Using insertional mutagenesis to identify breast cancer drivers and therapy resistance genes in mice = Insertie mutagenese voor het identificeren van genen betrokken bij de ontwikkeling van borsttumoren en therapie resistentie in muizen

Kas, S.M.

Chapter 5

In vivo oncogenicity of FGFR2 is modulated by multiple C-terminal domains

Sjors M. Kas[1,2], Julian R. de Ruiter[1,2,3], Eva Schut[1,2], Onno B. Bleijerveld[4], Sjoerd Klarenbeek[5], Ellen Wientjens[1,2], Martine H. van Miltenburg[1,2], Anne Paulien Drenth[1,2], Eline van der Burg[1,2], Ute Boon[1,2], Lodewyk F.A. Wessels[2,3,6], Jos Jonkers[1,2]

[1] Division of Molecular Pathology, The Netherlands Cancer Institute, Amsterdam, The Netherlands
[2] Oncode Institute, Amsterdam, The Netherlands
[3] Division of Molecular Carcinogenesis, The Netherlands Cancer Institute, Amsterdam, The Netherlands
[4] Proteomics/Mass Spectrometry Facility, The Netherlands Cancer Institute, Amsterdam, The Netherlands
[5] Experimental Animal Pathology, The Netherlands Cancer Institute, Amsterdam, The Netherlands
[6] Department of EEMCS, Delft University of Technology, Delft, The Netherlands

Manuscript in preparation

Abstract

Genetic aberrations of fibroblast growth factor receptors (FGFRs) are frequently observed in human cancers and include activating mutations, gene amplifications and gene fusions. The FGFR2 gene fusions are hallmarked by an intact kinase domain fused to various 3' fusion partners, which results in oncogenic constitutive kinase activity of FGFR2. Although the fusion partners generally replace the C-terminal tail of FGFR2, the role of the C-terminus in suppressing active FGFR signaling remains largely unknown. We previously identified *Fgfr2* as a key driver of invasive lobular carcinoma (ILC) in a transposon-based insertional mutagenesis screen in mice. Here, we observe that the majority of the *Fgfr2* transposon insertions in the mouse ILCs (mILCs) were clustered in the intron directly upstream of the last exon, resulting in the expression of C-terminal truncated *Fgfr2* transcripts. Mammary-specific expression of truncated FGFR2 variants induced rapid and multifocal ILC formation in mice, whereas no tumors were observed when full-length FGFR2 was expressed. To identify the critical tumor suppressive domains in the C-terminus of FGFR2, we developed a FACS-based assay to measure the S6 phosphorylation levels induced by FGFR2 variants *in vitro*, which correlated with their *in vivo* oncogenicity. Altogether, we provide strong evidence that multiple domains in the C-terminal tail modulate the tumorigenic potential of FGFR2 and we present a powerful approach to identify the critical motifs that are essential for tumor suppression.

Introduction

Receptor tyrosine kinases (RTKs) are a family of cell surface receptors that upon binding to a wide range of signaling moleculesactivate intracellular signaling cascades[1]. Fibroblast growth factor receptors (FGFRs) are members of the RTK family and drive many biological processes during embryonic development and throughout adult life, such as tissue morphogenesis, proliferation, angiogenesis and wound repair[2]. Given its important roles, tight regulation of FGFR activity is essential to maintain normal homeostasis and to prevent progression to a pathological state such as cancer.

A diversity of fibroblast growth factor (FGF) ligands can bind to FGFRs, which induces receptor dimerization followed by cross-phosphorylation of the cognate receptors tyrosine kinase domain and its intracellular tail. The resulting phosphorylated residues are docking sites for several adaptor proteins, such as FGFR substrate 2α (FRS2α) and phospholipase Cγ1 (PLCγ1)[2,3]. Once bound and phosphorylated, FRS2α can recruit growth factor receptor-bound 2 (GRB2) and son of sevenless (SOS) to form a complex that results in the activation of the RAS-MAPK signaling pathway. On the other hand, the recruitment of GRB2-associated binding protein 1 (GAB1) and phosphoinositide 3-kinase (PI3K) leads to the activation of the PI3K/AKT signaling pathway. Furthermore, phosphorylated FGFRs can also activate intracellular signaling cascades via PLCγ1 and signal transducers and activators of transcription (STATs).

In human cancer, FGFR signaling is frequently hyper-activated by deregulation of FGF ligands or by genetic alterations of the receptor, such as activating point mutations, gene amplifications and gene fusions[4]. Using RNA-sequencing based analyses, FGFR gene fusions were identified in patients with different tumor types, including cholangiocarcinoma, lung squamous cell, breast and gastric cancer[5–7]. In most of these cases, FGFRs with intact kinase domains are fused in-frame with different 3' fusion partners, which replace the C-terminal tail of the receptor[5,6]. In line with these observations, several studies reported the expression of FGFR2 variants lacking its carboxyl-terminus in human cancer cell lines and that the expression of these variants in cells resulted in enhanced transforming activity *in vitro*[8–11]. Moreover, recent studies show that GRB2 also binds to the last 10 amino acid residues in FGFR2 to control its kinase activity prior to ligand binding[12,13]. Collectively, these data indicate that the C-terminus is important in the regulation of FGFR signaling.

In previous work, we identified *Fgfr2* as a key driver of invasive lobular breast carcinoma (ILC) using a *Sleeping Beauty* (*SB*)-based transposon insertional mutagenesis screen in mice[14]. Here, we observed that the majority of transposon insertions in *Fgfr2* are clustered in the intron upstream of the last exon, which resulted in the expression of C-terminal truncated *Fgfr2* transcripts in *SB*-induced mouse ILCs (mILCs). Furthermore, we show that mammary-specific expression of FGFR2 C-terminal truncation variants induce rapid and multifocal ILC formation in mice, whereas expression of full-length FGFR2 did not show potent mammary

tumor development. Altogether, we demonstrate that the C-terminal tail of FGFR2 controls active receptor signaling and is therefore necessary to suppress malignant transformation of mammary epithelium.

Results

SB-induced tumors show two clusters of transposon insertions in Fgfr2

In previous work, we identified genes and pathways involved in the development of invasive lobular breast carcinoma (ILC) using a *Sleeping Beauty* (*SB*)-based insertional mutagenesis screen in mice with mammary-specific inactivation of *Cdh1* (encoding E-cadherin)[14]. Analysis of common insertion sites (CISs) in 123 *SB*-induced mouse ILCs (mILCs) showed that 64 of these tumors had insertions in *Fgfr2*, suggesting that *Fgfr2* is a driver of ILC formation.

Analysis of the *SB* insertions in *Fgfr2* revealed two clusters of insertions of which one cluster was located upstream of the transcription start site (hereafter referred to as 5' insertions) and the other was in the intron directly upstream of the last exon (hereafter referred to as 3' insertions) (Figure 1a). Closer analysis of the insertions showed that 5 of the 7 *SB*-induced tumors with a 5' insertion also contain a 3' insertion. Notably, the remaining two tumors with a 5' insertion harbored gene fusions between *Fgfr2* and other genes (*Kif16b* and *Tbc1d1*), which were

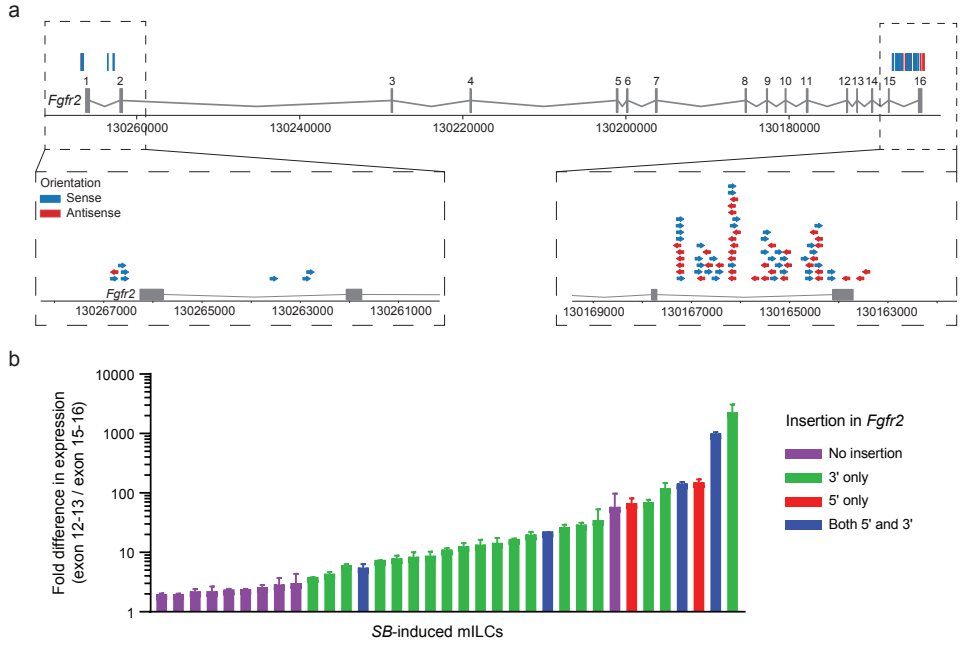


Figure 1. Overview of insertions in *Fgfr2* and their effects on gene expression over the insertion sites in *SB*-induced tumors. (a) Schematic overview of *SB* insertions in *Fgfr2* (*n* = 123 tumors). The bars represent the exact genomic location of the insertions, demonstrating defined clusters around the transcriptional start site and in the intron directly upstream of exon 16. **(b)** The fold difference in *Fgfr2* expression upstream (exons 12-13) and downstream of the insertion sites (exons 15-16). Data are mean ± standard deviation (s.d.) of three independent experiments, and in each experiment the samples were measured in triplicate. The expression of *Fgfr2* was determined by qRT-PCR and corrected for HPRT levels.

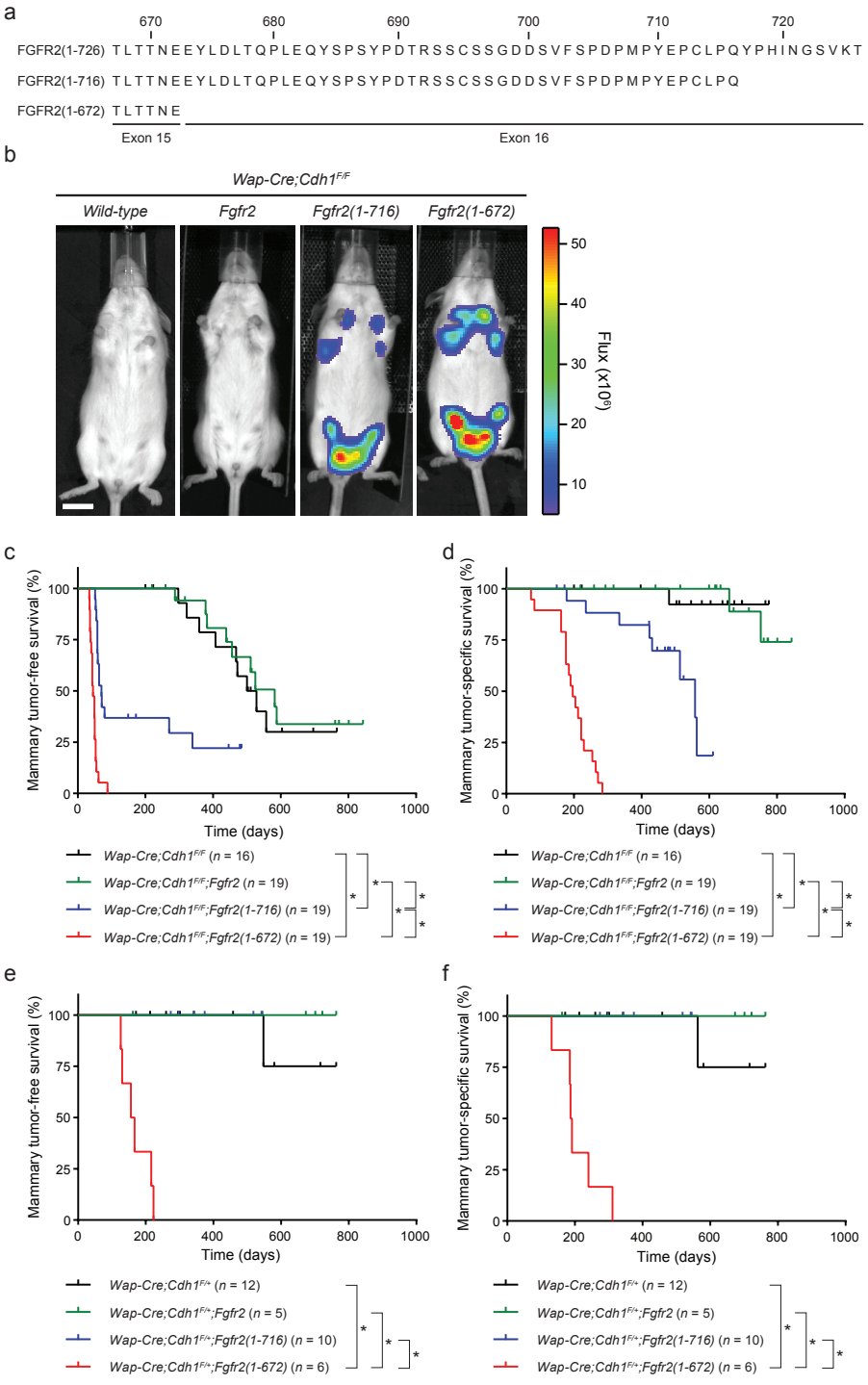
previously identified in RNA-sequencing data of these *SB*-induced mILCs[15]. These findings suggest that all *SB* insertions in *Fgfr2* were strongly selected for inducing expression of a truncated *Fgfr2* transcript lacking its last exon.

