Effects of *PIK3CA* mutations

on mammary cell fate and cancer

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Shany Koren

aus Deutschland

Basel, 2016

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Prof. Dr. Nancy Hynes

Dr. Mohamed Bentires-Alj

Dr. Matthew J. Smalley

Basel, den 8.12.2015

Prof. Dr. Jörg Schibler

Dekan

1 | Summary

The adult mouse mammary epithelium contains self-sustained cell lineages that form the inner luminal and outer basal cell layers, with stem and progenitor cells contributing to its proliferative and regenerative potential. A key issue in breast cancer biology is the effect of genomic lesions in specific mammary cell lineages on tumor heterogeneity and progression. The impact of transforming events on fate conversion in cancer cells-of-origins and thus their contribution to tumor heterogeneity remains largely elusive.

The phosphoinositide 3-kinase (PI3K) signaling pathway is crucial for cell growth, proliferation, metabolism and survival and is frequently deregulated in human cancer, including ~70% of breast tumors. The gene encoding for the alpha catalytic subunit of PI3K (*PIK3CA*) is mutated and/or amplified in over 30% of breast cancers. Mutations in either the kinase domain (H1047R) or the helical domain (E545K) are most common and result in a constitutively active enzyme with oncogenic capacity. $PIK3CA^{\text{H1047R}}$ was previously shown to induce heterogeneous mammary tumors in transgenic mouse models. Whether overexpression of the $PIK3CA^{\text{E545K}}$ mutant is sufficient to induce mammary tumors in transgenic mice has not been defined. Moreover, the origin of $PIK3CA^{\text{H1047R}}$-evoked tumor heterogeneity and the influence of the cell-of-origin on aggressiveness of breast cancer have remained elusive.

In my PhD studies, we demonstrate that expression of $PIK3CA^{\text{E545K}}$ in the mouse mammary gland induces heterogeneous mammary carcinomas but with a longer latency than $PIK3CA^{\text{H1047R}}$ suggesting that the helical domain mutant $PIK3CA^{\text{E545K}}$ is a less potent inducer of mammary tumors. Furthermore, by using *in situ* genetic lineage tracing, gene expression analyses and limiting dilution transplantation, we have unraveled the potential of $PIK3CA^{\text{H1047R}}$ to induce multipotency during tumorigenesis in the mammary gland. We show that expression of $PIK3CA^{\text{H1047R}}$ in lineage-committed basal Lgr5-positive and luminal keratin 8-positive cells of the adult mouse mammary gland evokes cell dedifferentiation into a

1

multipotent stem-like state, suggesting this to be a mechanism involved in the formation of heterogeneous, multi-lineage mammary tumors. Moreover, we show that the tumor cell-of-origin influences the frequency of malignant mammary tumors. Our results define a key effect of $PIK3CA^{H1047R}$ on mammary cell fate in the pre-neoplastic mammary gland and show that the cell-of-origin of $PIK3CA^{H1047R}$ tumors dictates their malignancy, thus revealing a mechanism underlying tumor heterogeneity and aggressiveness.

2 | Contents

Contents

3 | Introduction

3.1 The mouse mammary gland

3.1.1 The mouse mammary gland development and physiology

The mammary gland undergoes dynamic changes during development and throughout life. Derived from ectodermal cells, it forms a rudimentary ductal tree, the mammary primordium, before birth. Postnatal development is regulated by hormones, with puberty inducing the outgrowth of the ductal tree into the mammary fat pad under the influence of growth hormone, insulin-like growth factor 1 and estrogen. During pregnancy progesterone and prolactin stimulate further remodeling of the gland, generating the alveoli that secrete milk during lactation. Following lactation, the mammary gland ductal tree regresses to a pre-pregnancy-like stage, a process called involution (Macias and Hinck, 2012; Petersen and Polyak, 2010; Sternlicht, 2006).

The branched ductal-alveolar tree making up the postnatal mouse mammary gland is surrounded by a basement membrane and stromal cells and is composed of hierarchically organized cell types that contribute to tissue homeostasis. Two major cell lineages organized in a bi-layered tubular structure constitute the mammary epithelium **(Fig. 3-1)**. The luminal layer lining the ducts and the alveoli is composed of cells expressing keratin 8/18 (K8/18) and/or estrogen and/or progesterone receptor (ER/PR). The contractile myoepithelial layer with a basal location is composed of cells expressing keratin 5/14 (K5/14) and/or smooth muscle actin (SMA) and/or p63 (Bissell et al., 2003; Hennighausen and Robinson, 2001).

Figure 3-1 │Mouse mammary gland physiology
Scheme of an adult mouse mammary gland and duct. The branched mammary gland epithelium of adult nulliparous mice is surrounded by the mammary fat pad which comprises stromal cells, vascular cells and adipocytes (*left*). The mammary duct is organized in a bi-layered structure surrounded by a basement membrane (magnification, *right*). The luminal layer lining the ducts is composed of cells expressing keratin 8/18 and/or estrogen receptor/progesterone receptor. The basal layer is composed of cells expressing keratin 5/14 and/or smooth muscle actin and/or p63.

3.1.2   The mouse mammary gland hierarchy

The mouse mammary gland epithelium is hierarchically organized (Visvader and Stingl, 2014) **(Fig. 3-2)**. Multipotent cells that generate both the luminal and basal lineages are found in the mouse embryonic mammary gland (Spike et al., 2012; Van Keymeulen et al., 2011) but their existence in the adult gland is still under debate. Studies using serial transplantation into cleared mammary fat pad of mammary fragments, retroviral-infected mammary cells (Deome et al., 1959; Faulkin and Deome, 1960; Kordon and Smith, 1998), or cells isolated by fluorescence-activated cell sorting (FACS) using cell surface markers (e.g. CD24, Sca1, EpCAM, CD49f, CD29, CD61) (Prater et al., 2014; Shackleton et al., 2006; Sleeman et al., 2006; Sleeman et al., 2007; Stingl et al., 2006) **(Fig. 3-3)** supposed the existence in the adult mouse mammary gland of multipotent stem cells with myoepithelial features. Arguably, these assays reflected the regenerative potential of the transplanted cells rather than their properties

Figure 3-2 │Model of mammary cell hierarchy
During embryonic development, multipotent mammary stem cells (MaSCs) contribute to the luminal and basal lineages. Currently, the existence of adult multipotent MaSCs is under debate. Self-sustained unipotent luminal and basal stem cells as well as multipotent stem cells may maintain the homeostasis of the adult virgin mammary gland. The organization of their hierarchy is still unknown.
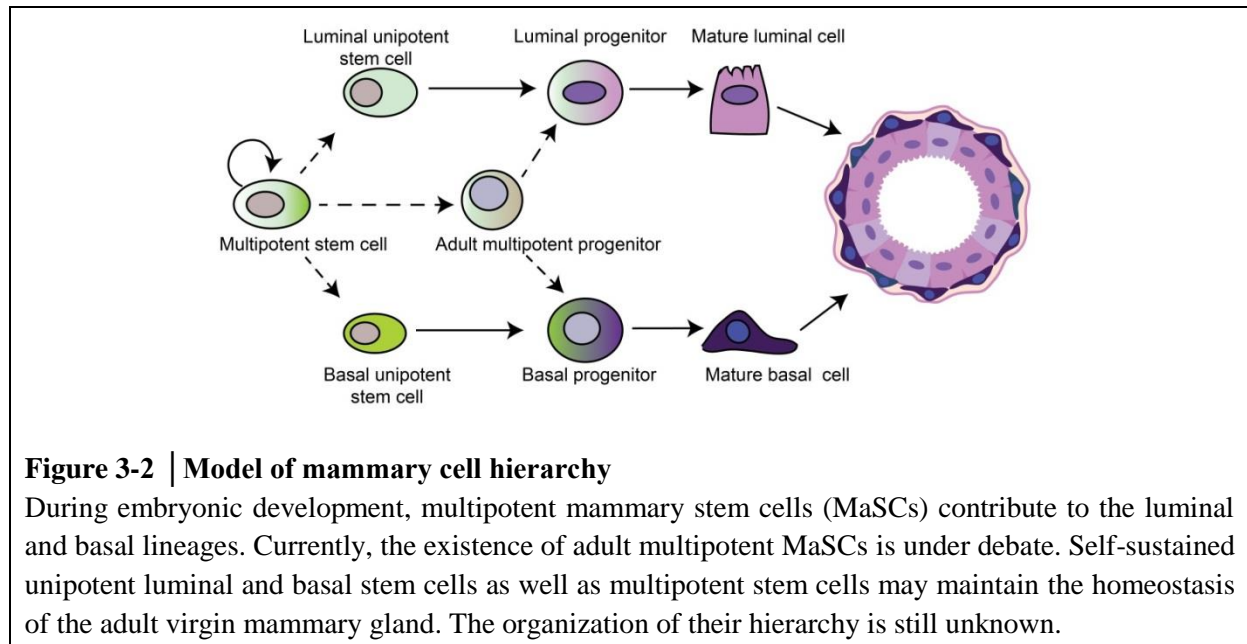
in situ (de Visser et al., 2012; van Amerongen et al., 2012; Van Keymeulen et al., 2011). Lineage-tracing studies, which permit targeted expression of a fluorescent reporter in a given cell and its progeny, showed that tissue homeostasis is maintained by unipotent luminal K8/18-positive and basal K5/14/Lgr5-positive stem cells after birth. Lineage-tracing of K8/18-, K5/14- and Lgr5-progeny found no evidence for the presence of multipotent stem cells in the adult mammary gland (Van Keymeulen et al., 2011) but did not exclude the possibility that rare cells not targeted by these reporters, or only at a very low frequency, have multipotent potential. Whilst tracing of the progeny of axin2-positive cells showed the presence of multipotent stem cells during puberty and pregnancy, this and other studies, revealed that the basal and luminal lineages are self-sustained in the adult virgin gland (de Visser et al., 2012; Tao et al., 2014; van Amerongen et al., 2012). In contrast, recent three-dimensional whole-mount imaging (Rios et al., 2014) and the identification of the Procr-positive subset (Wang et al., 2015) argue for the presence of multipotent stem cells in the adult virgin mouse mammary gland, thus reopening the debate. Delineating normal mammary cell hierarchy in mouse and humans is fundamental to the understanding of breast tumor heterogeneity.

