I. *et al.* (2017) Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell* 171, 1029–1041.e21
27. Aaltonen, L.A. *et al.* (2020) Pan-cancer analysis of whole genomes. *Nature* 578, 82–93
28. Formisano, L. *et al.* (2019) Aberrant FGFR signaling mediates resistance to CDK4/6 inhibitors in ER+ breast cancer. *Nat Commun* 10, 1373
29. Uehara, Y. *et al.* (2022) Targeting the FGF/FGFR axis and its co-alteration allies. *ESMO Open* 7, 100647
30. Liu, K. *et al.* (2025) Significant Response to Palbociclib Plus Lenvatinib as Second-line Treatment for CDKN2A/2B Deletion Intrahepatic Cholangiocarcinoma: A Case Report. *J Clin Transl Hepatol* 13, 169–172
31. Packer, L.M. *et al.* (2017) PI3K Inhibitors Synergize with FGFR Inhibitors to Enhance Antitumor Responses in FGFR2mutant Endometrial Cancers. *Mol Cancer Ther* 16, 637–648
32. Holzhauser, S. *et al.* (2019) In vitro antitumor effects of FGFR and PI3K inhibitors on human papillomavirus positive and negative tonsillar and base of tongue cancer cell lines. *Oncol Lett* 18, 6249–6260
33. Holzhauser, S. *et al.* (2020) Targeting Fibroblast Growth Factor Receptor (FGFR) and Phosphoinositide 3-kinase (PI3K) Signaling Pathways in Medulloblastoma Cell Lines. *Anticancer Research* 40, 53–66
34. Lukoseviciute, M. *et al.* (2024) Combined targeted therapy with PI3K and CDK4/6, or FGFR inhibitors show synergistic effects in a neuroblastoma spheroid culture model. *Biomed Pharmacother* 177, 116993
35. Juric, D. *et al.* (2025) PI3Ka inhibitor and degrader inavolisib can co-opt FGFR2 to enhance responses in patients with PIK3CA-mutated solid tumors and in preclinical models. *Clin Cancer Res* DOI: 10.1158/1078-0432.CCR-25-1459
36. Nakamura, R. *et al.* (2024) Triple combination therapy comprising osimertinib, an AXL inhibitor, and an FGFR inhibitor improves the efficacy of EGFR-mutated non-small cell lung cancer. *Cancer Letters* 598, 217124
37. Fernandes Neto, J.M. *et al.* (2020) Multiple low dose therapy as an effective strategy to treat EGFR inhibitor-resistant NSCLC tumours. *Nat Commun* 11, 3157
38. Silverman, I.M. *et al.* (2021) Clinicogenomic Analysis of FGFR2-Rearranged Cholangiocarcinoma Identifies Correlates of Response and Mechanisms of Resistance to Pemigatinib. *Cancer Discov* 11, 326–339
39. D'Souza, N. *et al.* (2024) Computational patterns in FGFR3-altered non-muscle invasive bladder cancer influence recurrence and progression. *Urologic Oncology: Seminars and Original Investigations* 42, S55
40. Itoh, H. *et al.* (1994) Preferential alternative splicing in cancer generates a K-sam messenger RNA with higher transforming activity. *Cancer Res* 54, 3237–3241
41. Ueda, T. *et al.* (1999) Deletion of the carboxyl-terminal exons of K-sam/FGFR2 by short homology-mediated recombination, generating preferential expression of specific messenger RNAs. *Cancer Res* 59, 6080–6086