To test this hypothesis, we performed quantitative RT-PCR (qRT-PCR) on *SB*-induced mILCs to compare expression of exons upstream (exons 12-13) and downstream (exons 15-16) of the 3' insertion sites, respectively. As expected, *SB*-induced tumors with insertions in *Fgfr2* generally showed increased mRNA expression upstream compared to downstream of the insertion sites (Figure 1b), whereas no difference was observed in tumors without insertions in *Fgfr2*. Altogether, these data suggest that the transposon insertions induce the expression of a carboxyl (C)-terminal truncation variant of FGFR2, which is frequently observed in *SB*-induced tumors and could therefore induce the formation of ILC.

C-terminal truncation variants of FGFR2 induced rapid and multifocal mammary tumor formation in Wap–Cre;Cdh1$^{F/F}$ mice

Next, we investigated whether different C-terminal truncation variants of FGFR2 are capable of driving ILC formation in conditional E-cadherin (*Cdh1*)-deficient mice. For this purpose, we generated genetically engineered mice with mammary-specific inactivation of E-cadherin and activation of native or truncated FGFR2 (Figure 2a). We included the truncated FGFR2 lacking the last exon (FGFR2(1-672)) and an FGFR2 variant that lacks the last 10 amino acid residues of the C-terminus, which has been shown to abrogate the binding of GRB2 (FGFR2(1-716))[12]. We introduced *invCAG-Fgfr2(1-726)-IRES-Luc*, *invCAG-Fgfr2(1-716)-IRES-Luc* and *invCAG-Fgfr2(1-672)-IRES-Luc* alleles into the *Col1a1* locus of *Wap–Cre*;*Cdh1$^{F/F}$* embryonic stem cells (ESCs) for Cre-inducible expression of firefly luciferase and full-length FGFR2, FGFR2(1-716) and FGFR2(1-672), respectively (Supplementary Figure 1a). Blastocyst injections of the modified ESCs were performed to produce chimeric mice, of which high quality male chimeras were mated with *Cdh1$^{F/F}$* females to generate cohorts of female *Wap–Cre*;*Cdh1$^{F/F}$*;*Col1a1$^{invCAG\text{-}Fgfr2(1\text{-}726)\text{-}IRES\text{-}Luc/+}$* (hereafter referred to as *Wap–Cre*;*Cdh1$^{F/F}$*

Figure 2. ► FGFR2 C-terminal truncation variants induce rapid and multifocal mammary tumor formation in *Wap-Cre;Cdh1$^{F/F}$* mice. (a) The amino acid sequences of full-length FGFR2 (726 amino acids) and the two FGFR2 C-terminal truncation variants, which either lack the last 54 (FGFR2(1-672)) or the last 10 amino acid residues (FGFR2(1-716). **(b)** Representative images of *in vivo* bioluminescence imaging of luciferase expression in 10-week-old *Wap-Cre;Cdh1$^{F/F}$*, *Wap-Cre;Cdh1$^{F/F}$;Fgfr2*, *Wap-Cre;Cdh1$^{F/F}$;Fgfr2(1-716)* and *Wap-Cre;Cdh1$^{F/F}$;Fgfr2(1-672)* females. Scale bar, 1 cm. **(c)** Kaplan-Meier curve showing the mammary tumor-free survival for the indicated genotypes. *Wap-Cre;Cdh1$^{F/F}$;Fgfr2(1-716)* (70 days) and *Wap-Cre;Cdh1$^{F/F}$;Fgfr2(1-672)* (45 days) show reduced survival compared to *Wap-Cre;Cdh1$^{F/F}$* (514 days) and *Wap-Cre;Cdh1$^{F/F}$;Fgfr2* (582 days) female mice. **P* < 0.05 by Mantel–Cox test. **(d)** Kaplan-Meier curve showing the mammary tumor-specific survival (as defined in the Materials and Methods section) for the indicated genotypes. *Wap-Cre;Cdh1$^{F/F}$;Fgfr2(1-672)* (196 days) females show reduced survival compared to *Wap-Cre;Cdh1$^{F/F}$;Fgfr2(1-716)* (558 days), *Wap-Cre;Cdh1$^{F/F}$* and *Wap-Cre;Cdh1$^{F/F}$;Fgfr2* females. **P* < 0.05 by Mantel–Cox test. **(e,f)** The mammary tumor-free **(e)** and tumor-specific **(f)** survival of females with the indicated genotypes as shown using Kaplan-Meier curves. *Wap-Cre;Cdh1$^{F/+}$;Fgfr2(1-672)* females show reduced tumor-free and tumor-specific survival compared to *Wap-Cre;Cdh1$^{F/+}$*, *Wap-Cre;Cdh1$^{F/+}$;Fgfr2* and *Wap-Cre;Cdh1$^{F/+}$;Fgfr2(1-716)* female mice. **P* < 0.05 by Mantel–Cox test.


a
670
680
690
700
710
720
FGFR2(1-726) T L T T N E E Y L D L T Q P L E Q Y S P S Y P D T R S S C S S G D D S V F S P D P M P Y E P C L P Q Y P H I N G S V K T
FGFR2(1-716) T L T T N E E Y L D L T Q P L E Q Y S P S Y P D T R S S C S S G D D S V F S P D P M P Y E P C L P Q
FGFR2(1-672) T L T T N E
Exon 15
Exon 16
b
Wap-Cre;Cdh1F/F
Wild-type
Fgfr2
Fgfr2(1-716)
Fgfr2(1-672)
50
40
30
20
10
Flux (x10⁶)
c
Mammary tumor-free survival (%)
Time (days)
Wap-Cre;Cdh1F/F (n = 16)
Wap-Cre;Cdh1F/F;Fgfr2 (n = 19)
Wap-Cre;Cdh1F/F;Fgfr2(1-716) (n = 19)
Wap-Cre;Cdh1F/F;Fgfr2(1-672) (n = 19)
d
Mammary tumor-specific survival (%)
Time (days)
Wap-Cre;Cdh1F/F (n = 16)
Wap-Cre;Cdh1F/F;Fgfr2 (n = 19)
Wap-Cre;Cdh1F/F;Fgfr2(1-716) (n = 19)
Wap-Cre;Cdh1F/F;Fgfr2(1-672) (n = 19)
e
Mammary tumor-free survival (%)
Time (days)
Wap-Cre;Cdh1F/+ (n = 12)
Wap-Cre;Cdh1F/+;Fgfr2 (n = 5)
Wap-Cre;Cdh1F/+;Fgfr2(1-716) (n = 10)
Wap-Cre;Cdh1F/+;Fgfr2(1-672) (n = 6)
f
Mammary tumor-specific survival (%)
Time (days)
Wap-Cre;Cdh1F/+ (n = 12)
Wap-Cre;Cdh1F/+;Fgfr2 (n = 5)
Wap-Cre;Cdh1F/+;Fgfr2(1-716) (n = 10)
Wap-Cre;Cdh1F/+;Fgfr2(1-672) (n = 6)

;*Fgfr2*), *Wap–Cre*;*Cdh1*$^{F/F}$;*Col1a1*$^{invCAG\text{-}Fgfr2(1\text{-}716)\text{-}IRES\text{-}Luc/+}$ (*Wap–Cre*;*Cdh1*$^{F/F}$; *Fgfr2(1-716)*) and *Wap–Cre*;*Cdh1*$^{F/F}$; *Col1a1*$^{invCAG\text{-}Fgfr2(1\text{-}672)\text{-}IRES\text{-}Luc/+}$ (*Wap–Cre*;*Cdh1*$^{F/F}$;*Fgfr2(1-672)*) mice[16], which were monitored for mammary tumor formation. In these mice, the expression of Cre recombinase is under the transcriptional control of the *Wap* gene promoter, which resulted in mammary-specific loss of E-cadherin and expression of native or truncated FGFR2 and luciferase (Figure 2b and Supplementary Figure 1b). *Wap–Cre*;*Cdh1*$^{F/F}$;*Fgfr2(1-672)* and *Wap–Cre*;*Cdh1*$^{F/F}$;*Fgfr2(1-716)* female mice showed increased bioluminescence signals compared to *Wap–Cre*;*Cdh1*$^{F/F}$;*Fgfr2* and *Wap–Cre*;*Cdh1*$^{F/F}$ females, which coincided with rapid and multifocal mammary tumor formation in the majority of these animals (Figure 2c). In contrast, expression of native FGFR2 did not accelerate the formation of mammary tumors in *Wap–Cre*;*Cdh1*$^{F/F}$ mice. Although there was no clear difference in tumor onset between *Wap–Cre*;*Cdh1*$^{F/F}$;*Fgfr2(1-672)* and *Wap–Cre*;*Cdh1*$^{F/F}$;*Fgfr2(1-716)* females, there was a strongly decreased mammary-tumor specific survival of *Wap–Cre*;*Cdh1*$^{F/F}$;*Fgfr2(1-672)* mice (196 days) compared to *Wap–Cre*;*Cdh1*$^{F/F}$;*Fgfr2(1-716)* mice (558 days) (Figure 2d).

To determine the mammary tumor burden induced by the different FGFR2 truncation variants, we performed histopathological analysis of all mammary glands in these animals. This analysis revealed that 94% of the glands contained mammary tumors in *Wap–Cre*;*Cdh1*$^{F/F}$;*Fgfr2(1-672)* mice compared to 73% in *Wap–Cre*;*Cdh1*$^{F/F}$;*Fgfr2(1-716)* mice (Supplementary Figure 1c, 1d), As expected, the majority of the tumors showed the typical ILC characteristics with single-strands of noncohesive E-cadherin-negative and cytokeratin 8 (CK8)-positive tumor cells infiltrating the surrounding tissue. Tumors with a spindle cell morphology were observed in 10% and 6% in mammary glands of *Wap–Cre*;*Cdh1*$^{F/F}$;*Fgfr2(1-672)* and *Wap–Cre*;*Cdh1*$^{F/F}$;*Fgfr2(1-716)* mice, respectively. In contrast, mammary glands of *Wap–Cre*;*Cdh1*$^{F/F}$;*Fgfr2* and *Wap–Cre*;*Cdh1*$^{F/F}$ female mice mainly consisted of normal epithelium and early (small) lesions with noncohesive cells.

Altogether, these data provide strong evidence that both FGFR2 truncation variants are able to initiate ILC formation in *Wap–Cre*;*Cdh1*$^{F/F}$ mice. However, mice expressing FGFR2(1-672) showed a decreased mammary tumor-specific survival and an increased mammary tumor burden compared to mice expressing FGFR2(1-716), indicating that *Fgfr2(1-672)* is a stronger oncogenic driver in *Wap–Cre*;*Cdh1*$^{F/F}$ mice.

Expression of FGFR2(1-672) also induced mammary tumor formation in E-cadherin proficient mammary epithelium

Next, we tested whether the two C-terminal truncation variants of FGFR2 could also induce malignant transformation of E-cadherin-proficient mammary epithelium. For this purpose, the modified *Wap–Cre*;*Cdh1*$^{F/F}$ ESCs derived chimeras carrying Cre-inducible *Fgfr2* variants were mated with *Cdh1*$^{F/+}$ females. We generated cohorts of *Wap–Cre*;*Cdh1*$^{F/+}$;*Fgfr2*, *Wap–Cre*;*Cdh1*$^{F/+}$;*Fgfr2(1-716)* and *Wap–Cre*;*Cdh1*$^{F/+}$;*Fgfr2(1-672)* female mice, which were compared to *Wap–Cre*;*Cdh1*$^{F/+}$ littermate controls for spontaneous mammary tumor formation. Of these cohorts, only *Wap–*

Cre;*Cdh1*$^{F/+}$;*Fgfr2(1-672)* female mice showed multiple palpable mammary tumors with a median mammary tumor-free survival of 162 days (Figure 2e). Notably, all of these animals were sacrificed due to a large tumor size (>1500 mm^3) within a few weeks after tumor detection (Figure 2f), indicating that these mammary tumors had a high proliferation rate. Histopathological analysis showed that 78% of the mammary glands in these mice contained tumors that were mainly classified as solid carcinoma, which was confirmed by E-cadherin and CK8-positive tumor cells (Supplementary Figure 1c, 1e).