Figure 3-3│ Separation of mammary epithelial cell subpopulations by FACS
a, Representative FACS-plot for separation of basal and luminal mammary subpopulations using anti-CD24 and anti-Sca1 cell surface markers. **b,** Representative immunofluorescence staining of isolated mammary epithelial cells for basal (K14, red) and luminal (K8/18, green) markers.
The red box indicates CD24^{low}Sca1^{negative} basal population. The orange and yellow boxes indicate CD24^{high}Sca1^{negative} and CD24^{high}Sca1^{positive} luminal populations, respectively.

3.2 Breast cancer

3.2.1 Breast cancer prevalence and classification

Breast cancer is the most frequently diagnosed cancer and the leading cause of death in female cancer patients with yearly over 1.5 million cases and 500,000 deaths worldwide (Torre et al., 2015). Breast cancer progresses from a premalignant disease (e.g., hyperplasia, ductal carcinoma *in situ*) to invasive carcinoma and metastasis. Differing histopathological parameters such as receptor status (estrogen-receptor, progesterone-receptor and/or *ERBB2/HER2* overexpression) and proliferation status (Ki67-expression) (Viale, 2012), and molecular profiling subdivides breast cancer into at least six subtypes (normal-like, luminal A, luminal B, HER2-enriched, claudin-low, and basal-like). These expression profiles mostly reflect different clinical prognoses (Perou et al., 2000; Prat et al., 2010; Sorlie et al., 2001; Sorlie et al., 2003) and to some extent responses to therapy (Troester et al., 2004). Integrated genomic and transcriptomic analysis of breast tumors has revealed further subgroups with distinct clinical outcomes (Curtis et al., 2012).

3.2.2   Heterogeneity in breast cancer

Breast cancer displays molecular, phenotypic, and functional diversity within a patient's tumor (intratumor heterogeneity) and among tumors from different patients (intertumor heterogeneity). This tumor heterogeneity challenges accurate prognosis and a single biopsy can misjudge the complexity of the disease. Pathologists biopsy multiple regions of a tumor and reach a diagnosis based on the most malignant region, yet further aggressive areas may be missed due to their scarcity and/or topological heterogeneity (Komaki et al., 2006). Generally, only a few histopathological parameters are assessed and this may overlook meaningful information. Furthermore, gene expression profiling of breast cancer is usually performed on the total biopsy, which may dilute information on aggressiveness. Breast cancer mortality essentially results from metastases in bone, lung, brain and liver, but a systematic and comprehensive assessment of the molecular makeup of metastases is still not available. Indeed, metastases may display varying genetic and non-genetic alterations, also when compared to the bulk of the primary tumor (Ding et al., 2010; Shah et al., 2009). This is a possible cause of a significant number of therapy failures. Therefore, imprecise sampling, particularly of metastases, and lack of in-depth molecular analysis are serious problems that need to be tackled if research discoveries are to be translated into long-lasting therapies.

The origins of intra- and intertumor heterogeneity are as yet not fully understood. Cell-autonomous (e.g., genetic and epigenetic) and non-cell-autonomous (e.g., tumor microenvironment) factors, as well as stochastic events (reviewed in (Marusyk et al., 2012)) are possible sources of cancer cell diversity and their delineation is important for a better understanding of tumor progression and could ultimately help design better therapies.

3.2.3 Tumor heterogeneity in preclinical cancer models

The use of preclinical models havs led to significant progress in understanding breast cancer. Mouse models of mammary tumors induced by the mouse mammary tumor virus (MMTV) provided a means to study tumor heterogeneity (Cohen et al., 1979). Numerous mouse models have been genetically engineered to investigate tumor initiation and progression and to test anti-cancer drugs. These were generated by genomic deletion of tumor suppressor genes and/or by expression of oncogenes under a mammary-specific promoter. Early studies expressing *ErbB2/Neu, Myc, H-ras* or *PyMT* (Guy et al., 1992; Muller et al., 1988; Sinn et al., 1987; Stewart et al., 1984), the generation of conditional mouse models using tamoxifen-inducible Cre/loxP (Jonkers et al., 2001) or doxycycline-inducible tet-on (Gunther et al., 2002), as well as advances in tissue- and cell type-specificity have demonstrated the importance of mouse models for studying inter- and intratumor heterogeneity of breast cancers (Cleary et al., 2014; Koren and Bentires-Alj, 2013; Liu et al., 2007; Melchor et al., 2014; Meyer et al., 2011; Meyer et al., 2013; Molyneux et al., 2010; Tao et al., 2015).

But the question remains how accurately murine models reflect human breast cancer heterogeneity. Histological phenotypes of several mouse mammary tumors do not resemble human breast cancers and the frequency of hormone-dependent mammary cancers is much lower in mouse models than in humans (Cardiff, 2001; Cardiff et al., 2000). Nonetheless these models recapitulate some aspects of molecular human breast tumor heterogeneity. Some mouse models resemble human breast cancer molecular subtypes and show conserved tumorigenic pathways (Herschkowitz et al., 2007; Hollern and Andrechek, 2014; Pfefferle et al., 2013). As expected, none of the current genetically engineered mouse models can recapitulate all characteristics of human breast cancer, yet such model systems are valuable for interrogating specific aspects of human disease and for testing hypotheses related to *in situ* tumor progression at the organismic level and in the presence of an intact immune system.

Human breast cancer cell lines are the most widely used preclinical model. Established human cell lines are derived from human breast cancer samples and pleural effusions and are propagated and manipulated *in vitro* or as xenograft transplants *in vivo*. They display many but not all of the recurrent genomic alterations found in human samples (Neve et al., 2006). Nevertheless, cell lines have proved useful as preclinical models in oncology (Neve et al., 2006; Voskoglou-Nomikos et al., 2003).
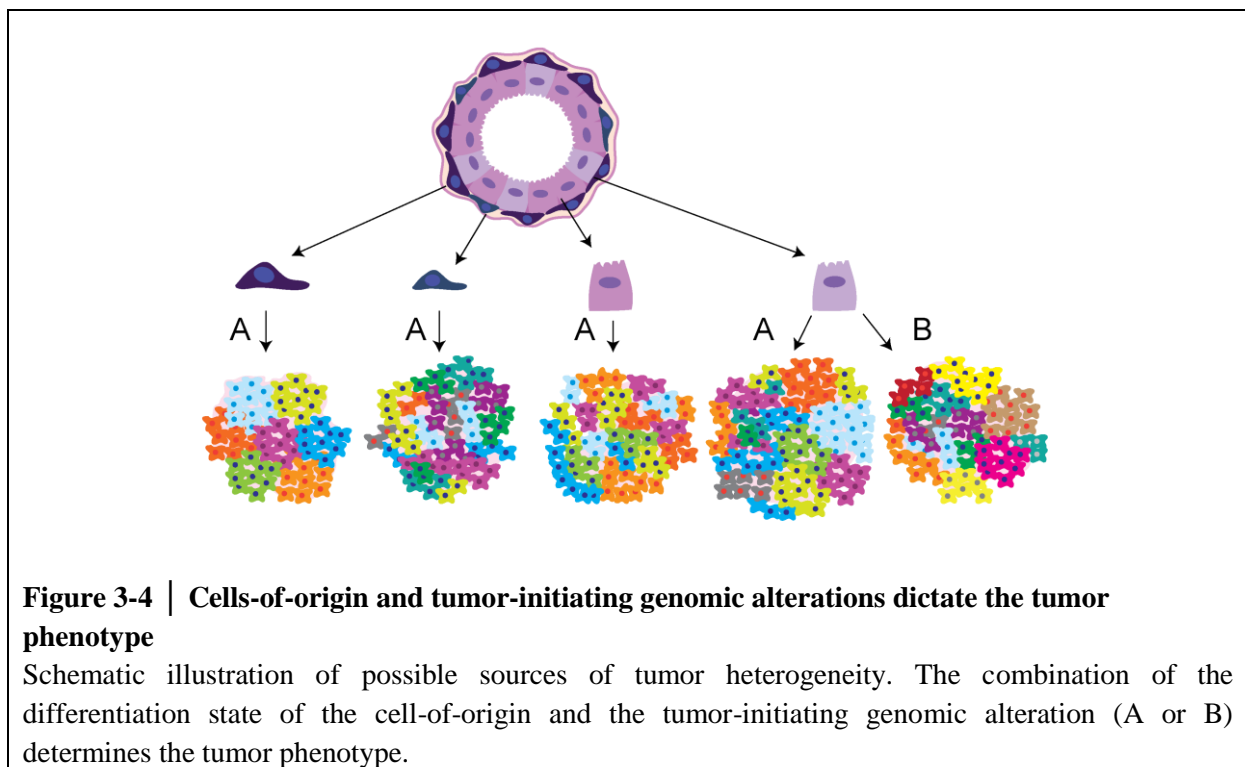
Primary human tumors can be transferred directly from patients into immunodeficient mice. These "patient-derived xenografts (PDXs)" generally show inter- and intra-tumor heterogeneity (Cassidy et al., 2015; DeRose et al., 2011; Zhang et al., 2013). PDXs display clonal dynamics (Eirew et al., 2015) and can be used to investigate patient-specific response to therapy (Hidalgo et al., 2014; Whittle et al., 2015). Breast PDXs mostly maintain the genomic and histopathological profiles and estrogen receptor-dependencies of the corresponding patient tumors during multiple passages *in vivo* (DeRose et al., 2011; Reyal et al., 2012). Yet because of cross-species incompatibilities and their transplantation into immunodeficient mice, human cell line xenografts and PDX models both lack the contribution of some non-cell autonomous drivers and the immune system to tumor heterogeneity. Even so, such model systems are promising tools to further investigate the heterogeneity of breast cancer and to develop precision therapy.