42. Tannheimer, S.L. *et al.* (2000) Characterization of fibroblast growth factor receptor 2 overexpression in the human breast cancer cell line SUM-52PE. *Breast Cancer Res* 2, 311–320
43. Moffa, A.B. *et al.* (2004) Transforming Potential of Alternatively Spliced Variants of Fibroblast Growth Factor Receptor 2 in Human Mammary Epithelial Cells¹. *Mol Cancer Res* 2, 643–652
44. Dai, S. *et al.* (2019) Fibroblast Growth Factor Receptors (FGFRs): Structures and Small Molecule Inhibitors. *Cells* 8, 614
45. Yemelyanenko, J. *et al.* (2025) C-terminal Truncation and Fusion Partner Determine Oncogenicity of FGFR3. *Cancer Res* DOI: 10.1158/0008-5472.CAN-24-2648
46. Lemmon, M.A. and Schlessinger, J. (2010) Cell signaling by receptor-tyrosine kinases. *Cell* 141, 1117–1134
47. Szybowska, P. *et al.* (2021) Negative Regulation of FGFR (Fibroblast Growth Factor Receptor) Signaling. *Cells* 10, 1342
48. Lin, C.-C. *et al.* (2023) The combined action of the intracellular regions regulates FGFR2 kinase activity. *Commun Biol* 6, 1–14
49. Zhang, X. *et al.* (2006) An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell* 125, 1137–1149
50. Schiering, N. *et al.* (2003) Crystal structure of the tyrosine kinase domain of the hepatocyte growth factor receptor c-Met and its complex with the microbial alkaloid K-252a. *Proc Natl Acad Sci U S A* 100, 12654–12659
51. Khazaie, K. *et al.* (1988) Truncation of the human EGF receptor leads to differential transforming potentials in primary avian fibroblasts and erythroblasts. *The EMBO Journal* 7, 3061–3071
52. Woolford, J. *et al.* (1988) Activation of the feline c-fms proto-oncogene: multiple alterations are required to generate a fully transformed phenotype. *Cell* 55, 965–977
53. Akiyama, T. *et al.* (1991) The transforming potential of the c-erbB-2 protein is regulated by its autophosphorylation at the carboxyl-terminal domain. *Mol Cell Biol* 11, 833–842
54. Bardelli, A. *et al.* (1999) A peptide representing the carboxyl-terminal tail of the met receptor inhibits kinase activity and invasive growth. *J Biol Chem* 274, 29274–29281
55. Lorenzi, M.V. *et al.* (1997) Ligand-independent activation of fibroblast growth factor receptor-2 by carboxyl terminal alterations. *Oncogene* 15, 817–826
56. Hart, K.C. *et al.* (2001) Identification of tyrosine residues in constitutively activated fibroblast growth factor receptor 3 involved in mitogenesis, Stat activation, and phosphatidylinositol 3-kinase activation. *Mol Biol Cell* 12, 931–942
57. Nadratowska-Wesolowska, B. *et al.* (2014) RSK2 regulates endocytosis of FGF receptor 1 by phosphorylation on serine 789. *Oncogene* 33, 4823–4836
58. Szybowska, P. *et al.* (2019) Cancer Mutations in FGFR2 Prevent a Negative Feedback Loop Mediated by the ERK1/2 Pathway. *Cells* 8, 518
59. Ahmed, Z. *et al.* (2010) Direct binding of Grb2 SH3 domain to FGFR2 regulates SHP2 function. *Cell Signal* 22, 23–33
60. Ahmed, Z. *et al.* (2013) Grb2 controls phosphorylation of FGFR2 by inhibiting receptor kinase and Shp2 phosphatase activity. *J Cell Biol* 200, 493–504
61. Lin, C.-C. *et al.* (2012) Inhibition of basal FGF receptor signaling by dimeric Grb2. *Cell* 149, 1514–1524
62. Cha, J.Y. *et al.* (2009) Aberrant receptor internalization and enhanced FRS2-dependent signaling contribute to the transforming activity of the fibroblast growth factor receptor 2 IIIb C3 isoform. *J Biol Chem* 284, 6227–6240
63. Zingg, D. *et al.* (2025) The C-terminal Kinase Domain-Binding and Suppression Motif Prevents Constitutive Activation of FGFR2. *Cancer Research* DOI: 10.1158/0008-5472.CAN-24-3349
64. Singh, D. *et al.* (2012) Transforming fusions of FGFR and TACC genes in human glioblastoma. *Science* 337, 1231–1235