Thus, expression of both FGFR2(1-672) and FGFR2(1-716) resulted in ILC formation in *Wap–Cre*;*Cdh1*$^{F/F}$ mice, whereas mammary tumors were only detected when FGFR2(1-672) was expressed in *Wap–Cre*;*Cdh1*$^{F/+}$ females. Interestingly, several studies reported oncogenic FGFR2 gene fusions in different human cancer types using RNA-sequencing-based analyses[5–7]. The majority of these in-frame fusions are composed of an intact FGFR2 kinase domain and a wide range of different 3' fusion partners, which in all cases replaced the C-terminal exon of *FGFR2*. These findings support our data which shows that the last exon of *Fgfr2* prevents malignant transformation of mammary epithelium. Furthermore, the difference in oncogenic potential between the two truncated FGFR2 variants suggests that multiple factors might play a role in suppressing tumor formation.

The C-terminal tail of FGFR2 in FGFR2-GFP fusions controls FGFR signaling to prevent malignant transformation of mammary epithelial cells

To investigate whether the deletion of the C-terminal tail in FGFR2 gene fusions is crucial for its oncogenic activity, we hypothesized that cells expressing an in-frame FGFR2 fusion including the C-terminal tail would lack the capacity to induce malignant transformation of mammary epithelial cell. Furthermore, FGFR2 fusions with and without a C-terminal tail could be used to investigate differences in interacting proteins and thereby provide mechanistic insights underlying the tumor suppressive function of the C-terminal tail. For this purpose, we generated in-frame fusions of native and truncated FGFR2 variants with green fluorescent protein (GFP) and transduced spontaneously immortalized normal mouse mammary epithelial (NMuMG) cells with lentiviruses encoding GFP, FGFR2-GFP, FGFR2(1-716)-GFP and FGFR2(1-672)-GFP, respectively. The GFP^{high}-expressing cells were separated using fluorescence activated cell sorting (FACS) and we subsequently cultured these cells to analyze FGFR pathway activation. In 10% fetal bovine serum (FBS)-containing medium, NMuMG cells expressing FGFR2(1-672)-GFP showed increased expression of phosphorylated FGFR, FRS2α, AKT and S6 compared to cells expressing GFP, FGFR2-GFP or FGFR2(1-716)-GFP (Figure 3a). Interestingly, these increased phosphorylation levels were sustained under serum-starved conditions for 48 hours. These data suggest that mammary epithelial cells expressing this truncated FGFR2-GFP variant might have an increased tumorigenic potential induced by active FGFR signaling.

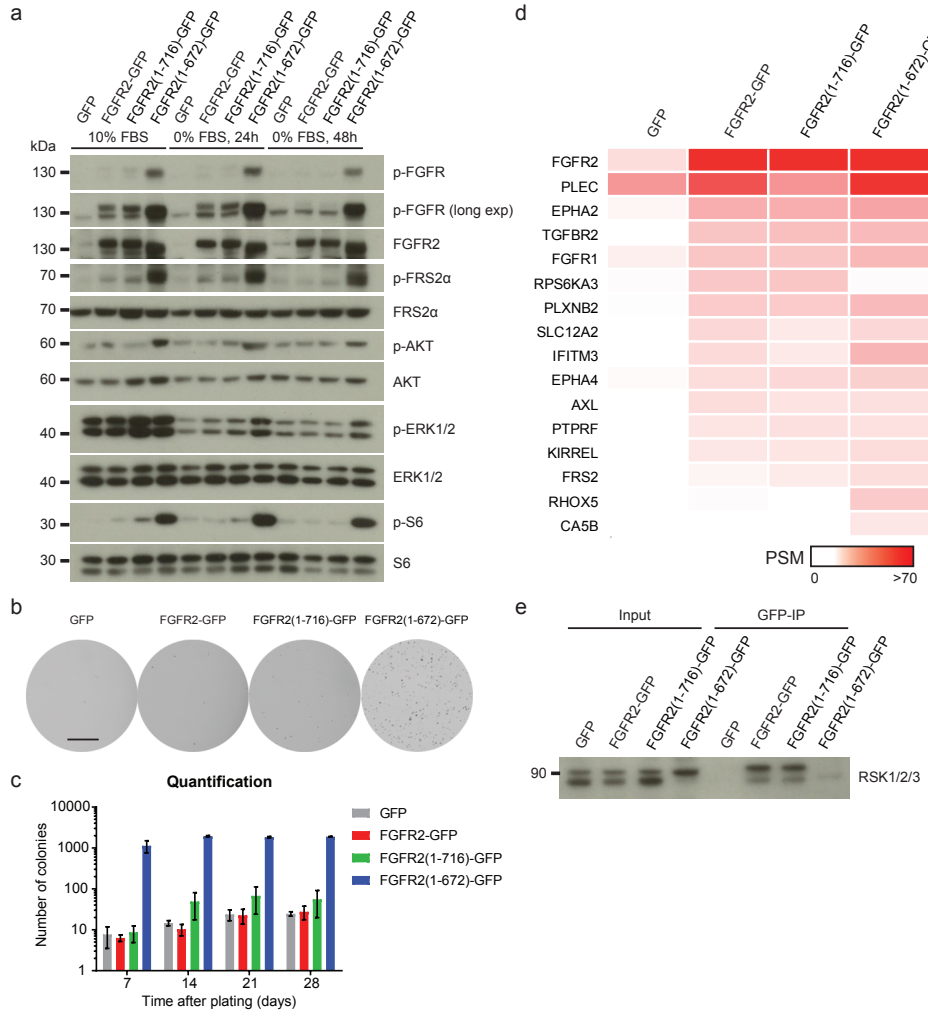


Figure 3. Mammary epithelial cells expressing FGFR2(1-672)-GFP show prolonged activation of FGFR signaling, which coincides with increased tumorigenic potential and reduced RSK2 binding compared to cells expressing GFP, FGFR2-GFP and FGFR2(1-716)-GFP. (a) Immunoblot for FGFR downstream signaling levels in NMuMG cells expressing GFP-only or the different FGFR2-GFP variants in normal and serum-starved (24h and 48h) conditions. **(b)** Representative whole well images of NMuMG cells expressing GFP-only or different FGFR2-GFP variants that were allowed to grow in soft agar for 28 days. Scale bar, 1 cm. **(c)** Quantification of the number of NMuMG colonies expressing GFP-only or different FGFR2-GFP variants in soft agar over time. **(d)** Heat map showing the abundance of FGFR2 interacting proteins in NMuMG cells expressing the different FGFR2-GFP variants compared to GFP-expressing control cells. All FGFR2 interacting proteins with a SAINT probability score ≥ 0.95 are shown. The data represent the averaged spectral counts of three independent experiments. PSM, peptide-spectrum matches. **(e)** Coimmunoprecipitation of RSK1,2,3 in NMuMG cells expressing GFP-only or different FGFR2-GFP variants, as visualized by immunoblotting using an anti-RSK1,2,3 antibody.

To assess the tumorigenic potential of cells expressing the different FGFR2-GFP *in vitro*, we performed a soft agar colony formation assay with NMuMG cells expressing GFP, FGFR2-GFP, FGFR2(1-716)-GFP and FGFR2(1-672)-GFP. Using this approach, we show that only expression of FGFR2(1-672)-GFP enabled anchorage-independent growth in NMuMG cells confirming the ability of this truncated FGFR2 to induce cellular transformation of mammary epithelial cells (Figure 3b, 3c).

Altogether, these data indicate that the C-terminal tail of FGFR2 in a FGFR2-GFP fusion is important for the regulation of FGFR signaling pathway. As a consequence, FGFR signaling activation is prolonged when the C-terminal tail is removed, which results in the cellular transformation of mammary epithelial cells.

RSK2 binds to the C-terminus of FGFR2 in mammary epithelial cells expressing FGFR2-GFP variants

To identify potential FGFR2-binding proteins involved in suppressing cellular transformation, we performed immunoprecipitation with a GFP-specific antibody followed by liquid chromatography–tandem mass spectrometry (LC-MS/MS) analysis in NMuMG mammary epithelial cells expressing GFP, FGFR2-GFP, FGFR2(1-716)-GFP and FGFR2(1-672)-GFP (Supplementary Figure 2a). This analysis showed that all FGFR2-GFP variants were able to bind the fibroblast growth factor receptor substrate 2 (FRS2α), which is the adaptor protein of FGFR2 and activates downstream signaling (Figure 3d). Surprisingly, we could not confirm the previously observed binding of GRB2 to the last 10 amino acids of FGFR2[12]. Nevertheless, we did identify several receptor tyrosine kinases (RTKs), including fibroblast growth factor 1 (FGFR1), transforming growth factor receptor beta 2 (TGFBR2) and various ephrin receptor family members, suggesting that the FGFR2-GFP variants are able to interact with other RTKs in mammary epithelial cells. Similar results were obtained when the cells were serum-starved for 48 hours, indicating that these interactions were ligand-independent (Supplementary Figure 2b).

Interestingly, the binding of ribosomal protein S6 kinase A3 (RPS6KA3, also known as RSK2) was significantly reduced in the cells expressing FGFR2(1-672)-GFP compared to cells expressing FGFR2-GFP or FGFR2(1-716)-GFP (Figure 3e), indicating that RSK2 binds to FGFR2(673-716). The interaction of RSK2 with the C-terminal tail of FGFR1 has been reported to regulate receptor signaling via endocytosis[17], suggesting that RSK2 plays a similar role in regulating the activity of FGFR2 and could thereby suppress the cellular transformation of mammary epithelial cells. To test whether abrogation of RSK2 binding results in increased tumorigenic potential of FGFR2 in mammary epithelial cells, we generated FGFR2-GFP mutants that lack RSK2 binding by deleting the two conserved serine residues (S696 and S697 in FGFR2) in FGFR2-GFP and FGFR2(1-716)-GFP, which have been shown to be important for its binding to FGFR1[17]. First, we confirmed that the binding of RSK2 was significantly reduced in cells expressing FGFR2$^{\Delta S696/697}$-GFP or FGFR2(1-716)$^{\Delta S696/697}$-GFP compared to cells expressing FGFR2-GFP