3.2.4 Sources of breast tumor heterogeneity

Different concepts relating to the origin of tumor diversity have been proposed, including the differentiation state of the initially transformed cell (cell-of-origin of cancer), cancer cell plasticity, genetic evolution and tumor microenvironment. Genetic and non-genetic alterations underlie these not mutually exclusive sources of heterogeneity.

3.2.4.1 Differentiation state of the cell-of-origin

Theoretically, each cell type within the mammary gland hierarchy can be the subject of genomic alterations **(Fig. 3-4)** but whether all cell subtypes are prone to cancer is unclear. Moreover, the contribution of the cell-of-origin to tumorigenesis has proved largely elusive. Human mammary epithelial cells (HMECs) with distinct differentiation states caused by different culture media were transformed with the same set of oncogenes and tested in xenograft transplantations (Ince et al., 2007). The resulting tumors exhibited differences in histopathology, tumorigenicity, and metastatic behavior. Moreover, analysis of FACS-sorted transformed HMEC subsets suggested that the differentiation state of the cell-of-origin is a determinant of the tumor phenotype (Chaffer et al., 2011). Similarly, transformation of human epithelial cell adhesion molecule (EpCAM$^+$)-positive cells led to the formation of ER-positive and -negative tumors, while transformation of CD10$^+$ cells resulted in metaplastic tumors reminiscent of claudin-low breast cancer (Keller et al., 2012). While it is difficult to analyze the tumor cell-of-origin in human breast cancer retrospectively, mouse models have proven

Figure 3-4 │ Cells-of-origin and tumor-initiating genomic alterations dictate the tumor phenotype
Schematic illustration of possible sources of tumor heterogeneity. The combination of the differentiation state of the cell-of-origin and the tumor-initiating genomic alteration (A or B) determines the tumor phenotype.

informative. By targeting different cell types, conditional inactivation of *p53* in cells expressing the Cre recombinase under the control of the mouse mammary tumor virus long terminal repeat (MMTV-Cre) evoked ER-negative tumors with a longer latency than its inactivation in Whey acidic protein (WAP-Cre) expressing models that developed ER-positive and ER-negative tumors (Lin et al., 2004). *Rb* deletion in MMTV-Cre but not in WAP-Cre mice resulted in mammary tumors (Jiang et al., 2010), emphasizing the importance of the cell-of-origin in tumor susceptibility.

It was long believed that basal-like breast cancers that express high levels of basal cell markers originate from transformed basal progenitor/stem cells and that luminal-type breast cancers with high levels of luminal cell markers originate from luminal progenitors. However, it has been shown that basal-like mammary cancer can arise from luminal progenitors (Molyneux et al., 2010). Similarly, analysis of pre-neoplastic human tissue from *BRCA1* mutation carriers revealed an expanded population of aberrant luminal progenitor cells whose expression profile associated with basal cancers, suggesting the luminal progenitor population as a target population in *BRCA1*-associated basal-like breast tumors (Lim et al., 2009). The use of the term "basal-like" is debatable as discussed in a review by B. Gusterson (Gusterson, 2009).

Just how different cell types yield different tumor types is yet little understood at the molecular level. Each cell type in the mammary gland hierarchy has its own signaling, transcriptional and epigenetic profile that determines cell identity (Kendrick et al., 2008; Lim et al., 2010; Pal et al., 2013). Further in-depth analysis is warranted to understand better the critical differences between cells-of-origin and heterogeneous tumors.

It is probable that mammary tumor phenotypes depend on interactions between cells-of-origin and initiating genetic alterations (Melchor et al., 2014). Multiple phenotypes can arise depending on the initiating depletion of a tumor suppressor gene (*Brca1/2*, *p53* and/or *Pten*) in basal (K14-Cre model) or luminal ER-negative cells (beta-lactoglobulin (Blg)-Cre
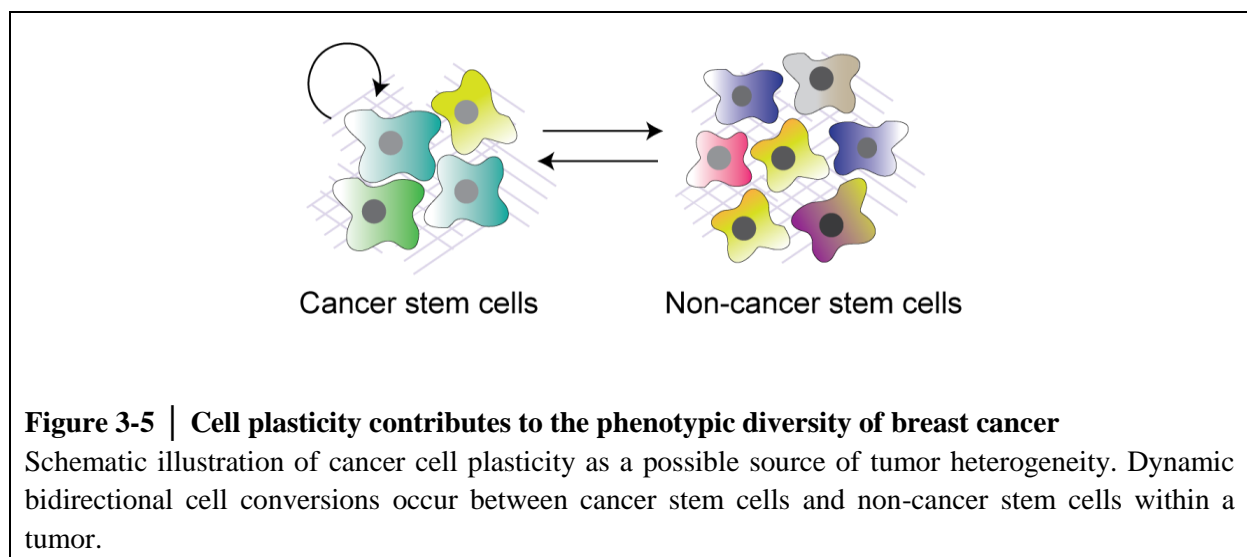
model). Luminal ER-negative cells can give rise to various phenotypes, including ER-negative and -positive mammary tumors. In contrast, tumors arising from basal cells show similar phenotypes irrespective of the depleted gene. For example, while *Brca2/p53* depletion in Blg-Cre cells evokes invasive ductal carcinomas of no special type and metaplastic spindle cell carcinomas, its depletion in K14-Cre cells evokes malignant adenomyoepitheliomas. The initiating mutation seems to be the prime determinant of the molecular profile of tumors of breast cancer, as *Brca1/2* and *Pten* depletion in any of the tested cells-of-origin generates basal-like or normal-like cancers, respectively (Melchor et al., 2014). A further recent study suggested that one cell-of-origin can give rise to different breast cancer subtypes depending on the oncogene expressed. While expression of *Neu* or *PyMT* gives rise to tumors with purely luminal differentiation originating from luminal WAP-Cre cells, expression of *Etv6-NTRK3* oncoprotein under the same promoter leads to tumors with basal differentiation (Tao et al., 2015), indicating that distinct oncogenes might have distinct effects on the tumor phenotype. Arguably, WAP can drive expression in ER-positive and ER-negative cells and conceivably *Etv6-NTRK3* preferentially transforms the ER-negative population while *Neu* or *PyMT* preferentially transform ER-positive cells.

Premalignant ductal or lobular carcinoma *in situ* already display heterogeneity (Clark et al., 2011) conceivably originating from several cell-of-origins. Important selection pressure during the transition from *in situ* to invasive carcinoma and to metastasis may further contribute to intra-tumor heterogeneity, highlighting the complexity of this disease.

These studies show that the combination of initiating transforming events and tumor cells-of-origin has a bearing on breast cancer diversity and that phenotypes do not reflect the cell-of-origin of a cancer. The recognition of these two contributors to tumor heterogeneity may have great implications for the diagnosis, prognosis and treatment of cancer.