65. Gavine, P.R. *et al.* (2012) AZD4547: an orally bioavailable, potent, and selective inhibitor of the fibroblast growth factor receptor tyrosine kinase family. *Cancer Res* 72, 2045–2056
66. Chen, L. *et al.* (2021) Fibroblast growth factor receptor fusions in cancer: opportunities and challenges. *Journal of Experimental & Clinical Cancer Research* 40, 345
67. Nakanishi, Y. *et al.* (2015) Mechanism of oncogenic signal activation by the novel fusion kinase FGFR3-BAIAP2L1. *Molecular Cancer Therapeutics* 14
68. Tao, J.J. *et al.* (2018) Basket Studies: Redefining Clinical Trials in the Era of Genome-Driven Oncology. *Annu Rev Med* 69, 319–331
69. Le, D.T. *et al.* (2017) Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 357, 409–413
70. Subbiah, V. *et al.* (2020) Dabrafenib plus trametinib in patients with BRAFV600E-mutated biliary tract cancer (ROAR): a phase 2, open-label, single-arm, multicentre basket trial. *Lancet Oncol* 21, 1234–1243
71. Subbiah, V. *et al.* (2022) Tumour-agnostic efficacy and safety of selpercatinib in patients with RET fusion-positive solid tumours other than lung or thyroid tumours (LIBRETTO-001): a phase 1/2, open-label, basket trial. *Lancet Oncol* 23, 1261–1273
72. Marabelle, A. *et al.* (2020) Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: prospective biomarker analysis of the multicohort, open-label, phase 2 KEYNOTE-158 study. *Lancet Oncol* 21, 1353–1365
73. Salama, A.K.S. *et al.* (2020) Dabrafenib and Trametinib in Patients With Tumors With BRAFV600E Mutations: Results of the NCI-MATCH Trial Subprotocol H. *J Clin Oncol* 38, 3895–3904
74. Wen, P.Y. *et al.* (2022) Dabrafenib plus trametinib in patients with BRAFV600E-mutant low-grade and high-grade glioma (ROAR): a multicentre, open-label, single-arm, phase 2, basket trial. *Lancet Oncol* 23, 53–64
75. Kim, C. and Prasad, V. (2015) Cancer Drugs Approved on the Basis of a Surrogate End Point and Subsequent Overall Survival: An Analysis of 5 Years of US Food and Drug Administration Approvals. *JAMA Intern Med* 175, 1992–1994
76. Davis, C. *et al.* (2017) Availability of evidence of benefits on overall survival and quality of life of cancer drugs approved by European Medicines Agency: retrospective cohort study of drug approvals 2009–13. *BMJ* 359, j4530
77. Sabnis, A.J. and Bivona, T.G. (2019) Principles of resistance to targeted cancer therapy: lessons from basic and translational cancer biology. *Trends Mol Med* 25, 185–197
78. Kas, S.M. *et al.* (2018) Transcriptomics and transposon mutagenesis identify multiple mechanisms of resistance to the FGFR inhibitor AZD4547. *Cancer Research* 78, 5668–5679
79. Katoh, M. (2019) Fibroblast growth factor receptors as treatment targets in clinical oncology. *Nat Rev Clin Oncol* 16, 105–122
80. Katoh, M. *et al.* (2024) FGFR-targeted therapeutics: clinical activity, mechanisms of resistance and new directions. *Nat Rev Clin Oncol* DOI: 10.1038/s41571-024-00869-z
81. Turner, N. *et al.* (2010) FGFR1 amplification drives endocrine therapy resistance and is a therapeutic target in breast cancer. *Cancer Res* 70, 2085–2094
82. Hanker, A.B. *et al.* (2017) HER2-Overexpressing Breast Cancers Amplify FGFR Signaling upon Acquisition of Resistance to Dual Therapeutic Blockade of HER2. *Clin Cancer Res* 23, 4323–4334
83. Zhou, Y. *et al.* (2020) FGF/FGFR signaling pathway involved resistance in various cancer types. *J Cancer* 11, 2000–2007
84. Mao, P. *et al.* (2020) Acquired FGFR and FGF Alterations Confer Resistance to Estrogen Receptor (ER) Targeted Therapy in ER+ Metastatic Breast Cancer. *Clin Cancer Res* 26, 5974–5989
85. Servetto, A. *et al.* (2021) Nuclear FGFR1 Regulates Gene Transcription and Promotes Antiestrogen Resistance in ER+ Breast Cancer. *Clin Cancer Res* 27, 4379–4396

86. Long, G.V. *et al.* (2014) Combined BRAF and MEK inhibition versus BRAF inhibition alone in melanoma. *N Engl J Med* 371, 1877–1888
87. Lu, S. *et al.* (2026) Savolitinib plus osimertinib versus chemotherapy for advanced, EGFR mutation-positive, MET-amplified non-small-cell lung cancer in China (SACHI): interim analysis of a multicentre, open-label, phase 3 randomised controlled trial. *Lancet* 407, 375–387
88. Dhanasekaran, R. *et al.* (2021) The MYC oncogene — the grand orchestrator of cancer growth and immune evasion. *Nature Reviews Clinical Oncology* 2021 DOI: 10.1038/s41571-021-00549-2
89. Hanahan, D. (2026) Hallmarks of cancer—Then and now, and beyond. *Cell* 0
90. Bhin, J. *et al.* (2023) MYC is a clinically significant driver of mTOR inhibitor resistance in breast cancer. *Journal of Experimental Medicine* 220, e20211743
91. Duffy, M.J. *et al.* (2025) MYC as a Target for Cancer Treatment: from Undruggable to Druggable? *Target Oncol* 20, 791–801
92. Whitfield, J.R. and Soucek, L. (2025) MYC in cancer: from undruggable target to clinical trials. *Nat Rev Drug Discov* 24, 445–457
93. Facchinetti, F. *et al.* (2024) Understanding and Overcoming Resistance to Selective FGFR Inhibitors across FGFR2-Driven Malignancies. *Clin Cancer Res* 30, 4943–4956
94. Ma, F. *et al.* (2016) ctDNA dynamics: a novel indicator to track resistance in metastatic breast cancer treated with anti-HER2 therapy. *Oncotarget* 7, 66020–66031
95. Cao, H. *et al.* (2020) Circulating Tumor DNA Is Capable of Monitoring the Therapeutic Response and Resistance in Advanced Colorectal Cancer Patients Undergoing Combined Target and Chemotherapy. *Front. Oncol.* 10
96. Ernst, S.M. *et al.* (2025) Utilizing ctDNA to discover mechanisms of resistance to targeted therapies in patients with metastatic NSCLC: towards more informative trials. *Nat Rev Clin Oncol* 22, 371–378
97. Priestley, P. *et al.* (2019) Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature* 575, 210–216
98. Alexandrov, L.B. *et al.* (2020) The repertoire of mutational signatures in human cancer. *Nature* 578, 94–101
99. Li, Y. *et al.* (2020) Patterns of somatic structural variation in human cancer genomes. *Nature* 578, 112–121
100. Rheinbay, E. *et al.* (2020) Analyses of non-coding somatic drivers in 2,658 cancer whole genomes. *Nature* 578, 102–111
101. Roepman, P. *et al.* (2021) Clinical Validation of Whole Genome Sequencing for Cancer Diagnostics. *J Mol Diagn* 23, 816–833
102. Rosenquist, R. *et al.* (2022) Clinical utility of whole-genome sequencing in precision oncology. *Seminars in Cancer Biology* 84, 32–39
103. Cuppen, E. *et al.* (2022) Implementation of Whole-Genome and Transcriptome Sequencing Into Clinical Cancer Care. *JCO Precis Oncol* 6, e2200245
104. Koboldt, D.C. (2020) Best practices for variant calling in clinical sequencing. *Genome Med* 12, 91
105. Kosugi, S. and Terao, C. (2024) Comparative evaluation of SNVs, indels, and structural variations detected with short- and long-read sequencing data. *Hum Genome Var* 11, 18
106. Chen, X. *et al.* (2025) Long-Read Sequencing Outperforms Short-Read Sequencing in Detecting Most Structural Variations. *Serican J Med* 2, 23671
107. Inaba, H. *et al.* (2017) Integration of Next-Generation Sequencing to Treat Acute Lymphoblastic Leukemia with Targetable Lesions: The St. Jude Children's Research Hospital Approach. *Front Pediatr* 5, 258
108. Rusch, M. *et al.* (2018) Clinical cancer genomic profiling by three-platform sequencing of whole genome, whole exome and transcriptome. *Nat Commun* 9, 3962