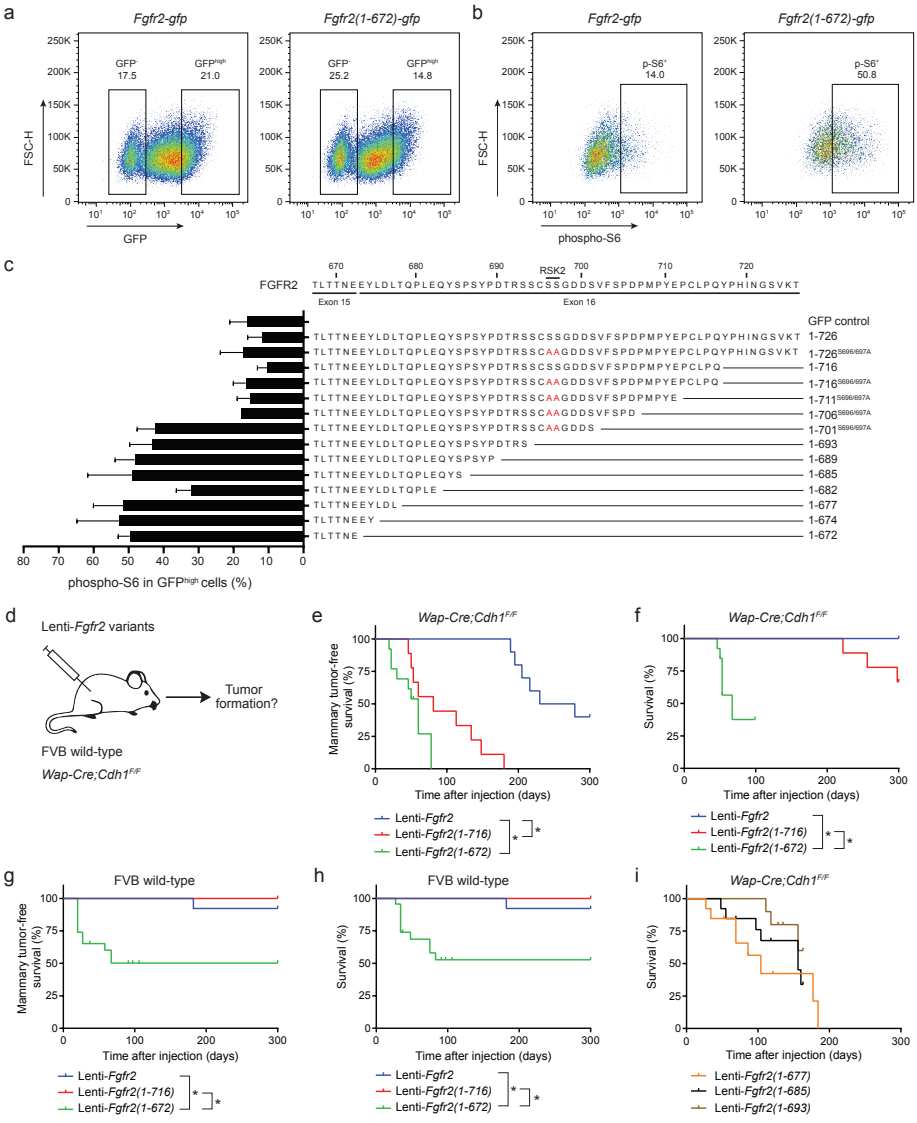
(Supplementary Figure 2c), indicating that RSK2 binds to FGFR2$^{S696/697}$. Next, NMuMG cells were transduced with lentiviruses encoding FGFR2$^{\Delta S696/697}$-GFP or FGFR2(1-716)$^{\Delta S696/697}$-GFP and GFP^{high} cells were separated using FACS. Finally, we tested the tumorigenic potential of these cells in a soft agar colony formation assay, which showed that the cells expressing FGFR2$^{\Delta S696/697}$-GFP or FGFR2(1-716)$^{\Delta S696/697}$-GFP were not able to induce anchorage-independent growth in NMuMG cells (Supplementary Figure 2d). Thus, abrogation of RSK2 binding alone is not sufficient to enable anchorage-independent growth of mammary epithelial cells expressing FGFR2-GFP, indicating that other factors contribute to the malignant transformation of these cells.

High levels of phosphorylated S6 determine the tumorigenic potential of FGFR2-GFP variants in mammary epithelial cells

To explore which other factors suppress oncogenicity of FGFR2 in NMuMG mammary epithelial cells, we first determined which part of the FGFR2 C-terminus is essential in suppressing cellular transformation of these cells. As increased levels of phosphorylated S6 (phospho-S6) coincided with the ability to show anchorage-independent growth in soft agar (Figure 3b, 3c), we investigated phospho-S6 levels induced by different FGFR2-GFP variants in mammary epithelial cells. We transduced NMuMG cells with lentiviruses encoding different FGFR2-GFP variants and measured the level of phospho-S6 in GFP^{high} cells using flow cytometry (Figure 4a-c and Supplementary Figure 3a-c). Our analysis included a GFP-only control, FGFR2-GFP, FGFR2(1-716)-GFP, FGFR2(1-672)-GFP and 11 additional FGFR2-GFP variants in which we progressively deleted amino acid residues from the C-terminus of FGFR2(1-716)-GFP. Furthermore, the two serine residues (S696 and S697) responsible for RSK2 binding were mutated to alanine residues to identify additional motifs in the C-terminal tail that suppress FGFR signaling. As expected, half of the GFP^{high} cells expressing FGFR2(1-672)-GFP were positive for phospho-S6 after 48 hours of serum starvation, whereas only 10-20% of the GFP^{high} cells expressing GFP-only, FGFR2-GFP and FGFR2(1-716)-GFP were positive.

Interestingly, high levels of S6 phosphorylation were still observed in GFP^{high} cells expressing FGFR2(1-701)$^{S696/697A}$-GFP, whereas this signal was decreased in cells expressing FGFR2(1-706)$^{S696/697A}$-GFP. These data suggest that the ^{702}VFSPD^{706}

Figure 4. ► Mammary epithelial cells expressing FGFR2(1-672)-GFP show an high level of S6 phosphorylation, which can be used as a marker for the tumorigenic potential of FGFR2 truncation variants in *Wap-Cre;Cdh1$^{F/F}$* and FVB wild-type female mice. (a) Representative flow cytometry dot plots of GFP expression in 48 hours serum-starved NMuMG cells transduced with Lenti-*Fgfr2-gfp* and Lenti-*Fgfr2(1-672)-gfp*, respectively. **(b)** Representative dot plots of phosphorylated S6 expression in GFP^{high} cells measured by flow cytometry. Serum-starved NMuMG cells transduced with Lenti-*Gfp* were used to determine basal expression of S6 phosphorylation (Supplementary Figure 3b). **(c)** Quantification of phospho-S6 expression in serum-starved NMuMG cells transduced with Lenti-*Gfp* or different Lenti-*Fgfr2-gfp* variants. The overview shows the amino acid residues of the FGFR2 C-terminal truncation variants fused in-frame to GFP. The two serine residues (S696 and S697) responsible for RSK2 binding were mutated to alanines (in red) in the corresponding FGFR2 truncation variants. Data are mean ± s.d. of three independent transductions. **(d)** Overview of the intraductal



injections of high-titer lentiviruses encoding different FGFR2 variants performed in *Wap-Cre;Cdh1$^{F/F}$* and FVB wild-type females. **(e)** The mammary tumor-free survival of glands injected with Lenti-*Fgfr2* (*n* = 10), Lenti-*Fgfr2(1-716)* (*n* = 9) and Lenti-*Fgfr2(1-672)* (*n* = 13) in *Wap-Cre;Cdh1$^{F/F}$* female mice is shown using a Kaplan Meier curve. **P* < 0.05 by Mantel–Cox test. **(f)** Kaplan Meier curve showing the survival until tumors exceeded the size of 200 mm^3 in the glands of *Wap-Cre;Cdh1$^{F/F}$* females injected with the indicated cDNA-encoding lentiviruses. **(g,h)** Kaplan Meier curves showing the mammary tumor-free survival **(g)** and survival until tumors exceeded the size of 200 mm^3 in the glands **(h)** of FVB wild-type females injected with Lenti-*Fgfr2* (*n* = 13), Lenti-*Fgfr2(1-716)* (*n* = 13) and Lenti-*Fgfr2(1-672)* (*n* = 23). **P* < 0.05 by Mantel–Cox test. (**i**) The survival until tumors exceeded the size of 200 mm^3 in the glands of *Wap-Cre;Cdh1$^{F/F}$* females injected with Lenti-*Fgfr2(1-677)* (*n* = 13), Lenti-*Fgfr2(1-685)* (*n* = 13) and Lenti-*Fgfr2(1-693)* (*n* = 12), as shown using a Kaplan Meier curve. The statistical analysis was conducted using the Mantel-Cox test and did not show any differences (*P* < 0.5).

sequence in FGFR2(1-706)$^{S696/697A}$-GFP is important in suppressing phospho-S6 levels in mammary epithelial cells. Furthermore, cells expressing FGFR2$^{S696/697A}$-GFP induced similar levels of phospho-S6 as GFP-only control cells. Thus, although RSK2 binding was clearly abrogated in cells expressing FGFR2(1-672)-GFP, these data indicate that other factors contribute to the cellular transformation of mammary epithelial cells.

Somatic expression of FGFR2 C-terminal truncation variants induces rapid mammary tumor formation in Wap–Cre;Cdh1$^{F/F}$ mice

To test whether FGFR2-GFP variants that show high levels of phospho-S6 in mammary epithelial cells could induce mammary tumor formation *in vivo*, we performed intraductal injections of lentiviruses encoding different FGFR2 C-terminal truncation variants in *Wap–Cre*;*Cdh1$^{F/F}$* and FVB female mice (Figure 4d). Within 11 weeks after injection, palpable tumors were detected in *Wap–Cre*;*Cdh1$^{F/F}$* mice injected with Lenti-*Fgfr2(1-716)* (n = 9) and Lenti-*Fgfr2(1-672)* (n = 13), whereas tumors developed sporadically in mammary glands injected with Lenti-*Fgfr2* (n = 10) 200 days post-injection (Figure 4e). Although we did not observe any difference in tumor initiation between the two *Fgfr2* truncation variants, the tumors induced by Lenti-*Fgfr2(1-672)* showed rapid proliferation resulting in a decreased median survival (67 days; when tumors reached the size of 200 mm^3) compared to the slow growing tumors induced by Lenti-*Fgfr2(1-716)* (Figure 4f and Supplementary Figure 3d). In addition, we only detected palpable tumors in mammary glands injected with Lenti-*Fgfr2(1-672)* (n = 23) in FVB wild-type mice (Figure 4g), which progressed rapidly over time (Figure 4h and Supplementary Figure 3e). These results confirm that FGFR2(1-672) has an increased tumorigenic potential compared to FGFR2(1-716) (n = 13), which is in line with our previous observations in *Wap-Cre;Cdh1$^{F/F}$* and *Wap-Cre;Cdh1$^{F/+}$* female mice carrying conditional *Fgfr2* knock-in alleles.

Next, we tested the *in vivo* oncogenicity of other FGFR2 C-terminal truncation variants that induce high phospho-S6 levels in mammary epithelial cells. To this end, we performed intraductal injections of Lenti-*Fgfr2(1-677)*, Lenti-*Fgfr2(1-685)* and Lenti-*Fgfr2(1-693)* in *Wap-Cre;Cdh1$^{F/F}$* and FVB wild-type mice. As expected, we detected palpable mammary tumors in *Wap-Cre;Cdh1$^{F/F}$* females and FVB wild-type mice (Supplemental Figure 3f-g), which progressed rapidly over time (Figure 4i and Supplementary Figure 3h).

Altogether, these data show that C-terminal FGFR2 truncation variants which induce a high level of phospho-S6 in mammary epithelial cells also promote formation of rapidly proliferating mammary tumors in both *Wap-Cre;Cdh1$^{F/F}$* and FVB wild-type mice. Therefore, our FACS-based assay to measure phospho-S6 levels induced by FGFR2 variants *in vitro* followed by assessing their tumorigenic potential *in vivo* provides a powerful approach to dissect which domains in the C-terminus of FGFR2 are essential for suppressing mammary tumor formation in mice.

Discussion

In this study, we found that *SB*-induced mILCs frequently carry transposon insertions in the last intron of *Fgfr2*, leading to premature transcriptional termination and expression of FGFR2 truncation variants lacking the last 54 amino acids. In contrast to full-length FGFR2, mammary-specific expression of two different FGFR2 truncation variants induced rapid ILC formation in *Wap-Cre;Cdh1*$^{F/F}$ mice. Intriguingly, removal of the last 54 amino acid residues resulted in an FGFR2 truncation variant that was more potent in driving mammary tumorigenesis compared to a variant that only lacked the last 10 amino acids. Furthermore, intraductal injections of lentiviruses encoding different C-terminal FGFR2 truncation variants in *Wap-Cre;Cdh1*$^{F/F}$ mice induced mammary tumors with different growth kinetics, indicating that the oncogenicity of FGFR2 is modulated by multiple domains in the C-terminus.

Mutations in oncogenes and tumor suppressor genes are typically expected to result in activation or inactivation of their respective genes. However, various mutations in tumor suppressor genes might also result in mutant proteins with distinctive roles in tumor cells, as exemplified by the different types of outcome for mutations in *TP53*. Hotspot mutations in *TP53* may result in the expression of mutant p53 isoforms that, in addition to exerting dominant-negative effects on wildtype p53, gain novel oncogenic properties[18,19]. These effects are not readily identifiable in human tumors and therefore require complementary approaches to unravel how genes may acquire their oncogenic potential. Insertional mutagenesis screens in mice can – in addition to identifying novel cancer drivers and pathways – provide biological insight into mechanisms of gene activation or inactivation. For example, retroviral insertional mutagenesis screens in *p53*$^{-/-}$, *p19*$^{ARF-/-}$ and wild-type mice induced tumors with clustered insertions within *Flt3* and *Notch1*, which resulted in the expression of N-terminal truncated oncogenic mutants with distinct biological properties[20]. The NOTCH1 mutants specifically lacked the destabilizing COOH-terminal PEST-domain, which may lead to impaired degradation and enhanced NOTCH1 signaling. Similarly, intragenic transposon insertions in *Ppp1r12a* and *Trp53bp2* in *SB*-induced mILCs result in the expression of truncated PPP1R12A and TRP53BP2 proteins that lack several negative regulatory phosphorylation sites, which could result in enhanced phosphatase activity[14]. In this study, we observed a strong clustering of transposon insertions in the last intron of *Fgfr2*, indicating that ILC formation is induced by expression of truncated FGFR2 rather than full-length FGFR2. These data suggest that tumor formation is suppressed by inhibitory domains within the C-terminus of FGFR2. These studies highlight insertional mutagenesis in mice as a powerful tool to identify distinct mechanisms of gene activation or inactivation involved in tumorigenesis.