3.2.4.2 Cell plasticity and tumor cell hierarchy

Sorting of cancer cell subpopulations has revealed that some tumor cells are tumorigenic but others are not. Notably, tumorigenic subpopulations can give rise to tumorigenic and non-tumorigenic progenies. The term "cancer stem cells" (CSCs) was coined to describe tumorigenic cells that can self-renew (i.e., form tumors when serially passaged at limiting dilutions) and give rise to tumors that display the phenotypic heterogeneity of the parental tumor. This concept implies a hierarchical organization of tumors reminiscent of normal tissue in which stem cells are at the apex of the hierarchy, giving rise both to further stem cells and to differentiated cells (Kreso and Dick, 2014). Even though the organization of tumor cell hierarchy is not yet clear, recent studies have proposed dynamic bidirectional cell conversion. Using transformed HMECs *in vitro*, Chaffer *et al.* showed not only that CSCs give rise to differentiated cells but that differentiated tumor cells are also able to dedifferentiate (Chaffer et al., 2011). Cancer cell interconversions appear to maintain the equilibrium of cell states within a tumor (Gupta et al., 2011) **(Fig. 3-5)**. Exome sequencing of CSCs from 12 breast cancer patients using paired primary tumor samples showed that the majority of mutations are shared between CSCs and the bulk primary tumor, which suggests a

Cancer stem cells   Non-cancer stem cells

Figure 3-5 │ Cell plasticity contributes to the phenotypic diversity of breast cancer
Schematic illustration of cancer cell plasticity as a possible source of tumor heterogeneity. Dynamic bidirectional cell conversions occur between cancer stem cells and non-cancer stem cells within a tumor.
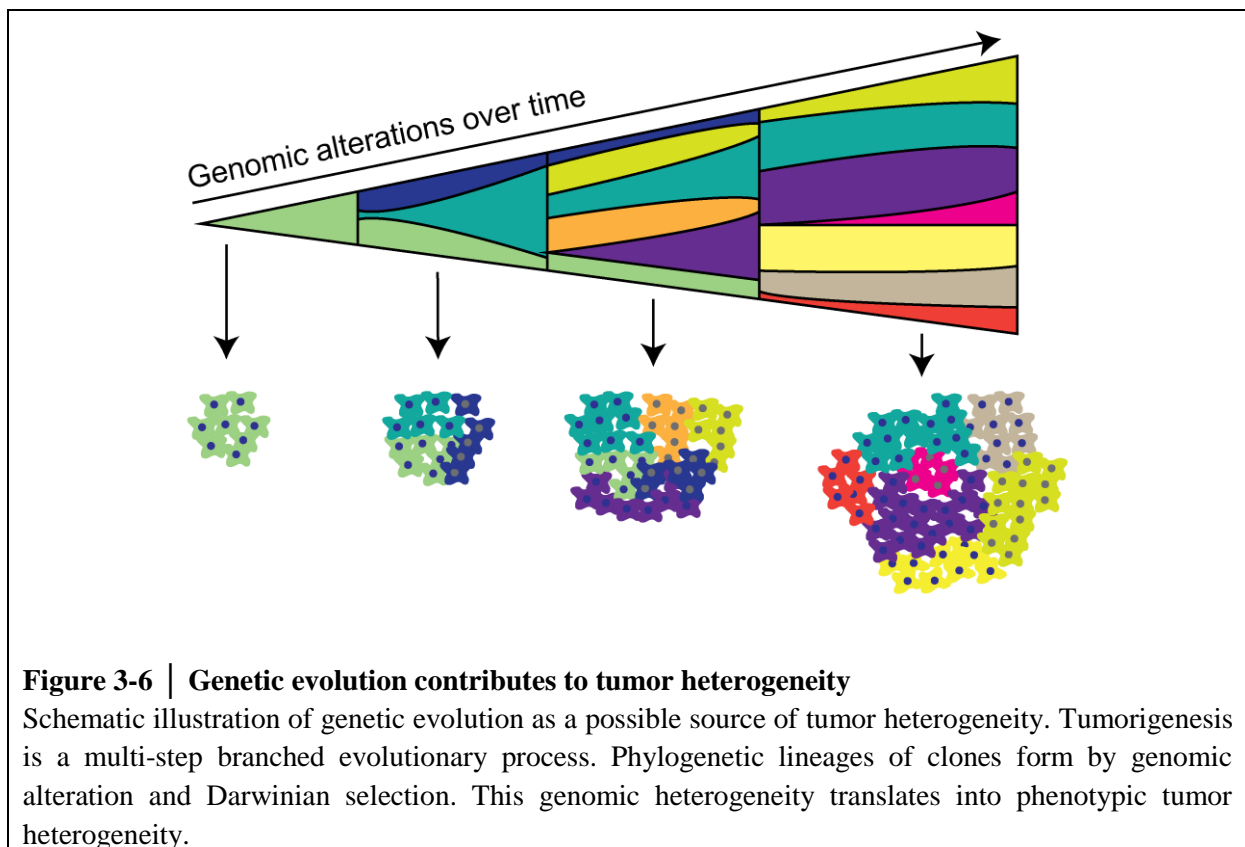
dynamic switch between CSCs and differentiated cell states (Klevebring et al., 2014). A further *in vitro* study reported culture medium-dependent phenotypic plasticity in DKAT cells derived from an aggressive, treatment-resistant triple-negative heterogeneous human breast cancer. Whereas DKAT cells cultured in serum-containing medium displayed epithelial morphology, growth in serum-free media evoked mesenchymal characteristics (D'Amato et al., 2012). *In vivo* studies in prostate (Choi et al., 2012; Wang et al., 2013) and intestine (Schwitalla et al., 2013) have demonstrated deregulation of epithelial differentiation during tumorigenesis.

The mechanism underlying this oncogene-triggered plasticity is not yet understood. The dedifferentiation processes might involve epithelial-to-mesenchymal transitions (EMT), given that overexpression of several EMT-markers (including *SNAI1*, *SNAI2*, *TWIST1*) has been found in basal-like cancers (Skibinski and Kuperwasser, 2015). For example, Slug and Sox9 were shown to induce interconversions between luminal progenitor and stem cell states and these factors are required for the maintenance of tumor-initiating capacity in human MDA-MB231 cells (Guo et al., 2012). Additionally, ablation of *Slug* in MMTV-*Myc* mice resulted in resistance to tumorigenesis, suggesting that plasticity is crucial to tumor development. Slug also interacts with chromatin-modifier lysine-specific demethylase 1 (LSD1) and, thus, epigenetic remodeling may be involved in these processes (Phillips et al., 2014). A further study showed that cell plasticity is triggered by epigenetic remodeling driven by the EMT-transcription factor ZEB1 (Chaffer et al., 2013). Hence, plastic tumor cells may maintain bivalent chromatin configurations and activate/repress transcriptional programs in response to cell-cell and microenvironmental cues, conceivably resulting in different cell fates and degrees of tumor aggressiveness. At present, it is still not known whether all tumor cells can undergo bidirectional cell conversion, and the exact molecular mechanisms underlying cancer cell plasticity need further investigation.

16

3.2.4.3 Genetic evolution

Substantial genetic diversity is found in human breast cancers (Banerji et al., 2012; Cancer Genome Atlas, 2012; Curtis et al., 2012; Ellis et al., 2012; Shah et al., 2012; Stephens et al., 2012). The concept of genetic evolution as a source of tumor heterogeneity has existed for several decades. Tumorigenesis is an evolutionary process driven by mutations and Darwinian selection (Nowell, 1976). Heritable beneficial mutations in a tumor cell may pass to the progeny, which gain survival and proliferation advantages. Such tumor cells with increased robustness may form clones that become dominant through the occurrence of further favorable mutations in a multi-step manner. Hence, clones within a tumor form independent phylogenetic lineages and this genetic heterogeneity translates into phenotypic tumor diversity (Kreso and Dick, 2014; Marusyk and Polyak, 2010; Nowell, 1976) **(Fig. 3-6)**. Molecular profiling of 21 human breast cancers by whole-genome sequencing was performed to gain insight into their genomic architecture. By bioinformatic analysis, Nik-Zainal *et al.*

Figure 3-6 │ Genetic evolution contributes to tumor heterogeneity
Schematic illustration of genetic evolution as a possible source of tumor heterogeneity. Tumorigenesis is a multi-step branched evolutionary process. Phylogenetic lineages of clones form by genomic alteration and Darwinian selection. This genomic heterogeneity translates into phenotypic tumor heterogeneity.

reconstructed a phylogenetic tree for one tumor, highlighting the evolution of breast cancer through the occurrence of driver mutations and clonal expansion (Nik-Zainal et al., 2012a; Nik-Zainal et al., 2012b). Single-cell sequencing of different sectors of a breast tumor revealed distinct clonal subpopulations suggesting sequential clonal expansions with few persistent intermediates. Moreover, analysis of single cells from a monogenomic primary tumor and its liver metastases indicated that a single clonal expansion formed the primary tumor and seeded the metastases (Navin et al., 2011). Similarly, multi-region genetic analysis (exome sequencing, chromosome aberration analysis, and ploidy profiling) of renal carcinoma and matched metastases has revealed branched evolutionary tumor progression (Gerlinger et al., 2012). The results of a further recent study investigating the genetic landscape of multifocal lesions in breast cancer concur with the notion of genetic evolution. Oncogenic variants were found to be shared frequently between lesions, especially proximal rather than distal lesions, indicating the common origin and evolutionary process of tumor progression (Desmedt et al., 2015). Sequencing of multifocal cancers from another study revealed that distinct foci were clonally related suggesting that tumor subclones are capable of transiting distances through normal tissue and during progression (Yates et al., 2015).