109. Frank, M.O. *et al.* (2019) Sequencing and curation strategies for identifying candidate glioblastoma treatments. *BMC Med Genomics* 12, 56
110. Walter, W. *et al.* (2021) Clinical application of whole transcriptome sequencing for the classification of patients with acute lymphoblastic leukemia. *BMC Cancer* 21, 886
111. Horak, P. *et al.* (2021) Comprehensive Genomic and Transcriptomic Analysis for Guiding Therapeutic Decisions in Patients with Rare Cancers. *Cancer Discov* 11, 2780–2795
112. Newman, S. *et al.* (2021) Genomes for Kids: The Scope of Pathogenic Mutations in Pediatric Cancer Revealed by Comprehensive DNA and RNA Sequencing. *Cancer Discov* 11, 3008–3027
113. Möhrmann, L. *et al.* (2022) Comprehensive genomic and epigenomic analysis in cancer of unknown primary guides molecularly-informed therapies despite heterogeneity. *Nat Commun* 13, 4485
114. Schwarze, K. *et al.* (2018) Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the literature. *Genetics in Medicine* 20, 1122–1130
115. Schwarze, K. *et al.* (2020) The complete costs of genome sequencing: a microcosting study in cancer and rare diseases from a single center in the United Kingdom. *Genet Med* 22, 85–94
116. Pasmans, C.T.B. *et al.* (2021) Micro-costing diagnostics in oncology: from single-gene testing to whole- genome sequencing. *Expert Review of Pharmacoeconomics & Outcomes Research* 21, 403–413
117. Ji, H.J. *et al.* (2026) Annotating genomes at increased scale and resolution. *Nat Rev Genet* DOI: 10.1038/s41576-026-00937-3
118. Clark, A.J. and James W. Lillard, J. (2024) A Comprehensive Review of Bioinformatics Tools for Genomic Biomarker Discovery Driving Precision Oncology. *Genes* 15
119. Goldman, M.J. *et al.* (2020) A user guide for the online exploration and visualization of PCAWG data. *Nat Commun* 11, 3400
120. Landrum, M.J. *et al.* (2025) ClinVar: updates to support classifications of both germline and somatic variants. *Nucleic Acids Res* 53, D1313–D1321
121. Sondka, Z. *et al.* (2024) COSMIC: a curated database of somatic variants and clinical data for cancer. *Nucleic Acids Res* 52, D1210–D1217
122. Griffith, M. *et al.* (2017) CIViC is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. *Nat Genet* 49, 170–174
123. Chakravarty, D. *et al.* (2017) OncoKB: A Precision Oncology Knowledge Base. *JCO Precis Oncol* 2017, PO.17.00011
124. Suehnholz, S.P. *et al.* (2024) Quantifying the Expanding Landscape of Clinical Actionability for Patients with Cancer. *Cancer Discov* 14, 49–65
125. The AACR Project GENIE Consortium *et al.* (2017) AACR Project GENIE: Powering Precision Medicine through an International Consortium. *Cancer Discovery* 7, 818–831
126. Acebedo, A. *et al.* (2025) Collaborating across sectors in service of open science, precision oncology, and patients: an overview of the AACR Project GENIE (Genomics Evidence Neoplasia Information Exchange) Biopharma Collaborative (BPC). *ESMO Real World Data Digit Oncol* 7, 100097
127. Riba, M. *et al.* (2024) The 1+Million Genomes Minimal Dataset for Cancer. *Nat Genet* 56, 733–736
128. Chang, T.-G. *et al.* (2025) Hallmarks of artificial intelligence contributions to precision oncology. *Nat Cancer* 6, 417–431
129. Reardon, B. *et al.* (2026) Convergence of machine learning and genomics for precision oncology. *Nat Rev Cancer* 26, 217–229
130. Kehl, K.L. *et al.* (2021) Artificial intelligence-aided clinical annotation of a large multi-cancer genomic dataset. *Nat Commun* 12, 7304
131. Tran, T.N. *et al.* (2025) AI cancer driver mutation predictions are valid in real-world data. *Nat Commun* 16, 8509