In human cancer, an important class of cancer drivers involves the activation of kinases due to mutations, amplifications or gene fusions. Several studies focusing on the identification of gene fusions involving kinases have uncovered various fusions with FGFR family members in different types of cancer, including glioblastoma multiforme (GBM), breast cancer, cholangiocarcinoma and bladder

cancer[5–7,21]. *FGFR3-TACC3* fusions are mainly observed in small subsets of GBMs and bladder carcinomas[5,21], whereas *FGFR2* is fused to a diversity of 3' fusion partners in cholangiocarcinoma and breast cancer[6]. Overall, FGFR fusions are characterized by an intact kinase domain fused in-frame with various 3' partners, which thereby remove the C-terminal tail of the receptor. Several studies have proposed different mechanisms for the oncogenic activity of FGFR fusion proteins, which mainly involve increased oligomerization due to the presence of known dimerization motifs in the fusion partners[6]. However, we postulate that oncogenicity of FGFR2 fusions may also derive from removal of the C-terminal tail of FGFR2. In line with this, we find that expression of C-terminal FGFR2 truncation variants induces rapid and multifocal mammary tumor formation in mice, whereas full-length FGFR2 does not induce tumor formation. Our notion is further supported by previous studies showing that alternative splicing in human cancer cell lines results in the expression of FGFR2 isoforms lacking the C-terminal tail, which enhance the transforming activity of cells *in vitro*[8–11]. Altogether, these findings indicate that truncated FGFR2 proteins – independent of a 3' fusion partner – promote cellular transformation. Therefore, we propose that oncogenicity of FGFR2 fusion proteins is primarily due to the absence of the C-terminus rather than presence of the 3' fusion partner. Testing of this hypothesis requires comparison of the oncogenic potential of full-length and C-terminally truncated FGFR2 fused to the same 3' fusion partner.

Molecular characterization of human ILCs (hILCs) has shown frequent activation of RAS-MAPK and PI3K-AKT signaling pathways in these tumors[22–24]. Although these large studies did not reveal frequent hyper-activation of FGFR-related genes, *FGFR1* amplification has been observed in a subset of hILCs[25]. Alternative splicing might constitute yet another mechanism to generate oncogenic FGFR2 truncation variants[8–11]. Together, these studies warrant in-depth analysis of RNA-sequencing data from hILCs to investigate whether these tumors might reveal rearrangements, fusions and/or alternative splicing of *FGFR* genes resulting in the expression of truncated proteins lacking the C-terminal tail.

Future studies are required to identify the domains in the C-terminal tail of FGFR2 that are essential in suppressing oncogenic FGFR signaling, and which interactors or mechanisms could be involved in this process. These studies will require combined proteomic and functional analyses of large numbers of FGFR2 mutants. For this purpose, we developed a FACS-based assay to measure the S6 phosphorylation levels induced by FGFR2 variants in mammary epithelial cells, which correlated with their tumorigenic potential in mice. Using this sensitive approach, it should be possible to identify the minimal mutations that are required to generate a full-length FGFR2 mutant with a comparable oncogenic potential as FGFR2(1-672). A similar approach combining flow cytometry with structural modeling has been successfully used to identify several inhibitory interactions of the C-terminal tail and the kinase domain of the epidermal growth factor receptor (EGFR)[26], which are predicted to be lost in lung tumors driven by EGFR fusions[27]. Although two out of three amino acid sequences responsible for the inhibitory interactions in EGFR are not observed in the C-terminal tail of FGFR2, similar

conformational changes underlying these interactions might also be involved in regulating FGFR signaling. The third inhibitory interaction involves the clathrin adaptor AP2, which is responsible for endocytosis of the receptor. Although the first of the two putative AP2 binding motifs (^{674}YXXL^{677}) in the FGFR2 C-terminus seems to be nonessential for tumor suppression in our study, the lack of the second motif (^{717}YXXI^{720}) in mice expressing FGFR2(1-716) might explain the formation of ILCs in *Wap-Cre;Cdh1$^{F/F}$* mice.

In summary, we provide strong evidence that C-terminal truncation variants of FGFR2 can drive mammary tumor formation in mice, which emphasizes the role of the C-terminal tail in regulating FGFR signaling and highlights the utility of insertional mutagenesis in mice to uncover novel mechanisms of gene activation involved in human cancer development. Our findings also indicate that oncogenicity of FGFR2 fusion proteins observed in human tumors is caused by loss of the C-terminal tail rather than the presence of the 3' fusion partner.

Acknowledgements

We are grateful to Micha Nethe, Joppe Nieuwenhuis, Stefano Annunziato, Tanya Braumuller and Linda Henneman for providing technical suggestions and/or assistance with the experiments. We thank Peter Bouwman for critical reading of the manuscript. We thank the NKI animal facility, RHPC computing facility, the animal pathology facility, the mouse clinic transgenic and intervention unit, the core facility molecular pathology and biobanking (CFMPB), and the genomics core facility for their expert technical support. This work was carried out on the Dutch national e-infrastructure with the support of SURF Cooperative (e-infra160136). Financial support was provided by the Netherlands Organization for Scientific Research (NWO: Cancer Genomics Netherlands (CGCNL), Cancer Systems Biology Center (CSBC), Netherlands Genomics Initiative (NGI) grant Zenith 93512009, infrastructural grant from the National Roadmap for Large-Scale Research Facilities, Proteins@Work (184.032.201)), the EU Seventh Framework Program (Infrafrontier-I3 project 312325), and the European Research Council (ERC Synergy project CombatCancer). This work is part of the Oncode Institute which is partly financed by the Dutch Cancer Society (KWF).

Materials and methods

Generation of mouse models
The *Fgfr2* variants were isolated from *Fgfr2* (NM_201601) cDNA (OriGene, MC221076), sequence verified and inserted with FseI-PmeI fragments into the *Frt-invCag-IRES-Luc* vector (shuttle vector) using the following primer sequences: forward 5'-AAAAGGCCGGCCATGGGATTAC-3' combined with reverse *Fgfr2* 5'-AAAAGTTTAAACTCATGTTTTAACACTGCC-3'; reverse *Fgfr2*$^{ex1\text{-}16\Delta30bp}$ 5'-AAAAGTTTAAACTCACTGAGGCAGACAGGG-3'; reverse *Fgfr2*$^{ex1\text{-}15}$ 5'-AAAAGTTTAAACTCACTCATTGGTTGTGAG-3', resulting in *Frt-invCAG-Fgfr2(1-726)-IRES-Luc*, *Frt-invCAG-Fgfr2(1-716)-IRES-Luc* and *Frt-invCAG-Fgfr2(1-672)-IRES-Luc* alleles, respectively. Flp-mediated integration of the shuttle vectors in *WapCre;Cdh1*$^{F/F}$;*Col1a1*$^{frt/+}$ GEMM-ESC clones and subsequent blastocyst injections of the modified ESCs were performed as previously described[16]. Chimeric animals were mated with *Cdh1*$^{F/F}$ and *Cdh1*$^{F/+}$ animals to generate the experimental cohorts. *WapCre* and *Cdh1*^{F} alleles were detected using PCR as previously described[28,29]. The *Col1a1*$^{invCAG\text{-}Fgfr2(1\text{-}726)\text{-}IRES\text{-}Luc/+}$, *Col1a1*$^{invCAG\text{-}Fgfr2(1\text{-}716)\text{-}IRES\text{-}Luc/+}$, *Col1a1*$^{invCAG\text{-}Fgfr2(1\text{-}672)\text{-}IRES\text{-}Luc/+}$ and wild-type alleles were detected using standard PCR with an annealing temperature of 58°C using the following primer sequences: *Col1a1-Fgfr2* forward 5'-TGGCCAGGGATATCAACAAC-3' and reverse 5'-ACACCGGCCTTATTCCAAGC-3' (585 bp, 554 and 420 bp, respectively); wild-type forward 5'-CTCGCACGTACTTCATTC-3' and reverse 5'-CTGCTTGAATCCCTTTGAG-3' (234 bp).

The experimental cohorts were monitored twice a week and the mammary tumor-free survival was scored (event) when the first palpable tumor was detected, whereas mice that did not develop any mammary tumors during life were censored. Mammary tumor-specific survival was scored for mice that were sacrificed when the total mammary tumor burden reached a size of 1500 mm^3 (tumor volume: length × width^2 × 0.5) or the mice suffered from clinical signs of distress caused by the mammary tumor burden or metastatic disease (such as respiratory distress, ascites, distended abdomen, rapid weight loss and severe anemia). Mice that were sacrificed due to other circumstances were censored. Lungs, heart, liver, spleen, kidneys, mammary glands and tumor-draining lymph-nodes were collected and analyzed for histological abnormalities. All 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.

Lentiviral vectors and virus production
The Lenti-*Fgfr2-gfp* variants were constructed as follows. The three different *Fgfr2* variants were isolated with NotI-FseI fragments from *Fgfr2* (NM_201601) cDNA (OriGene, MC221076). *Gfp* was isolated with FseI-XhoI overhangs from the SIN. LV.SF vector (Lenti-GFP)[30], which was a gift of Eugenio Montini. *Gfp* and the different *Fgfr2* variants were simultaneously cloned into the pcDNA™3.1/Zeo $^{(+)}$ Mammalian Expression Vector (ThermoFisher, V86020). The FseI sequence in the different pcDNA™3.1-*Fgfr2-gfp* variants was removed using site-directed mutagenesis

(Agilent QuikChange Lightning mutagenesis kit) according to manufacturer's protocol. The resulting *Fgfr2-gfp* cDNAs were isolated with AgeI-SalI overhangs and inserted into the SIN.LV.SF vector, resulting in SIN.LV.SF-*Fgfr2-gfp* (Lenti-*Fgfr2-gfp*), SIN.LV.SF-*Fgfr2(1-716)-gfp* (Lenti-*Fgfr2(1-716)-gfp*) and SIN.LV.SF-*Fgfr2(1-672)-gfp* (Lenti-*Fgfr2(1-672)-gfp*). The Lenti-*Fgfr2* variants were constructed as follows. The *Fgfr2* variants were isolated with AgeI-SalI overhangs from *Fgfr2* cDNA and inserted into the SIN.LV.SF vector, resulting in SIN.LV.SF-*Fgfr2* (Lenti-*Fgfr2*), SIN.LV.SF-*Fgfr2(1-716)* (Lenti-*Fgfr2(1-716)*) and SIN.LV.SF-*Fgfr2(1-672)* (Lenti-*Fgfr2(1-672)*). All primers are listed in Supplementary Table 1. The mutations and nucleotide deletions/additions in the different lentiviral vectors were generated using site-directed mutagenesis according to manufacturer's protocol. The primers used for site-directed mutagenesis were designed using the QuikChange Primer Design Program. All cDNAs were isolated using Q5 High-Fidelity DNA Polymerase (NEB) according to manufacturer's protocol and verified using Sanger sequencing. Concentrated lentiviral stocks were produced by transient co-transfection of four plasmids in 293T cells as previously described[31]. Viral titers were determined using the qPCR lentivirus titration kit (Abm).

Intraductal injections

Intraductal injections of high-titer lentiviruses were performed as previously described[32,33]. All lentiviral stocks used had viral titers ranging between 2 x 10^8 transducing units (TU)/mL to 2 x 10^9 TU/mL.

In vivo bioluminescence imaging

In vivo bioluminescence imaging of luciferase expression was performed as previously described[34]. Signal intensity was measured on the whole body of the mouse (excluding the head and tail) using a size-fixed square and quantified as flux (photons per seconds per cm^2 per steradian).

Histology and immunohistochemistry

Tissues were formalin-fixed and paraffin-embedded (FFPE) by routine procedures. Immunohistochemical stainings of E-cadherin, vimentin and cytokeratin-8 were processed as previously described[34,35]. All slides were digitally processed using the Aperio ScanScope (Aperio, Vista, CA, USA) and captured using ImageScope software version 12.3.2.8013 (Aperio). The mammary glands were reviewed by a European College of Veterinary Pathologists (ECVP) certified veterinary pathologist (Sjoerd Klarenbeek) in a blinded manner according to international consensus of mammary pathology[36].