The concept of genetic evolution by itself is insufficient to explain tumor heterogeneity as it only considers homogeneous genetic pools within clones and not the functional diversity of cell states (e.g., CSCs) of clone constituents. The genetic evolution and CSC models are not necessarily mutually exclusive and a unifying model has been proposed in which CSCs may evolve and change in frequency under the influence of clonal genetic evolution during tumor progression (Kreso and Dick, 2014).

3.2.4.4   Tumor stroma and clonal cooperation

In healthy tissue, cell identity and homeostasis are tightly controlled not only by cell autonomous mechanisms but also by reciprocal interactions with the environment. Failure to preserve tissue stability may result in neoplastic transformation (Goubran et al., 2014). Cell plasticity and tumor heterogeneity are also influenced by extrinsic factors. Interactions between tumor cells and their microenvironment, which includes stromal cells, blood vessels and the immune system via paracrine factors, contribute to tumor progression (Allinen et al., 2004; Bissell and Hines, 2011) **(Fig. 3-7)**. In cancer, physiological interactions between cells and the microenvironment are disrupted as tumor cells are not sensitive to growth restrictive cues from their surroundings (Joyce and Pollard, 2009; Quail and Joyce, 2013). Cancer-associated fibroblasts (CAFs) differ from normal fibroblasts and are among the most abundant cell types in the tumor stroma. While normal fibroblasts from reduction mammoplasties retain the epithelial morphology of HMECs, co-culture with CAFs evokes a mesenchymal morphology (Dumont et al., 2013), highlighting the influence of the tumor stroma on cell
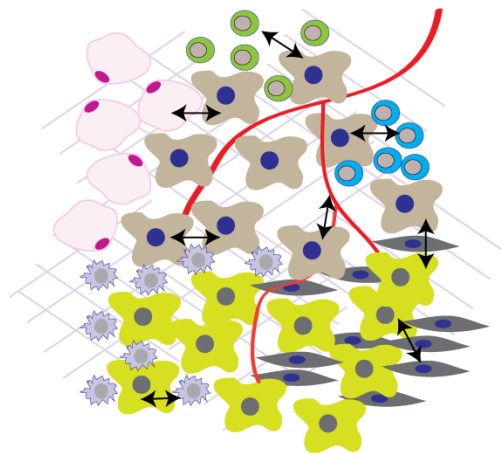

Figure 3-7 │ Cross-talk of tumor cells and their microenvironment influences tumor heterogeneity
Schematic illustration of the interaction and cross-talk between tumor cells and their microenvironment including stromal cells, blood vessels and immune cells. Interclonal and heterotypic cell interactions contribute to tumor heterogeneity.

identity and breast cancer heterogeneity. Similarly, the tumor microvasculature influences disseminated tumor cells, as demonstrated by the promotion of tumor growth by sprouting of neovasculature; in contrast, a stable microvasculature induces tumor cell quiescence and is, thus, tumor suppressive (Ghajar et al., 2013).

The interaction of tumor clones also may affect the tumor landscape (Marusyk et al., 2014; Tabassum and Polyak, 2015). Clonal interaction has been reported between hierarchically different clonal populations in MMTV-*Wnt1* mammary tumors, in which basal and luminal subclones were both required for tumorigenesis. This process was highly dependent on Wnt1 production by the luminal cells, which affected basal clones (Cleary et al., 2014; Tabassum and Polyak, 2015). These examples show that not only cell-intrinsic determinants such as genomic alterations, cell plasticity, and cell-of-origin have a crucial impact on the development of heterogeneous breast cancer, but also interclonal, interlineage and heterotypic interactions.

3.2.5   Consequences of breast tumor heterogeneity

Heterogeneity enhances the robustness of tumors (Marusyk and Polyak, 2010). It can confuse diagnosis and prognosis and challenge cancer therapies. Indeed, while targeted therapy (e.g., Trastuzumab and Lapatinib, Everolimus combined with endocrine therapy) and cytotoxic drugs are efficient strategies to treat some subtypes of breast cancer (Zardavas et al., 2013), high variability in therapeutic response and modest clinical benefit in terms of overall survival occur. As the progression of heterogeneous breast tumors rarely depends on a single pathway, this can be the basis of resistance to targeted therapies (De Palma and Hanahan, 2012; Ramos and Bentires-Alj, 2015). Mutations resulting in resistance may already be present at a very low frequency in the primary tumor or may occur during treatment (Bhang et al., 2015; Ramos and Bentires-Alj, 2015). Despite the elimination of responsive clones, the presence of

resistant clones may reduce treatment success and lead to tumor relapse and therapy failure (Marusyk and Polyak, 2010).

3.2.6 Conclusion and counter-measures of breast tumor heterogeneity

Our knowledge of tumor heterogeneity is presumably only the peak of the iceberg. The complexity and diversity of breast cancers needs to be characterized at the level of individual patients if efficient prognostic and predictive decisions are to achieve long-lasting therapeutic responses. This in-depth understanding is critical to the efficacy of personalized medicine (De Palma and Hanahan, 2012). Predictive preclinical trials may prove useful in which several PDXs from the same patient are established and characterized at the genomic and proteomic levels. The PDXs would be subjected to an *ex vivo* educated screen of drugs or drug combinations based on the results of genomic/proteomic analysis. The most efficient treatments revealed by such a screen would be compared to the current standard of care *ex vivo* and *in vivo*. Based on these data, treatment could be adjusted using approved drugs or patients may be directed to relevant clinical trials. Generating PDXs from different sites within primary tumors and metastases would permit assessment of tumor heterogeneity at different stages of tumor progression and facilitate identification of biomarkers in individual patients. Moreover, personalized PDX models might reveal therapy resistance before this emerges in the patient (Hidalgo et al., 2014; Nardella et al., 2011). Arguably, PDXs established in immunodeficient mice are not subject to the effects of the immune system or species-specific heterotypic signaling interactions between neoplastic cells and recruited stromal cells. The development of humanized mouse models may overcome this potential limitation. In-depth longitudinal studies in which tumor and liquid biopsies at different time points of tumor progression (particularly for metastases) should also help define tumor evolution and allow adjustment of the therapy as needed. These predictive preclinical trials

and longitudinal studies should be applied as a priority to patients who fail first-line therapies. Such a personalized approach has been performed (Bousquet et al., 2014), but it remains costly and is time- and resource demanding. To overcome the cost issue and avoid a two-tiered society, health authorities, the pharmaceutical industry, and health insurance companies should work together to identify a new paradigm for personalized medicine.

Distinct cellular and molecular mechanisms account for the development of tumor heterogeneity. Here, several processes are highlighted that contribute to diversity in breast cancer. It is clear that none of these mechanisms alone explain all facets of breast cancer heterogeneity and, thus, a combination of several factors is most likely involved. These include tumor-initiating alterations in distinct cells-of-origin, cancer cell plasticity, genomic evolution during tumor progression and treatment, as well as reciprocal interaction with the tumor environment. Systems-medicine investigations are warranted in the future to clarify and model all aspects of these multipronged processes during tumor progression and, as a consequence, to develop long-lasting, beneficial therapies.