132. Nitta, N. *et al.* (2026) Clinical-grade autonomous cytopathology through whole-slide edge tomography. *Nature* DOI: 10.1038/s41586-025-10094-y
133. Zhao, W. *et al.* (2026) An agentic system for rare disease diagnosis with traceable reasoning. *Nature* DOI: 10.1038/s41586-025-10097-9
134. Bixi, G. *et al.* (2026) Genome modelling and design across all domains of life with Evo 2. *Nature* DOI: 10.1038/s41586-026-10176-5
135. Tangermann, C. *et al.* (2025) Saturation mutagenesis identifies activating and resistance-inducing FGFR kinase domain mutations. *Nat Genet* DOI: 10.1038/s41588-025-02431-8
136. Ng, P.K.-S. *et al.* (2018) Systematic Functional Annotation of Somatic Mutations in Cancer. *Cancer Cell* 33, 450–462.e10
137. Muiños, F. *et al.* (2021) In silico saturation mutagenesis of cancer genes. *Nature* 596, 428–432



APPENDICES

ENGLISH SUMMARY

Precision oncology seeks to deliver the right treatment to the right patient at the right time by integrating the molecular characteristics of individual tumors with patient-specific factors. Despite major advances in genomic profiling and the development of targeted therapies, significant challenges remain. One key challenge is the accurate distinction between genomic alterations that function as oncogenic drivers—shaping tumor behavior and therapeutic response—and those that represent passenger events with little or no functional relevance to disease progression or treatment outcome. Another important challenge is to anticipate resistance mechanisms that may pre-exist or emerge during treatment and influence therapeutic response. This thesis contributes to addressing these challenges by investigating the functional and clinical relevance of genomic alterations in the fibroblast growth factor receptor (FGFR) family—particularly in FGFR2 and FGFR3—and by elucidating mechanisms of therapeutic resistance within the PI3K–AKT–mTOR signaling pathway.

The first part of this thesis focuses on *FGFR* alterations, which are increasingly recognized as important therapeutic targets across multiple cancer types. While *FGFR* amplifications and common hotspot mutations are relatively well characterized, another important class remains underexplored: *FGFR* rearrangements, including both canonical in-frame fusions and complex noncanonical variants. In **Chapter 2**, we review the biological and clinical relevance of both canonical FGFR fusions and non-canonical structural variants, with a particular focus on their oncogenic mechanisms, biomarker potential, and therapeutic implications. A key concept discussed in this chapter is the critical role of C-terminal exon 18 (E18) truncations in *FGFR2* and *FGFR3* as drivers of oncogenic activation.

As investigated in **Chapter 3**, C-terminal truncation of *FGFR2* constitutes a potent oncogenic event. Using functional genomic approaches, *in vivo* tumor modeling, and analysis of human oncogenomic datasets, we show that these truncations arise through diverse genomic mechanisms—including rearrangements, partial amplifications, and truncating mutations—all converging on the production of oncogenic *FGFR2* transcripts lacking the E18-encoded C-terminal tail. In contrast, amplification of full-length *FGFR2* alone is insufficient to drive tumorigenesis without additional cooperating alterations. Importantly, the presence of E18-truncated *FGFR2* predicts sensitivity to FGFR-targeted therapies, highlighting its potential as a clinically actionable biomarker. The mechanistic basis of this oncogenic activation is elucidated in **Chapter 4**, where we identify a previously unrecognized regulatory function of the FGFR2 C-terminal tail. Specifically, we uncover a critical motif that mediates autoinhibition of the kinase domain under physiological conditions. Loss or disruption of this motif abolishes this regulatory constraint, resulting in constitutive activation of FGFR2 and downstream signaling pathways, including PI3K–

AKT–mTOR and STAT3. These findings provide a structural explanation for how C-terminal truncations drive oncogenic signaling and suggest potential therapeutic strategies, including combined inhibition of FGFR2 and downstream pathways to overcome resistance to FGFR-targeted therapies.