Cell culture

NMuMG cells were cultured in DMEM-F12/Glutamax medium containing 10% fetal bovine serum (FBS), 100 IU/mL penicillin, 100 µg/mL streptomycin (all from Life Technologies). 293T cells were cultured in Iscove's medium (Life Technologies) containing 10% FBS, 2mM glutamine, 100 IU/ml penicillin and 100 µg/ml streptomycin. All cell lines were routinely tested for Mycoplasma contamination using the MycoAlert mycoplasma detection kit (Lonza).

FACS-sorting of GFP^{high} NMuMG cells expressing GFP or different FGFR2-GFP variants

NMuMG cells were transduced with diluted viral supernatants at equal multiplicity of infections (MOIs) in the presence of 8 µg/mL polybrene (Sigma). At least 6 days after transduction, the cells were collected using trypsine, resuspended in PBS containing 1% FBS. The GFP^{high} cells were sorted on an Astrios Moflo sorter system. The sorted cells were collected in DMEM-F12/Glutamax medium containing 10% FBS, 100 IU/mL penicillin, 100 µg/mL streptomycin and allowed to recover for at least 6 days before they were used in the experiments.

Soft agar colony formation assay

The soft agar assays were performed as previously described[37]. After the base layer was solidified overnight at 4°C, 25000 NMuMG cells expressing GFP or the different FGFR2-GFP variants were suspended in 2.5 ml of 0.35% low-gelling temperature agarose (Sigma-Aldrich) in DMEM-F12/Glutamax medium containing 10% FBS, 100 IU/mL penicillin, 100 µg/mL streptomycin. Anchorage-independent growth was assessed by counting colonies using the GelCount instrument (Oxford Optronix). All soft agar assays were performed in at least three independent experiments, using at least two independently transduced and FACS-sorted cell cultures.

Immunoblotting

Protein lysates were made using lysis buffer (25 mM Tris-HCl pH 7.6, 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS in milli-Q) complemented with protease and phosphatase inhibitors (Roche) and quantified using the BCA protein assay kit (Pierce). Equal amounts of proteins were separated on a 4-12% Bis-Tris gradient gel (Invitrogen) and transferred overnight onto nitrocellulose membrane (Bio-Rad) in 1x transfer buffer (25 mM Tris, 2 M Glycine, 20% methanol in demineralized water). Membranes were blocked in 5% w/v bovine serum albumin (BSA) in PBS-T (pH 7.5, 0.005% Tween-20 in demineralized water) and incubated overnight using the following primary antibodies: FGFR2 (1:1000, GeneTex 10648), phospho-FGFR (1:1000, CST 3471), FRS2 (1:1000, ProteinTech 11503-1-AP), phospho-FRS2 (Tyr436) (1:1000, Abcam 193363), AKT1 (1:1000, Cell Signaling Technology [CST] 2938), phospho-AKT(Ser473) (1:1000, CST 4060), p44/42 MAP kinase (1:1000, CST 4695), phospho-p44/42 MAPK ERK1/ERK2 (Thr202/Tyr204) (1:1000 CST 9101), S6 (1:1000, CST 2217), phospho-S6 (1:1000, CST 2211), RSK1,2,3 (1:1000, CST 9355), GFP (1:1000, Roche 11814460001) and β-actin (1:50000, Sigma A5441) in 5% w/v BSA in PBS-T. Membranes were washed three times and incubated with the secondary antibodies goat anti-rabbit-HRP (1:2000, Dako P0448), rabbit anti-mouse-HRP (1:5000, Dako P0260), or donkey anti-mouse IRDye 680nm (1:5000, Li_COR 926-32222, light protected) in 5% w/v BSA in PBS-T. Stained membranes were washed three times in PBS-T and then developed using ECL (Pierce 32209) or captured using the Li-Cor Odyssey Infrared Imaging System and analyzed using Odyssey Application software version 3.0.16.

Flow cytometry and intracellular staining of phospho-S6

NMuMG cells were transduced with diluted viral supernatants of expression vectors encoding GFP or different FGFR2-GFP variants at equal MOIs in the presence of 8 µg/mL polybrene (Sigma). Three days after transduction, 500.000 cells were seeded per 6-well in DMEM-F12/Glutamax medium containing 10% FBS, 100 IU/mL penicillin, 100 µg/mL streptomycin. After 24 hours, the cells were washed with PBS and serum-free DMEM-F12/Glutamax medium was added for 48 hours. To measure the expression of phospho-S6 using flow cytometry, the cells were processed as previously described[38]. Briefly, the cells were harvested using trypsin and collected in DMEMF12/glutamax containing 1% FBS. After centrifugation at 2000 rpm for 3 min, the pellet was washed with PBS. The cells were pelleted again to discard the PBS and fixed with 100 µl of BD fix buffer I (BD Biosciences) at 37°C for 10 min. Cells were washed with PBS containing 10% FBS (FACS buffer), pelleted using centrifugation and subsequently permeabilized for 30 min on ice using 100 µl of ice-cold BD permeabilization buffer III (BD Biosciences). After washing and pelleting the cells twice with FACS-buffer, the cells were incubated with the phospho-S6 antibody (1:400, CST 4856) in FACS-buffer for 1 hour rotating at room temperature. The suspension was washed with FACS-buffer and pelleted twice before the cells were incubated with anti-rabbit AlexaFluor 647 (1:500, Invitrogen A-21244) in FACS-buffer for 1 hour rotating at room temperature, protected from light. The cells were washed and pelleted twice, resuspended in FACS buffer and 50.000 single cells were analyzed per sample, which were gated on size and shape using forward and side scatter. All experiments were performed using a BD LSRII flow cytometer using Diva software. Data analyses were performed using FlowJo Software version 10.4.2.

GFP immunoprecipitation and LC/MS-MS analysis

NMuMG cells expressing GFP, FGFR2-GFP, FGFR2(1-716)-GFP and FGFR2(1-672)-GFP were seeded in 10 cm dishes until they were 80% confluent (normal condition), or the cells were starved for 48 hours with serum-free medium when the dishes were 60% confluent. Protein lysates were made using a NP40 lysis buffer (50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1% NP40 and 10% Glycerol). GFP-Trap beads (ChromoTek, Gta-20) were prepared according to manufacturer's protocol. The supernatant was incubated overnight with 25 µl of beads and after the incubation, the beads were washed five times with NP40 lysis buffer. The beads were reconstituted in sample loading buffer and heated at 95^0C for 7 min. The eluates were run into the stacking of a 4-12% Bis-Tris gel and coomassie-stained bands were excised for LC/MS-MS analysis. Proteins were reduced with 6.5mM DTT, alkylated with 54 mM iodoacetamide and digested in-gel with trypsin (Gold, mass spectrometry grade, Promega, 3 ng/µL) overnight at 37^0C. Extracted peptides were vacuum dried, reconstituted in 10% formic acid and analyzed by nanoLC-MS/MS on an Orbitrap Fusion Tribrid mass spectrometer equipped with a Proxeon nLC1000 system (Thermo Scientific). Peptides were loaded directly on the analytical column (Agilent Poroshell EC-C18 120 2.7 µm, 50 µmx500 mm, packed in-house) and separated in a 90-min gradient containing a 74-min linear increase from 6-30% solvent B (0.1% formic acid/80% acetonitrile), with 0.1% formic acid/water as Solvent A. Further settings were as described previously[39].

Raw data files were processed using Proteome Discoverer (version 2.2.0.388) (Thermo Scientific) using default settings unless stated otherwise. MS/MS spectra were searched against the *M. musculus* Swissprot database (16,954 entries, release 2017_12) using the Mascot search engine (Matrix Science, version 2.6.1), with trypsin chosen as cleavage specificity allowing two missed cleavages and with carbamidomethylation (C) and oxidation (M) set as fixed and variable modification, respectively. A decoy database defined in the Percolator node was used to validate and filter Peptide-Spectrum Matches (PSMs) <1% FDR. The protein PSM data from all IP samples (4 constructs, 3 biological replicates per construct) were combined by PD2.2 in a multiconsensus report, which was exported for determination of potential FGFR2 protein interactors using the SAINT computational tool through the CRAPome interface (http://www.crapome.org)[40,41]. SAINT interaction probability ≤ 0.95 was taken as cutoff to discriminate between true and false interactions.

Nucleic acid isolation
DNA and RNA were isolated from frozen mammary tumor pieces as previously described[14,42].

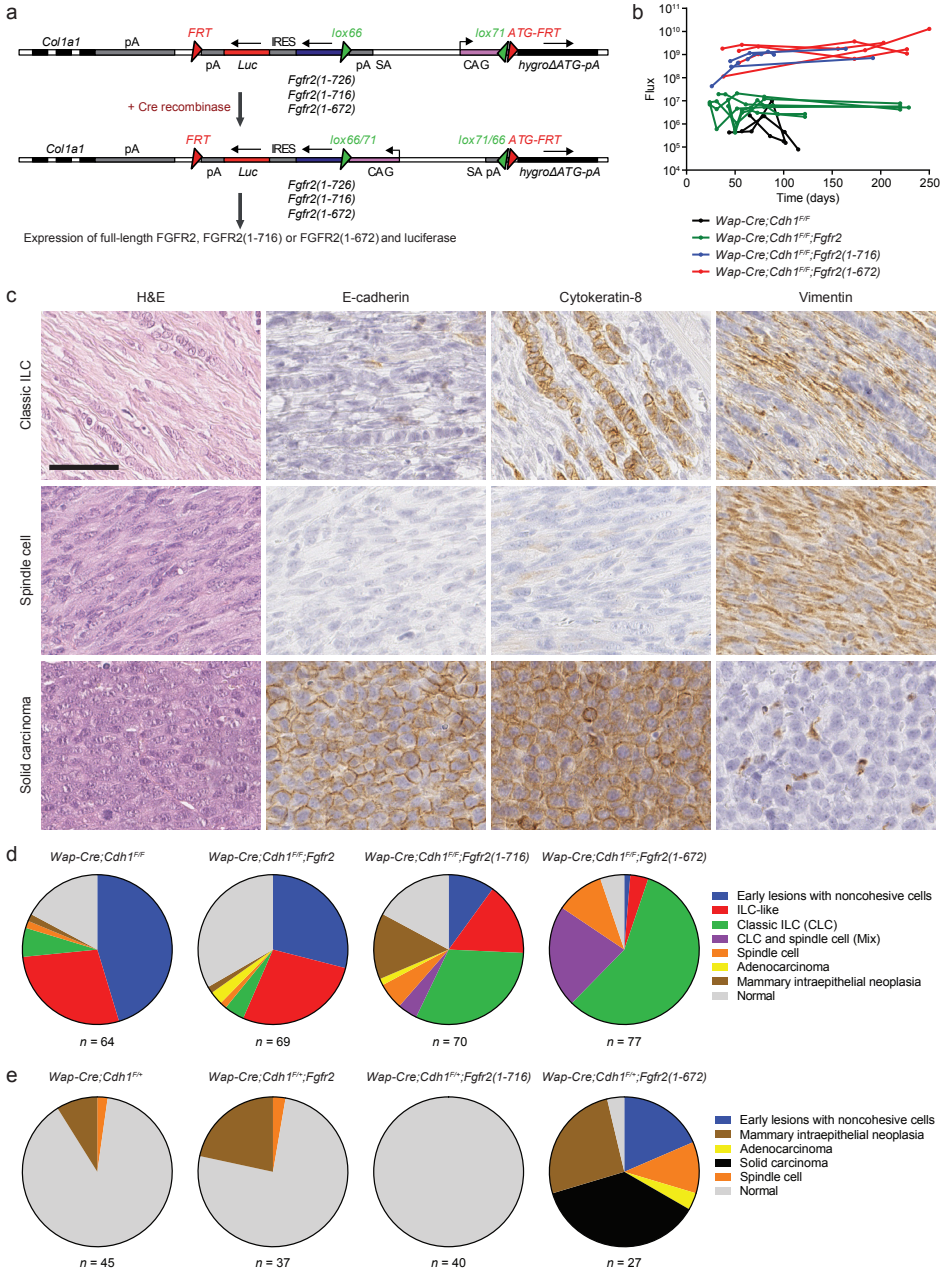
Mapping of insertion sites and the identification of common insertion sites
Genomic DNA from the *SB*-induced tumors was processed using the ShearSplink protocol and transposon insertions were amplified as previously described[43]. The insertions were aligned to the genome and the common integration sites were identified as previously described[14].