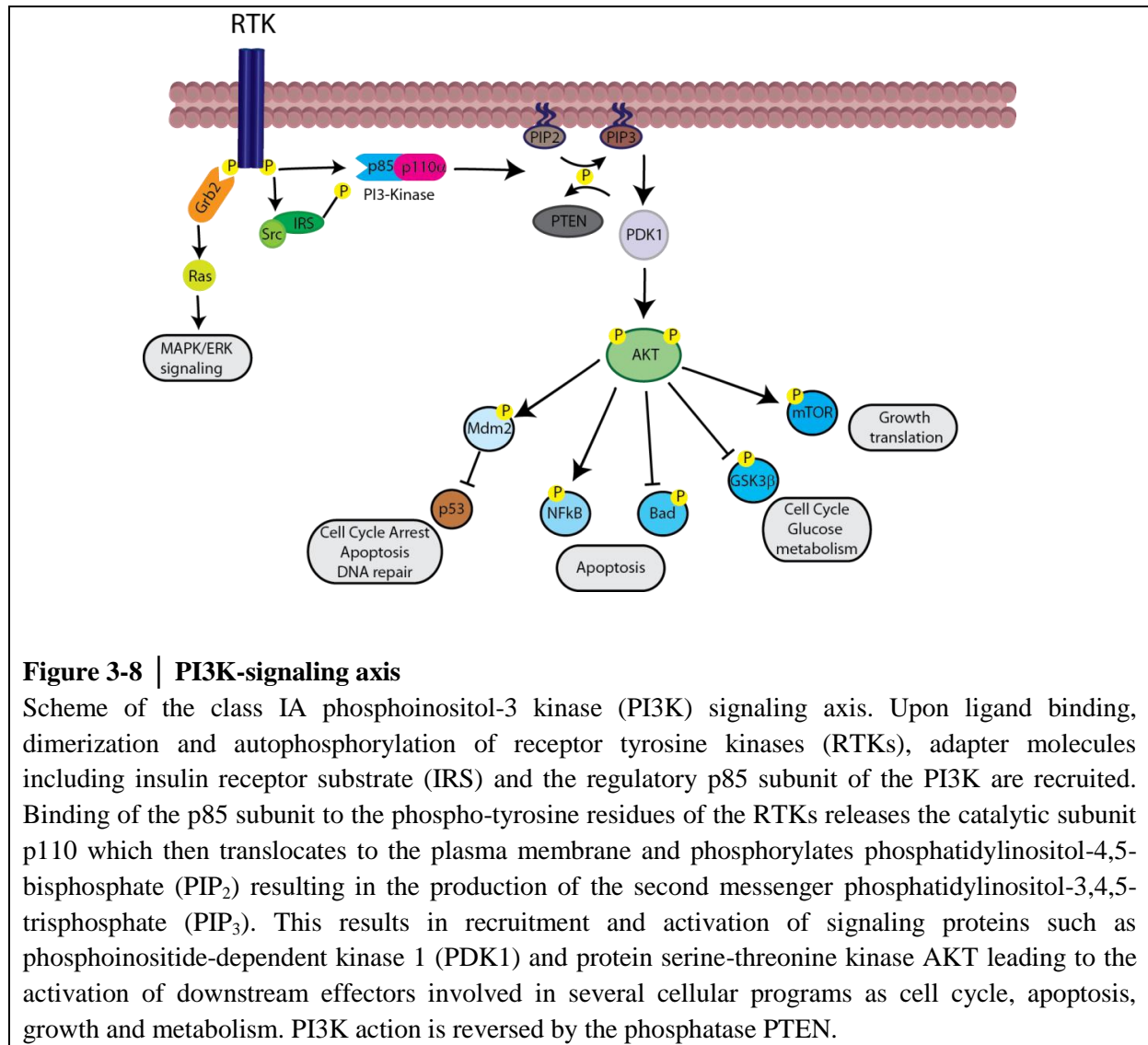## 3.3   Phosphoinositol 3-Kinase (PI3K) signaling

3.3.1   PI3K classification

Phosphoinositide 3-kinases (PI3Ks) belong to a family of lipid kinases involved in metabolism, growth, proliferation and survival signaling. PI3Ks are heterodimers of regulatory (p85α, p85β, p50α, p55α, p55γ) and catalytic (p110α, p110β, p110γ or p110δ) subunits (Yuan and Cantley, 2008; Zhao and Vogt, 2008a). PI3Ks are classified in three classes (I-III) according to their substrate preference and sequence homology. Each class has different roles in cellular signaling (Engelman et al., 2006). Class I PI3Ks preferentially phosphorylate phosphatidylinositol-4,5-bisphosphate (PIP$_2$) to phosphatidylinositol-3,4,5-trisphosphate (PIP$_3$) and are divided into two subgroups depending on the receptors which

activate the signaling cascade. Class IA PI3Ks are activated by growth factor receptor tyrosine kinases (RTKs) while class IB is activated by G-protein-coupled receptors (GPCRs). Class II PI3Ks preferentially phosphorylate phosphatidylinositol and to a lesser extent phosphatidylinositol-4-phosphate (PI-4-P). This class can be activated by RTKs, cytokine receptors and integrins. Class III was initially identified in yeast. The enzyme Vsp34 and its mammalian homologue hVsp34 are predominantly involved in membrane trafficking, endosomal protein sorting, endosome-lysosome maturation, autophagy and cytokinesis (Engelman et al., 2006; Jean and Kiger, 2014)**.**

3.3.2   Class IA PI3K–signaling axis

Class IA PI3Ks are heterodimers of a p85 regulatory subunit and a p110 (α, β or δ) catalytic subunit. In an inactive state, the regulatory p85 subunit interferes with the kinase activity of the catalytic subunit p110. Class IA PI3Ks are activated by growth factor receptor tyrosine kinases (RTKs) including insulin receptor, members of the epidermal growth factor receptor family and platelet-derived growth factor receptor. Ligand binding to the respective receptors leads to receptor dimerization and autophosphorylation, which recruits adapter molecules (e.g., insulin receptor substrate (IRS) 1 or 2) and the regulatory subunit p85. By binding to the phospho-tyrosine residues of the RTKs or adapter molecules, p85 releases the catalytic subunit p110 which then translocates to the plasma membrane and phosphorylates the 3-hydroxyl group of phosphatidylinositol-4,5-bisphosphate (PIP_2), resulting in the production of the second messenger phosphatidylinositol-3,4,5-trisphosphate (PIP_3). This recruits and activates several signaling proteins, including the phosphoinositide-dependent kinase 1 (PDK1) and protein serine-threonine kinase AKT, leading to the activation of their downstream effectors. PI3K action is reversed by the PTEN phosphatase (Cantley, 2002; Engelman, 2009; Yuan and Cantley, 2008) (**Fig. 3-8**).

Figure 3-8 │ PI3K-signaling axis

Scheme of the class IA phosphoinositol-3 kinase (PI3K) signaling axis. Upon ligand binding, dimerization and autophosphorylation of receptor tyrosine kinases (RTKs), adapter molecules including insulin receptor substrate (IRS) and the regulatory p85 subunit of the PI3K are recruited. Binding of the p85 subunit to the phospho-tyrosine residues of the RTKs releases the catalytic subunit p110 which then translocates to the plasma membrane and phosphorylates phosphatidylinositol-4,5-bisphosphate (PIP$_2$) resulting in the production of the second messenger phosphatidylinositol-3,4,5-trisphosphate (PIP$_3$). This results in recruitment and activation of signaling proteins such as phosphoinositide-dependent kinase 1 (PDK1) and protein serine-threonine kinase AKT leading to the activation of downstream effectors involved in several cellular programs as cell cycle, apoptosis, growth and metabolism. PI3K action is reversed by the phosphatase PTEN.

3.4 *PIK3CA* mutations in breast cancer

Genomic alterations of components of the PI3K pathway are found in over 70% of breast cancers (Cancer Genome Atlas, 2012; Miller et al., 2011). The gene *PIK3CA* encodes the catalytic subunit p110α and its amplification and/or mutation is associated with several kinds of human solid tumors (Bachman et al., 2004; Kadota et al., 2009; Levine et al., 2005; Samuels et al., 2004; Wu et al., 2005). Activating somatic mutations in *PIK3CA* are present in approximately 30% of human breast cancers at all stages (Bachman et al., 2004; Barbareschi et al., 2007; Miller, 2012; Saal et al., 2005; Samuels et al., 2004). In 47% of these cases, mutations occur in the kinase domain, the most frequent one is H1047R in exon 20. In 33% of