In **Chapter 5**, we examine whether similar mechanisms apply to FGFR3 and reveal important functional differences between these closely related receptors. In contrast to FGFR2, we show that C-terminal truncation alone is insufficient for full oncogenic activation of FGFR3 and that fusion to a partner gene providing a dimerization domain is required. We further validate the oncogenic potential of several previously uncharacterized C-terminally truncated *FGFR3* fusions involving partner genes containing dimerization domains. These alterations confer resistance to certain receptor tyrosine kinase inhibitors, yet remain sensitive to pharmacological FGFR inhibition. Together, these findings refine our understanding of FGFR3-driven cancers and have important implications for the future interpretation of genomic alterations in clinical settings.

The second part of this thesis focuses on mechanisms of resistance to targeted therapies within the PI3K–AKT–mTOR pathway. In **Chapter 6**, we investigate resistance to mTOR inhibitors in breast cancer using multi-omic profiling of treatment-naïve and resistant tumor models. We identify activation of the MYC oncogene as a key driver of resistance, demonstrating that MYC sustains protein synthesis and tumor growth despite mTOR inhibition. Importantly, resistant tumors remain dependent on MYC activity, and combined inhibition of mTOR and MYC restores treatment sensitivity. These findings are supported by clinical data showing that MYC activation is associated with poor response to the mTOR inhibitor everolimus in patients with metastatic breast cancer. Together, these results establish MYC as both a mediator of resistance and a potential biomarker for patient stratification, suggesting that combination therapeutic strategies may improve clinical outcomes.

Finally, in **Chapter 7**, the main findings of this thesis are discussed in a broader context, highlighting their implications for precision oncology, remaining challenges, and directions for future research.

In summary, this thesis provides new insights into the molecular mechanisms underlying FGFR-driven oncogenesis and resistance to mTOR inhibition. By identifying functionally relevant genomic alterations and uncovering key regulatory processes, this work contributes to a more precise definition of actionable cancer drivers and resistance mechanisms, while highlighting new opportunities for therapeutic intervention. Ultimately, these findings may improve patient selection for targeted therapies and inform strategies to overcome drug resistance, thereby advancing the goals of precision oncology.

NEDERLANDSE SAMENVATTING

Precisie-oncologie heeft als doel om de juiste behandeling op het juiste moment aan de juiste patiënt toe te dienen, door de moleculaire kenmerken van individuele tumoren te integreren met patiëntspecifieke factoren. Ondanks belangrijke vooruitgang in genomische profilering en de ontwikkeling van doelgerichte therapieën, blijven er aanzienlijke uitdagingen bestaan. Een belangrijke uitdaging is het nauwkeurig onderscheiden van genomische afwijkingen die kanker veroorzaken—en daarmee tumorgedrag en therapierespons sturen—van afwijkingen die slechts bijkomende DNA-schade vertegenwoordigen zonder enig effect op ziekteprogressie of behandeluitkomst. Een andere belangrijke uitdaging is het anticiperen op resistentiemechanismen die reeds aanwezig kunnen zijn of zich tijdens de behandeling ontwikkelen en de therapierespons beïnvloeden. Dit proefschrift draagt bij aan het aanpakken van deze uitdagingen door de functionele en klinische relevantie van genomische afwijkingen in de fibroblastgroeifactorreceptoren FGFR2 en FGFR3 te onderzoeken, evenals door mechanismen van therapeutische resistentie binnen de PI3K–AKT–mTOR-signaalroute te ontrafelen.

Het eerste deel van dit proefschrift richt zich op *FGFR*-afwijkingen, die in toenemende mate worden erkend als belangrijke therapeutische doelwitten in verschillende kankertypen. Hoewel *FGFR*-amplificaties en veelvoorkomende hotspotmutaties relatief goed gekarakteriseerd zijn, blijft een andere belangrijke klasse onderbelicht: *FGFR*-herschikkingen, waaronder zowel canonieke fusies tussen *FGFR* genen en andere genen als complexe niet-canonieke varianten. In **Hoofdstuk 2** bespreken we de biologische en klinische relevantie van zowel canonieke FGFR-fusies als niet-canonieke structurele varianten, met specifieke aandacht voor hun oncogene mechanismen, biomarkerpotentieel en therapeutische implicaties. Een centraal concept in dit hoofdstuk is de cruciale rol van verlies van het C-terminale exon 18 (E18) in *FGFR2* en *FGFR3* als drijvende kracht achter oncogene activatie.

Zoals onderzocht in **Hoofdstuk 3**, vormt C-terminale truncatie van *FGFR2* een krachtige oncogene gebeurtenis. Met behulp van functionele experimenten, *in vivo* tumormodellen en analyse van humane oncogenomische datasets laten we zien dat deze truncaties via diverse genomische mechanismen kunnen ontstaan—waaronder herschikkingen, partiële amplificaties en truncerende mutaties—die allemaal leiden tot de expressie van oncogene *FGFR2*-transcripten zonder de door E18 gecodeerde C-terminale staart. Daarentegen is amplificatie van volledig *FGFR2* op zichzelf onvoldoende om tumorvorming te veroorzaken zonder aanvullende coöpererende genetische veranderingen. Belangrijk is dat de aanwezigheid van E18-getrunceerd *FGFR2* gevoeligheid voor FGFR-gerichte therapieën voorspelt, wat het potentieel onderstreept als klinisch bruikbare biomarker.