Quantitative RT-PCR
Tumor RNA (1000 ng) was converted to complementary DNA (cDNA) with a Moloney murine leukemia virus reverse transcriptase using Oligo(dT) primers according to manufacturer's protocol (Tetro cDNA synthesis kit, Bioline). Real-Time PCR was performed using the SensiFast SYBR Hi-ROX kit (Bioline) on a QuantStudio 6 flex system (Applied Biosystems) with QuantStudio™ Real-Time PCR Software (Version 1.1). The expression levels were determined using the comparative C_T ($\Delta\Delta C_T$) method and values were corrected for HPRT. Three independent experiments were performed in which the samples were measured in triplicate. The following primer sequences were used: *Fgfr2* exon 12 forward 5´-GCCAGAAACGTGTTGGTAAC-3´, *Fgfr2* exon 13 reverse 5´-TTCAGGAGCCATCCACTTG-3´, *Fgfr2* exon 15 forward 5´-GGATCGAATT CTGACTCTCAC-3´, *Fgfr2* exon 16 reverse 5´-GGGTTCATAAGGCATGGG-3´, *Hprt* forward 5´-CTGGTGAAAAGGACCTCTCG-3´, *Hprt* reverse 5´-TGAAGTACT CATTATAGTCAAGGGCA-3´.

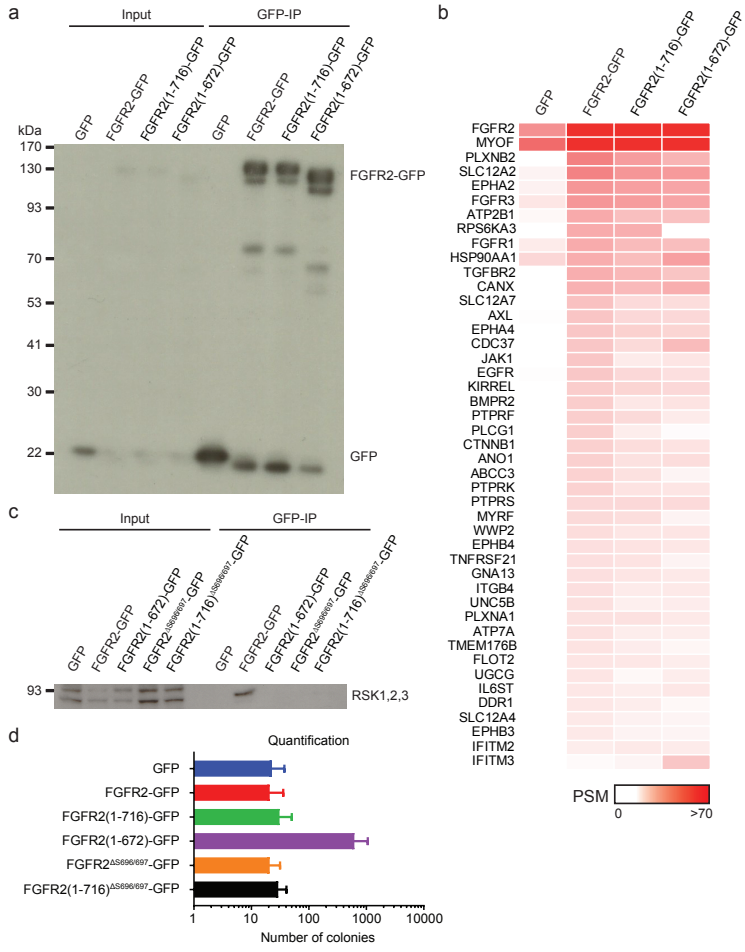
References

1. Blume-Jensen, P. & Hunter, T. Oncogenic kinase signalling. *Nature* **411,** 355–365 (2001).
2. Ornitz, D. M. & Itoh, N. The Fibroblast Growth Factor signaling pathway. *Wiley Interdiscip. Rev. Dev. Biol.* **4,** 215–266 (2015).
3. Turner, N. & Grose, R. Fibroblast growth factor signalling: from development to cancer. *Nat. Rev. Cancer* **10,** 116–129 (2010).
4. Babina, I. S. & Turner, N. C. Advances and challenges in targeting FGFR signalling in cancer. *Nat. Rev. Cancer* **17,** 318–332 (2017).
5. Singh, D. *et al.* Transforming fusions of FGFR and TACC genes in human glioblastoma. *Science* **337,** 1231–1235 (2012).
6. Wu, Y.-M. *et al.* Identification of targetable FGFR gene fusions in diverse cancers. *Cancer Discov.* **3,** 636–647 (2013).
7. Stransky, N., Cerami, E., Schalm, S., Kim, J. L. & Lengauer, C. The landscape of kinase fusions in cancer. *Nat. Commun.* **5,** 4846 (2014).
8. Itoh, H. *et al.* Preferential alternative splicing in cancer generates a K-sam messenger RNA with higher transforming activity. *Cancer Res.* **54,** 3237–3241 (1994).
9. Ishii, H., Yoshida, T., Oh, H., Yoshida, S. & Terada, M. A truncated K-sam product lacking the distal carboxyl-terminal portion provides a reduced level of autophosphorylation and greater resistance against induction of differentiation. *Mol. Cell. Biol.* **15,** 3664–3671 (1995).
10. Cha, J. Y., Lambert, Q. T., Reuther, G. W. & Der, C. J. Involvement of fibroblast growth factor receptor 2 isoform switching in mammary oncogenesis. *Mol. Cancer Res. MCR* **6,** 435–445 (2008).
11. Cha, J. Y., Maddileti, S., Mitin, N., Harden, T. K. & Der, C. J. 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 (2009).
12. Ahmed, Z. *et al.* Direct binding of Grb2 SH3 domain to FGFR2 regulates SHP2 function. *Cell. Signal.* **22,** 23–33 (2010).
13. Lin, C.-C. *et al.* Inhibition of basal FGF receptor signaling by dimeric Grb2. *Cell* **149,** 1514–1524 (2012).
14. Kas, S. M. *et al.* Insertional mutagenesis identifies drivers of a novel oncogenic pathway in invasive lobular breast carcinoma. *Nat. Genet.* **49,** 1219–1230 (2017).
15. de Ruiter, J. R. *et al.* Identifying transposon insertions and their effects from RNA-sequencing data. *Nucleic Acids Res.* **45,** 7064–7077 (2017).
16. Huijbers, I. J. *et al.* Using the GEMM-ESC strategy to study gene function in mouse models. *Nat. Protoc.* **10,** 1755–1785 (2015).
17. Nadratowska-Wesolowska, B. *et al.* RSK2 regulates endocytosis of FGF receptor 1 by phosphorylation on serine 789. *Oncogene* **33,** 4823–4836 (2014).
18. Lang, G. A. *et al.* Gain of function of a p53 hot spot mutation in a mouse model of Li-Fraumeni syndrome. *Cell* **119,** 861–872 (2004).
19. Olive, K. P. *et al.* Mutant p53 gain of function in two mouse models of Li-Fraumeni syndrome. *Cell* **119,** 847–860 (2004).
20. Uren, A. G. *et al.* Large-scale mutagenesis in p19(ARF)- and p53-deficient mice identifies cancer genes and their collaborative networks. *Cell* **133,** 727–741 (2008).
21. Williams, S. V., Hurst, C. D. & Knowles, M. A. Oncogenic FGFR3 gene fusions in bladder cancer. *Hum. Mol. Genet.* **22,** 795–803 (2013).
22. Ciriello, G. *et al.* Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell* **163,** 506–519 (2015).
23. Michaut, M. *et al.* Integration of genomic, transcriptomic and proteomic data identifies two biologically distinct subtypes of invasive lobular breast cancer. *Sci. Rep.* **6,** 18517 (2016).
24. Desmedt, C. *et al.* Genomic Characterization of Primary Invasive Lobular Breast Cancer. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **34,** 1872–1881 (2016).
25. Reis-Filho, J. S. *et al.* FGFR1 emerges as a potential therapeutic target for lobular breast carcinomas. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **12,** 6652–6662 (2006).
26. Kovacs, E. *et al.* Analysis of the Role of the C-Terminal Tail in the Regulation of the Epidermal Growth Factor Receptor. *Mol. Cell. Biol.* **35,** 3083–3102 (2015).
27. Konduri, K. *et al.* EGFR Fusions as Novel Therapeutic Targets in Lung Cancer. *Cancer Discov.* **6,** 601–611 (2016).
28. Derksen, P. W. B. *et al.* 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 (2006).
29. Derksen, P. W. B. *et al.* Mammary-specific inactivation of E-cadherin and p53 impairs functional

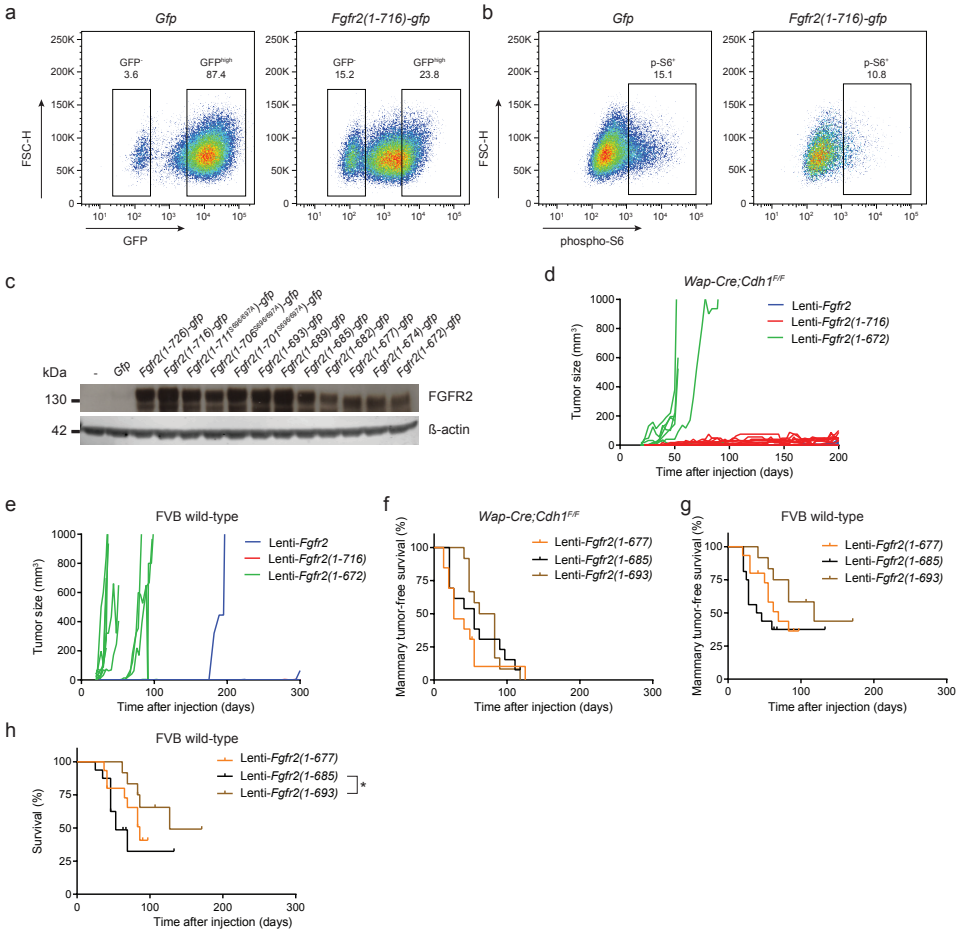
gland development and leads to pleomorphic invasive lobular carcinoma in mice. *Dis. Model. Mech.* **4,** 347–358 (2011).
30. Montini, E. *et al.* The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. *J. Clin. Invest.* **119,** 964–975 (2009).
31. Follenzi, A., Ailles, L. E., Bakovic, S., Geuna, M. & Naldini, L. Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences. *Nat. Genet.* **25,** 217–222 (2000).
32. Krause, S., Brock, A. & Ingber, D. E. Intraductal injection for localized drug delivery to the mouse mammary gland. *J. Vis. Exp. JoVE* (2013). doi:10.3791/50692
33. Annunziato, S. *et al.* Modeling invasive lobular breast carcinoma by CRISPR/Cas9-mediated somatic genome editing of the mammary gland. *Genes Dev.* **30,** 1470–1480 (2016).
34. Henneman, L. *et al.* 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 (2015).
35. Doornebal, C. W. *et al.* A preclinical mouse model of invasive lobular breast cancer metastasis. *Cancer Res.* **73,** 353–363 (2013).
36. Cardiff, R. D. *et al.* The mammary pathology of genetically engineered mice: the consensus report and recommendations from the Annapolis meeting. *Oncogene* **19,** 968–988 (2000).
37. Ikink, G. J., Boer, M., Bakker, E. R. M. & Hilkens, J. IRS4 induces mammary tumorigenesis and confers resistance to HER2-targeted therapy through constitutive PI3K/AKT-pathway hyperactivation. *Nat. Commun.* **7,** 13567 (2016).
38. Brockmann, M. *et al.* Genetic wiring maps of single-cell protein states reveal an off-switch for GPCR signalling. *Nature* **546,** 307–311 (2017).
39. Ameziane, N. *et al.* A novel Fanconi anaemia subtype associated with a dominant-negative mutation in RAD51. *Nat. Commun.* **6,** 8829 (2015).
40. Mellacheruvu, D. *et al.* The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. *Nat. Methods* **10,** 730–736 (2013).
41. Choi, H. *et al.* SAINT: probabilistic scoring of affinity purification-mass spectrometry data. *Nat. Methods* **8,** 70–73 (2011).
42. Boelens, M. C. *et al.* 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 (2016).
43. Koudijs, M. J. *et al.* High-throughput semiquantitative analysis of insertional mutations in heterogeneous tumors. *Genome Res.* **21,** 2181–2189 (2011).