these cases, mutations occur in the helical domain, the most frequent ones are E545K and E542K in exon 9 (Bader et al., 2005; Samuels et al., 2004). These mutations lead to a constitutively active enzyme with oncogenic capacity *in vitro* and enhance tumorigenicity in xenograft models (Bader et al., 2006; Isakoff et al., 2005; Zhao et al., 2005). It was suggested that kinase and helical domain mutations may trigger gain of function through different mechanisms. While E545K-mutations are independent of binding to the adaptor molecule p85 but require interaction with Ras-GTP, the H1047R-mutants are highly dependent on p85 for its oncogenic capacity but are independent of Ras-GTP (Zhao and Vogt, 2008b). Moreover, it was shown that in contrast to kinase mutants, the helical domain mutants directly associate with IRS1 without growth factor stimulation and without the interaction with p85 (Hao et al., 2013). Additionally, studies showed that alterations in distinct exons of *PIK3CA* have varying impacts on tumor development and progression. One study demonstrated *PIK3CA*$^{\text{H1047R}}$ to be more potent in inducing tumors in transplantation assays *in vivo* (Bader et al., 2006) while another study found no trend (Zhao et al., 2005). The exact impact of these mutations on breast cancer has remained controversial.
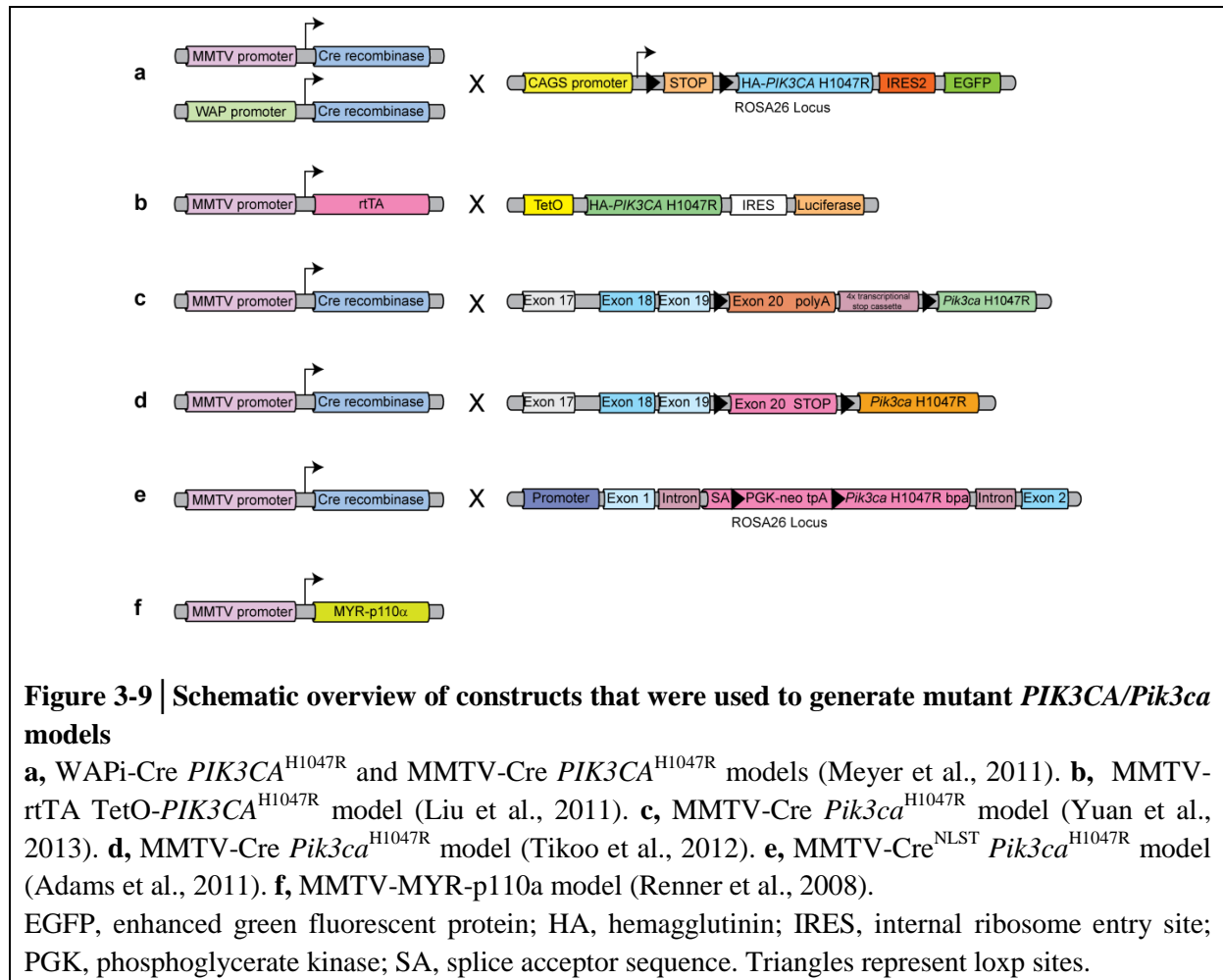
Alterations in *PIK3CA* are found at similar frequencies in pure ductal carcinoma *in situ* (DCIS), DCIS adjacent to invasive ductal carcinoma (IDC), and in IDC, indicating that *PIK3CA* mutations occur early in carcinoma development (Miron et al., 2010). In addition, mutant p110α has been found in distinct human breast cancer subtypes such as ERα-positive, PR-positive, HER2/Neu positive and triple-negative breast cancers (Boyault et al., 2012; Saal et al., 2005), but the correlation between *PIK3CA* mutations and pathological parameters remains controversial (Bachman et al., 2004; Campbell et al., 2004; Dunlap et al., 2010; Dupont Jensen et al., 2011; Kalinsky et al., 2009; Li et al., 2010; Li et al., 2006; Perez-Tenorio et al., 2007; Saal et al., 2005). Also assessment of the clinical outcome associated with these hotspot mutations showed contradictory results: some studies reported poor prognosis in breast cancer patients harboring *PIK3CA* exon 20 (Lai et al., 2008; Mangone et

al., 2012) or exon 9 mutations (Barbareschi et al., 2007) whereas others reported favorable prognosis with improved overall survival in patients with exon 20 mutations (Barbareschi et al., 2007; Kalinsky et al., 2009). Notably, *PIK3CA* mutations were shown to reduce the efficacy of HER2- and ER-targeted therapies (Eichhorn et al., 2008; Kataoka et al., 2010; Miller et al., 2009).

3.5   Mouse models of *PIK3CA* mutations

3.5.1   Tumor formation in *PIK3CA*^{H1047R} mutant mice

Adding to the list of mammary tumor mouse models (Borowsky, 2011), several groups have generated transgenic mice expressing *PIK3CA*^{H1047R} in the mammary gland (Adams et al., 2011; Liu et al., 2011; Meyer et al., 2011; Tikoo et al., 2012; Yuan et al., 2013) **(Fig. 3-9, Table 3-1)**. In contrast to mouse models such as Neu, Myc or polyoma middle T-antigen that evoke tumors with a very specific phenotype, expression of *PIK3CA*^{H1047R} in mice induces heterogeneous tumors. In previous studies two different promoters for conditional mammary-specific expression of human *PIK3CA*^{H1047R} or murine *Pik3ca*^{H1047R} were used to drive Cre recombinase expression. First, Cre driven by the MMTV-Cre promoter results in mosaic expression of mutant *PIK3CA/Pik3ca* in differentiated mammary luminal cells and progenitor cells, and in further organs, depending on the MMTV-Cre line (Andrechek et al., 2005; Soriano, 1999; Srinivas et al., 2001; Tikoo et al., 2012; Wagner et al., 2001; Wagner et al., 1997). Second, Cre driven by the WAPi-Cre (Wintermantel et al., 2002) results in expression of mutant *PIK3CA*^{H1047R} in alveolar progenitor cells and differentiated secretory luminal cells (Meyer et al., 2011) **(Fig. 3-9a)**. Tetracycline-inducible promoter systems (combined with MMTV-rtTA (Gunther et al., 2002)) were also used to drive overexpression of H1047R leading to a 7-8 fold change in expression of mutant *PIK3CA* compared to endogenous *Pik3ca* (Liu et al., 2011) **(Fig. 3-9b).** Other groups have used a knock-in system to express

Figure 3-9 │ Schematic overview of constructs that were used to generate mutant *PIK3CA/Pik3ca* models

a, WAPi-Cre *PIK3CA*[H1047R] and MMTV-Cre *PIK3CA*[H1047R] models (Meyer et al., 2011). **b,** MMTV-rtTA TetO-*PIK3CA*[H1047R] model (Liu et al., 2011). **c,** MMTV-Cre *Pik3ca*[H1047R] model (Yuan et al., 2013). **d,** MMTV-Cre *Pik3ca*[H1047R] model (Tikoo et al., 2012). **e,** MMTV-Cre[NLST] *Pik3ca*[H1047R] model (Adams et al., 2011). **f,** MMTV-MYR-p110a model (Renner et al., 2008).

EGFP, enhanced green fluorescent protein; HA, hemagglutinin; IRES, internal ribosome entry site; PGK, phosphoglycerate kinase; SA, splice acceptor sequence. Triangles represent loxp sites.

endogenous levels of *Pik3ca*[H1047R] under the control of the native promoter (combined with MMTV-Cre (Wagner et al., 2001; Wagner et al., 1997)) (Tikoo et al., 2012; Yuan et al., 2013) **(Fig. 3-9c,d)**.

Meyer *et al.* showed mammary-tumor independent high lethality (~75%) in MMTV-Cre *PIK3CA*[H1047R] mice. Whilst the cause of death could not be identified, promoter leakiness leading to expression of *PIK3CA*[H1047R] mutant in other tissues was suggested (Meyer et al., 2011). Using two different MMTV-Cre lines (Soriano, 1999; Srinivas et al., 2001; Wagner et al., 1997) to induce expression of the *PIK3CA*[H1047R] mutation **(Fig. 3-9e)**, Adams *et al.* found that some of MMTV-Cre[lineA] *Pik3ca*[H1047R] and MMTV-Cre[NLST] *Pik3ca*[H1047R] mice reached endpoint (e.g., lethargy, impaired breathing, tumors) independently of mammary tumors