De mechanistische basis van deze oncogene activatie wordt verder uitgewerkt in **Hoofdstuk 4**, waarin we een voorheen onbekende regulerende functie van de C-terminale staart van FGFR2 identificeren. We tonen aan dat een specifiek eiwitmotief binnen deze regio verantwoordelijk is voor autoinhibitie van het kinasedomein onder fysiologische omstandigheden. Verlies of verstoring van dit motief heft deze regulatie op, wat leidt tot constitutieve activatie van FGFR2 en de downstream signaalroutes, waaronder PI3K-AKT-mTOR en STAT3. Deze bevindingen bieden een structurele verklaring voor hoe C-terminale truncaties oncogene signalering aandrijven en suggereren mogelijke therapeutische strategieën, waaronder gecombineerde remming van FGFR2 en downstream signaalroutes, om resistentie tegen FGFR-gerichte therapieën te overwinnen.

In **Hoofdstuk 5** onderzoeken we of vergelijkbare mechanismen van toepassing zijn op FGFR3 en tonen we belangrijke functionele verschillen tussen deze nauw verwante receptoren aan. In tegenstelling tot FGFR2 is C-terminale truncatie van FGFR3 alléén onvoldoende voor volledige oncogene activatie, en is fusie met een partnergen dat een dimerisatiedomein bevat noodzakelijk. Daarnaast valideren we het oncogene potentieel van verschillende ongekaracteriseerde *FGFR3*-fusies met partnergenen die dimerisatiedomeinen bevatten. Deze fusie-eiwitten veroorzaken resistentie tegen bepaalde receptor-tyrosinekinaseremmers maar blijven gevoelig voor farmacologische remming van FGFR. Samen verfijnen deze resultaten ons begrip van FGFR3-aangedreven kankers en hebben ze belangrijke implicaties voor de interpretatie van deze genomische afwijkingen in de kliniek.

Het tweede deel van dit proefschrift richt zich op resistentiemechanismen tegen therapeutische remmers van de PI3K-AKT-mTOR-signaalroute. In **Hoofdstuk 6** onderzoeken we resistentie tegen mTOR-remmers bij borstkanker met behulp van multi-omicsprofilering van onbehandelde en therapie-resistente muizentumoren. We laten zien dat activatie van het MYC-oncogen resistentie veroorzaakt, en tonen aan dat MYC eiwitsynthese en tumorgroei in stand houdt ondanks remming van mTOR. Belangrijk is dat resistente tumoren afhankelijk blijven van MYC-activiteit en dat gecombineerde remming van mTOR en MYC de gevoeligheid voor behandeling herstelt. Deze preklinische bevindingen worden ondersteund door klinische gegevens die aantonen dat MYC-activatie geassocieerd is met een slechte respons op de mTOR-remmer everolimus bij patiënten met gemetastaseerde borstkanker. Samen positioneren deze resultaten MYC als een mediator van resistentie en een potentiële biomarker voor patiëntenselectie, en suggereren zij dat combinatietherapieën de klinische uitkomsten kunnen verbeteren.

Tot slot worden in **Hoofdstuk 7** de belangrijkste bevindingen van dit proefschrift in een bredere context geplaatst, waarbij de implicaties voor precisie-oncologie, de resterende uitdagingen en mogelijke toekomstige onderzoeksrichtingen worden besproken.

Samenvattend biedt dit proefschrift nieuwe inzichten in de moleculaire mechanismen die ten grondslag liggen aan FGFR-gedreven tumorvorming en resistentie tegen mTOR-remmers. Door functioneel relevante genomische afwijkingen te identificeren en belangrijke regulerende mechanismen te ontrafelen, draagt dit werk bij aan een meer precieze definitie van klinisch relevante oncogene 'drivers' en resistentiemechanismen, en worden nieuwe mogelijkheden voor therapeutische interventie aangereikt. Uiteindelijk kunnen deze bevindingen bijdragen aan een betere patiëntselectie voor doelgerichte therapieën en aan strategieën om geneesmiddelresistentie te overwinnen, en daarmee de verdere ontwikkeling van precisie-oncologie ondersteunen.

CURRICULUM VITAE

Julia Yemelyanenko was born on November 21, 1993, in Sevastopol, Ukraine. In 2002, she moved to Spain with her parents. After completing her secondary education in 2012, she enrolled in the Bachelor's degree in Human Biology at the Pompeu Fabra University in Barcelona, Spain. During her undergraduate studies, she undertook two short research internships, at the University of Debrecen (Hungary) and the University of Adelaide (Australia).