a
Col1a1
pA
FRT
IRES
lox66
lox71 ATG-FRT
pA
Luc
pA SA
CAG
hygroΔATG-pA
Fgfr2(1-726)
Fgfr2(1-716)
Fgfr2(1-672)
+ Cre recombinase
lox66/71
lox71/66 ATG-FRT
SA pA
Expression of full-length FGFR2, FGFR2(1-716) or FGFR2(1-672) and luciferase
b
Flux
Time (days)
Wap-Cre;Cdh1F/F
Wap-Cre;Cdh1F/F;Fgfr2
Wap-Cre;Cdh1F/F;Fgfr2(1-716)
Wap-Cre;Cdh1F/F;Fgfr2(1-672)
c
H&E
E-cadherin
Cytokeratin-8
Vimentin
Classic ILC
Spindle cell
Solid carcinoma
d
Wap-Cre;Cdh1F/F
Wap-Cre;Cdh1F/F;Fgfr2
Wap-Cre;Cdh1F/F;Fgfr2(1-716)
Wap-Cre;Cdh1F/F;Fgfr2(1-672)
n = 64
n = 69
n = 70
n = 77
Early lesions with noncohesive cells
ILC-like
Classic ILC (CLC)
CLC and spindle cell (Mix)
Spindle cell
Adenocarcinoma
Mammary intraepithelial neoplasia
Normal
e
Wap-Cre;Cdh1F/+
Wap-Cre;Cdh1F/+;Fgfr2
Wap-Cre;Cdh1F/+;Fgfr2(1-716)
Wap-Cre;Cdh1F/+;Fgfr2(1-672)
n = 45
n = 37
n = 40
n = 27
Early lesions with noncohesive cells
Mammary intraepithelial neoplasia
Adenocarcinoma
Solid carcinoma
Spindle cell
Normal

Supplementary Figure 1. ◄ Overview of Cre-conditional alleles of the different *Fgfr2* variants, representative images and the distribution of the tumor morphologies observed in female mice expressing the different FGFR2 variants. (a) The schematic depiction of Cre-conditional *invCAG-Fgfr2(726)-IRES-Luc*, *invCAG-Fgfr2(1-716)-IRES-Luc* and *invCAG-Fgfr2(1-672)-IRES-Luc* alleles into in the *Col1a1* locus of *Wap-Cre;Cdh1$^{F/F}$* ESCs. Mammary-specific expression of Cre-recombinase results in the inversion of the CAG promoter and subsequent expression of the different FGFR2 variants accompanied by luciferase expression in mammary epithelial cells. **(b)** Quantification of bioluminescence imaging of luciferase expression over time in *Wap-Cre;Cdh1$^{F/F}$;Fgfr2* (*n* = 6), *Wap-Cre;Cdh1$^{F/F}$;Fgfr2(1-716)* (*n* = 4) and *Wap-Cre;Cdh1$^{F/F}$;Fgfr2(1-672)* females (*n* = 4). *Wap-Cre;Cdh1$^{F/F}$* (*n* = 3) female mice show background luminescence. **(c)** Representative images for hematoxylin and eosin (H&E) staining and immunohistochemical stainings of E-cadherin, cytokeratin-8 (CK8) and vimentin in classic ILCs, tumors with spindle-shaped cells and solid carcinomas. Scale bar, 50 µm. **(d,e)** The distribution of tumor morphologies as observed in the mammary glands of the indicated genotypes (normal epithelium and early lesions with noncohesive cells were excluded from the total mammary tumor burden). The ILC-like classification refers to lesions that are too small to show the typical invasive growth pattern of classic ILCs.



Supplementary Figure 2. Immunoprecipitation of GFP and the identification of FGFR2 interacting proteins in serum-starved NMuMG cells expressing GFP or the different FGFR2-GFP variants. **(a)** Immunoprecipitation of GFP in NMuMG cells expressing GFP-only, FGFR2-GFP, FGFR2(1-716)-GFP and FGFR2(1-672)-GFP, as visualized by immunoblot using an anti-GFP antibody. **(b)** The heat map showing the abundance of FGFR2 interacting proteins in serum-starved (for 48 hours) mammary epithelial cells expressing the different FGFR2-GFP variants compared to GFP-expressing control cells. Spectral counts of three independent biological experiments were averaged and represented with a color code. PSM, peptide-spectrum matches. **(c)** Coimmunoprecipitation of RSK1,2,3 in NIH3T3 fibroblasts expressing GFP-only or different FGFR2-GFP variants with or without the deletion of S696/697, as visualized by immunoblotting using an anti-RSK1,2,3 antibody. **(d)** Quantification of the colonies formed by NMuMG cells expressing GFP or the different FGFR2-GFP variants (with or without the deletion of S696 and S697, which are responsible for RSK2 binding), which were allowed to grow in soft agar for 14 days.



Supplementary Figure 3. The expression of phospho-S6 in mammary epithelial cells transduced with Lenti-*Fgfr2-gfp* variants can be used as a marker for the tumorigenic potential of different FGFR2 truncation variants in mice. **(a)** Representative flow cytometry dot plots of GFP expression in 48 hours serum-starved NMuMG cells transduced with Lenti-*Gfp* or Lenti-*Fgfr2(1-716)-gfp*, respectively. **(b)** Representative dot plots showing phosphorylated S6 expression in GFP^{high} mammary epithelial cells expressing GFP or FGFR2(1-716)-GFP, respectively. GFP^{high} cells expressing GFP show basal expression of phospho-S6 after 48 hours of serum-starvation. **(c)** FGFR2 expression in NMuMG cells transduced with different Lenti-*Fgfr2-gfp* variants after 6 days. β-actin is used as loading control. **(d)** Growth kinetics of tumors induced by intraductal injections of Lenti-*Fgfr2* (n = 10), Lenti-*Fgfr2(1-716)* (n = 9) and Lenti-*Fgfr2(1-672)* (n = 13) in *Wap-Cre;Cdh1$^{F/F}$* females. **(e)** Growth kinetics of tumors induced by intraductal injections of Lenti-*Fgfr2* (n = 13), Lenti-*Fgfr2(1-716)* (n = 13) and Lenti-*Fgfr2(1-672)* (n = 23) in FVB wild-type females. **(f,g)** Kaplan Meier curves showing the tumor-free survival of mammary glands injected with Lenti-*Fgfr2*, Lenti-*Fgfr2(1-716)* and Lenti-*Fgfr2(1-672)* in *Wap-Cre;Cdh1$^{F/F}$* **(f)** and FVB wild-type **(g)** female mice. The curves were not significantly different as calculated using a Mantel–Cox test. **(h)** The survival until tumors exceeded the size of 200 mm^3 in the glands of FVB wild-type female mice injected with Lenti-*Fgfr2(1-677)* (n = 15), Lenti-*Fgfr2(685)* (n = 16) and Lenti-*Fgfr2(1-693)* (n = 12), as shown using a Kaplan Meier curve. *P < 0.05 by Mantel–Cox test.

Supplementary Table 1: Overview of the primers used for generating the lentiviral constructs.

Information	Template	Experiment		Sequence
Fgfr2 in pcDNA3.1/Zeo (+) (without stopcodon)	*Fgfr2* (NM_201601) cDNA	Add NotI-FseI overhangs	FW	AAAAGCGGCCGCATGGGATTAC
Fgfr2 in pcDNA3.1/Zeo (+) (without stopcodon)	*Fgfr2* (NM_201601) cDNA	Add NotI-FseI overhangs	RV	AAAAGGCCGGCCTGTTTTAACACTGCC
Fgfr2(1-716) in pcDNA3.1/Zeo (+) (without stopcodon)	*Fgfr2* (NM_201601) cDNA	Add NotI-FseI overhangs	FW	AAAAGCGGCCGCATGGGATTAC
Fgfr2(1-716) in pcDNA3.1/Zeo (+) (without stopcodon)	*Fgfr2* (NM_201601) cDNA	Add NotI-FseI overhangs	RV	AAAAGGCCGGCCCTGAGGCAGACAGGG
Fgfr2(1-672) in pcDNA3.1/Zeo (+) (without stopcodon)	*Fgfr2* (NM_201601) cDNA	Add NotI-FseI overhangs	FW	AAAAGCGGCCGCATGGGATTAC
Fgfr2(1-672) in pcDNA3.1/Zeo (+) (without stopcodon)	*Fgfr2* (NM_201601) cDNA	Add NotI-FseI overhangs	RV	AAAAGGCCGGCCCTCATTGGTTGTGAG
Gfp (with ATG) in pcDNA3.1/Zeo (+)	SIN.LV.SF vector	Add NotI-XhoI overhangs	FW	GCGGCCGCATGGTGAG
Gfp (with ATG) in pcDNA3.1/Zeo (+)	SIN.LV.SF vector	Add NotI-XhoI overhangs	RV	ACTCGAGTTACTTGTACAGCTCGTCCATG
Gfp (in *Fgfr2-gfp* - without ATG) in pcDNA3.1/Zeo (+)	SIN.LV.SF vector	Add FseI-XhoI overhangs	FW	AGGCCGGCCGTGAGCAAG
Gfp (in *Fgfr2-gfp* - without ATG) in pcDNA3.1/Zeo (+)	SIN.LV.SF vector	Add FseI-XhoI overhangs	RV	ACTCGAGTTACTTGTACAGCTCGTCCATG
Gfp in SIN.LV.SF	pcDNA™3.1-*Gfp*	Add AgeI-SalI overhangs	FW	AAAACCGGTGCGGCCGCATGGTGAG
Gfp in SIN.LV.SF	pcDNA™3.1-*Gfp*	Add AgeI-SalI overhangs	RV	AAAGTCGACCTCGAGTTACTTGTACAGCTCGTCCATG
The different *Fgfr2-gfp* variants in SIN.LV.SF	corresponding pcDNA™3.1-*Fgfr2-gfp*	Add AgeI-SalI overhangs	FW	AAAACCGGTGCGGCCGCATGGGATTAC
The different *Fgfr2-gfp* variants in SIN.LV.SF	corresponding pcDNA™3.1-*Fgfr2-gfp*	Add AgeI-SalI overhangs	RV	AAAGTCGACCTCGAGTTACTTGTACAGCTCGTCCATG
Fgfr2 in SIN.LV.SF	*Fgfr2* (NM_201601) cDNA	Add AgeI-SalI overhangs	FW	AAAACCGGTGCGGCCGCATGGGATTAC
Fgfr2 in SIN.LV.SF	*Fgfr2* (NM_201601) cDNA	Add AgeI-SalI overhangs	RV	AAAAGTCGACCTCGAGTCATGTTTTAACACTGCC
Fgfr2(1-716) in SIN.LV.SF	*Fgfr2* (NM_201601) cDNA	Add AgeI-SalI overhangs	FW	AAAACCGGTGCGGCCGCATGGGATTAC
Fgfr2(1-716) in SIN.LV.SF	*Fgfr2* (NM_201601) cDNA	Add AgeI-SalI overhangs	RV	AAAAGTCGACCTCGAGTCACTGAGGCAGACAGGG
Fgfr2(1-672) in SIN.LV.SF	*Fgfr2* (NM_201601) cDNA	Add AgeI-SalI overhangs	FW	AAAACCGGTGCGGCCGCATGGGATTAC
Fgfr2(1-672) in SIN.LV.SF	*Fgfr2* (NM_201601) cDNA	Add AgeI-SalI overhangs	RV	AAAAGTCGACCTCGAGTCACTCATTGGTTGTGAG