Upon obtaining her Bachelor's degree, she moved to the Netherlands in 2016 to pursue a two-year Master's degree in Oncology at the Vrije Universiteit Amsterdam. During this period, she completed two extended research internships that sparked her strong interest in oncology research, particularly in its translational aspects. Her first internship was at the VUmc Cancer Center Amsterdam (CCA), in the groups of Prof. Dr. Tanja de Gruijl and Prof. Dr. Victor van Beusechem, where she investigated combination therapies involving an oncolytic virus and immune-modulating compounds in melanoma. For her second internship, she joined the group of Prof. Dr. Jos Jonkers at the Netherlands Cancer Institute (NKI), where she worked on DNA damage response and PARP inhibitor resistance.

Following her Master's graduation in 2018, she continued at the NKI as a laboratory technician in the group of Prof. Dr. Jos Jonkers. During this time, she focused on the validation of a novel potential mechanism of resistance to mTOR inhibitors in invasive lobular carcinoma. In 2021, she started her PhD in the same group, shifting her focus to the functional characterization of somatic variants of unknown significance in clinically actionable kinases, with particular emphasis on FGFR2 and FGFR3. The findings of this work are presented in this thesis.

LIST OF PUBLICATIONS

Yemelyanenko J, Jonkers J#, Zingg D#. FGFR rearrangements: oncogenic drivers and therapeutic targets. *Trends Cancer*. 2026 May;12(5):461-476. doi: 10.1016/j.trecan.2026.01.010.

Yemelyanenko J*, Bhin J*, van der Burg E, Drenth AP, Lee JK, Lutz C, Dörner L, Wientjens E, Klarenbeek S, Song JY, Moon H, Annunziato S, Proost N, Siteur B, Ross JS, van de Ven M, van Tellingen O, Ganesan S, Wessels LFA, Zingg D#, Jonkers J#. C-terminal Truncation and Fusion Partner Determine Oncogenicity of FGFR3. *Cancer Res*. 2026 Mar 16;86(6):1372-1391. doi: 10.1158/0008-5472.CAN-24-2648.

Zingg D*#, Lin CC*, **Yemelyanenko J**, Wieteska L, Kas SM, Bleijerveld OB, Chao X, Bhin J, Lutz C, Wientjens E, Klarenbeek S, Zanec G, Annunziato S, Siteur B, van der Burg E, Drenth AP, van de Ven M, Wessels LFA, Altelaar M, Ladbury JE#, Jonkers J#. The C-Terminal Kinase Domain-Binding and Suppression Motif Prevents Constitutive Activation of FGFR2. *Cancer Res*. 2025 Sep 2;85(17):3234-3257. doi: 10.1158/0008-5472.CAN-24-3349.

Bhin J*, **Yemelyanenko J***, Chao X, Klarenbeek S, Opdam M, Malka Y, Hoekman L, Kruger D, Bleijerveld O, Brambillasca CS, Sprengers J, Siteur B, Annunziato S, van Haren MJ, Martin NI, van de Ven M, Peters D, Agami R, Linn SC, Boven E, Altelaar M, Jonkers J#, Zingg D#, Wessels LFA#. MYC is a clinically significant driver of mTOR inhibitor resistance in breast cancer. *J Exp Med*. 2023 Nov 6;220(11):e20211743. doi: 10.1084/jem.20211743.

Zingg D*, Bhin J*, **Yemelyanenko J***, Kas SM*, Rolfs F, Lutz C, Lee JK, Klarenbeek S, Silverman IM, Annunziato S, Chan CS, Piersma SR, Eijkman T, Badoux M, Gogola E, Siteur B, Sprengers J, de Klein B, de Goeij-de Haas RR, Riedlinger GM, Ke H, Madison R, Drenth AP, van der Burg E, Schut E, Henneman L, van Miltenburg MH, Proost N, Zhen H, Wientjens E, de Bruijn R, de Ruiter JR, Boon U, de Korte-Grimmerink R, van Gerwen B, Féliz L, Abou-Alfa GK, Ross JS, van de Ven M, Rokenberg S, Cuppen E, Chessex AV, Ali SM, Burn TC, Jimenez CR, Ganesan S#, Wessels LFA#, Jonkers J#. Truncated FGFR2 is a clinically actionable oncogene in multiple cancers. *Nature*. 2022 Aug;608(7923):609-617. doi: 10.1038/s41586-022-05066-5.

Klarenbeek S, Doornebal CW, Kas SM, Bonzanni N, Bhin J, Braumuller TM, van der Heijden I, Opdam M, Schouten PC, Kersten K, de Bruijn R, Zingg D, **Yemelyanenko J**, Wessels LFA, de Visser KE, Jonkers J. Response of metastatic mouse invasive lobular carcinoma to mTOR inhibition is partly mediated by the adaptive immune system. *Oncoimmunology* 2020 Feb 12;9(1):1724049. doi: 10.1080/2162402X.2020.1724049.

*These authors contributed equally as first authors.

#These authors contributed equally as corresponding authors.