Mouse model	Mean age at tumor onset	Pathology	Ref.
MMTV- myr-*PIK3CA* and MMTV-myr-*PIK3CA* ; p53[+/-]	Not reported	Nulliparous: adenosquamous carcinoma; Parous: 66% adenosquamous carcinoma, 34% carcinoma; (all ER+, cathepsin D+)	(Renner et al., 2008)
MMTV-myr-*PIK3CA* CDK4(R24C)	Not reported	Nulliparous: adenosquamous carcinoma, papillary adenocarcinoma, carcinoma, sarcoma; Parous: adenosquamous carcinoma, complex adenocarcinoma, carcinoma, sarcoma (all ER+ except sarcoma)	(Renner et al., 2008)
MMTV-Cre[NLST] *PIK3CA*[H1047R]	5 months	Adenosquamous carcinoma (51%; ER+, K8+, K14+, K8/14+, K8/14-, N-Cadherin+, Vimentin+, Atf3+, K10+, β-Catenin+), adenomyoepithelioma (45%; ER+, K8+, K14+, N-Cadherin+, Atf3+, Desmin+, β-Catenin+), spindle cell tumors (1%), poorly differentiated adenocarcinoma (3%)	(Adams et al., 2011)
MMTV-Cre[NLST] *PIK3CA*[H1047R] p53[fl/+]	< 5 months	Spindle cell/EMT tumors (33%, ER+, K8+, K14+, N-Cadherin+, Desmin+), adenosquamous carcinoma (52%, ER+, K8+, K14+, K8/14+, K8/14-, N-Cadherin+, Desmin-, K10+), radial scar type lesions (10%) and poorly differentiated adenocarcinoma (5%, ER+, K8+, K14+, K8/14+, K8/14-)	(Adams et al., 2011)
MMTV-rtTA TetO-*PIK3CA*[H1047R]	7 months	Adenocarcinoma and adenosquamous carcinoma	(Liu et al., 2011)
WAPi-Cre *PIK3CA*[H1047R]	Parous: 140.3 (±6.9) days (=36.8±4.9 days after delivery); nulliparous: 219 (± 12) days	Adenosquamous carcinoma (54.6%), adenomyoepithelioma (22.7%, PR+), adenocarcinoma with squamous metaplasia (13.6%), adenocarcinoma (9.1%) (all ER+, K14+, K18+, K14/18+)	(Meyer et al., 2011)
MMTV-Cre *PIK3CA*[H1047R]	214 (±22.6) days	Adenomyoepithelioma (100%) (ER+, PR+, K14+, K18+ and α-SMA)	(Meyer et al., 2011)
MMTV-Cre *PIK3CA*[H1047R]	Parous: 465 days; nulliparous: 492 days	Fibroadenoma (76.9%), adenocarcinoma (15.4%) (both K5+, K18+, ER+ and PR+); spindle cell neoplasia (7.7%, ER-,PR- ,K5-, K18+ ,Vimentin+)	(Yuan et al., 2013)
MMTV-Cre *PIK3CA*[H1047R]	Parous: 393 days; nulliparous: 484 days	Benign fibroadenoma (45%), carcinosarcoma (both ER+, K5/6+, K8/18+, K5/K8+, K8/E-Cadherin+) or sarcoma (42.5%); adenosquamous carcinoma (10%) (K5/6+, K8-, E-Cadherin-); osteosarcoma (2.5%)	(Tikoo et al., 2012)

Table 3-1│ Summary of mouse models of *PIK3CA* alterations

(Adams et al., 2011). These observations raise the concern whether MMTV-Cre is the optimal promoter system to study *PIK3CA*-induced mammary cancer in mice. Each of these systems leads to the development of heterogeneous mammary tumors. The most prominent phenotypes, adenosquamous carcinoma and adenomyoepithelioma express ERα, as well as basal (e.g., keratin 5 and 14) and luminal markers (e.g., keratin 8 and 18). The *Pik3ca*[H1047R] knock-in models of Yuan *et al.* and Tikoo *et al.* led mostly to hormone receptor-positive fibroadenomas (76.9% and 45%, respectively) or sarcomas (42.5%). Other histopathological features such as adenocarcinoma, carcinosarcoma and osteosarcoma were also observed (Tikoo et al., 2012; Yuan et al., 2013) **(Table 3-1)**. Heterogeneity, a feature of human breast cancer, was also reported in mouse models of *Pten* inactivation (Li et al., 2002; Stambolic et al., 2000). In another study, however, loss of *PTEN* resulted only in adenomyoepithelioma (Dourdin et al., 2008) (for a Review of PTEN mouse models, please see (Hollander et al., 2011)).

Tumor heterogeneity and the observation that *PIK3CA*[H1047R] mutants develop keratin 5/14- and keratin 8/18-positive mammary carcinomas (Adams et al., 2011; Meyer et al., 2011; Tikoo et al., 2012; Yuan et al., 2013) suggest either a luminal and basal tumor cell-of-origin, transdifferentiation or the dedifferentiation of cell lineage committed tumor cells to multipotent progenitors that then give rise to keratin 5/14- and keratin 8/18-positive cells. Tikoo *et al.* found that expression of *Pik3ca*[H1047R] results in an expansion of the luminal progenitor population. Furthermore, the putative mammary stem cell-enriched basal population and the luminal progenitors of mutants display enhanced colony-forming ability and a larger colony size (Tikoo et al., 2012), but the molecular mechanisms underlying these effects have not yet been defined. The current mouse models do not definitely address the question what causes heterogeneity in *PIK3CA*[H1047R]-evoked tumors and which cell type gives rise to which subtype of mammary cancer. Lineage-tracing experiments should provide further information about the cell-of-origin and cellular hierarchy in tumors.

Nulliparous animals developed tumors at an average of 219 days in WAPi-Cre *PIK3CA*[H1047R] animals (S.K. and M.B.A., unpublished data) and 484 and 492 days in MMTV-Cre *Pik3ca* knock-in H1047R animals (Tikoo et al., 2012; Yuan et al., 2013). Notably, tumor latencies in parous animals ranged from an average of 140 days in WAPi-Cre *PIK3CA*[H1047R] animals (Meyer et al., 2011) to 392 and 465 days in MMTV-Cre *Pik3ca* knock-in H1047R animals (Tikoo et al., 2012; Yuan et al., 2013), showing that pregnancy accelerates tumorigenesis in these models. An increase in the number of H1047R-expressing cells after pregnancy and a delay in involution, due to a reduced number of apoptotic cells, was observed in WAPi-Cre *PIK3CA*[H1047R] mice (Meyer et al., 2011).

Tumor formation was also investigated using MMTV-driven expression of non-mutated p110α fused to a Src myristolyation sequence (**Fig. 3-9f**), which results in the recruitment of p110α to the membrane and constitutive activation of PI3K signaling (MMTV-MYR-p110α). Transgenic mice developed heterogeneous ER-positive mammary tumors but at a frequency lower than mice expressing mutant *PIK3CA* (Renner et al., 2008).

3.5.2 Synergism between *PIK3CA*[H1047R] and *P53* alterations

Whole-exome capture and sequencing of mammary tumors from MMTV-Cre knock-in *Pik3ca*[H1047R] mice of various histotypes has revealed an increase in somatic mutations in spindle cell tumors (~44-88) and adenocarcinoma (~4-61) more than in fibroadenoma (~2-13) (Yuan et al., 2013). Moreover, comparative genomic hybridization (CGH)-array profiling showed a greater accumulation of chromosomal copy number alterations in spindle cell tumors than in adenocarcinoma and fibroadenoma, compared with normal mammary glands (Yuan et al., 2013). Functional validation and examination of the clinical relevance of these secondary genomic alterations are warranted.

The most common alterations in breast cancer are found in the PI3K- and p53-signaling pathways (Cancer Genome Atlas, 2012). Some human breast tumors harbor alterations in *PIK3CA* in combination with mutant *P53* (Boyault et al., 2012; Buttitta et al., 2006; Maruyama et al., 2007). *P53* mutations (R245H, A135V, I192N) were among the secondary mutations identified by Yuan *et al.* (Yuan et al., 2013) in adenocarcinoma and spindle cell neoplasia. It is likely that these mutations prevented the well-established p53-dependent tumor suppression. Notably, *Pik3ca* mutant mouse models were used to investigate the interaction of *Pik3ca*^{H1047R} and *p53* (Adams et al., 2011; Yuan et al., 2013). Heterozygosity in *p53* was shown to accelerate tumor onset in MMTV-Cre *Pik3ca*^{H1047R} mutant mice (Adams et al., 2011). The tumor histotype in double mutants consisted mostly of ER-, K14- and K8-positive spindle cell tumors that express EMT markers, or adenosquamous carcinoma (Adams et al., 2011) **(Table 3-1)**.

p53 is found inactivated in MMTV-MYR-p110α-evoked tumors suggesting that p53 loss is important for tumorigenesis in this model. No difference was found in tumor latency or tumor phenotype between MMTV-MYR-p110α mice in a heterozygous *p53* background and MMTV-MYR-p110α mice (Renner et al., 2008). Notably, MMTV-MYR-p110α mice in an inactive pRB background (CDK4 R24C knock-in line (Quereda et al., 2007; Sotillo et al., 2001)) showed enhanced mammary tumorigenesis. These data suggest that tumor suppression mechanisms can be circumvented by inactivation of either p53 or pRB in *PIK3CA*^{H1047R} mutant tumors (Renner et al., 2008).

3.5.3 *PIK3CA*^{H1047R} mutations and metastasis

Mouse models of altered PI3K pathway can increase our understanding of breast cancer progression and metastatic spread. Metastases were reported in *Pten* heterozygous mice: One study found a metastatic tumor in the regional lymph node of one mouse, while three other

animals had lung metastases. Their morphological appearance was similar to the primary tumor (Stambolic et al., 2000). In contrast, metastases are rarely found in *PIK3CA/Pik3ca*[H1047R] mouse models. There is only one report of isolated lung metastasis in MMTV-Cre[NLST] *Pik3ca*[H1047R] mice (Adams et al., 2011). However, *PIK3CA* mutations occur at high frequencies in metastatic human breast cancer. The comparison of primary tumors with matched metastases revealed 42% and 20% of exon 20 (H1047R/H1047L) or exon 9 (E545K/E542K) mutations in the corresponding metastatic tumors, respectively. Cases were reported in which matched metastases from *PIK3CA* wildtype primary tumors gained *PIK3CA* mutations or in which metastases from *PIK3CA* mutant primary tumors lost its mutation. Moreover, mutations in both exons occur (Dupont Jensen et al., 2011). Surprisingly, oncogenic *PIK3CA*-driven breast tumors have a longer time to recurrence after surgery (Dupont Jensen et al., 2011) and some clinical studies reported a good prognosis (Barbareschi et al., 2007; Kalinsky et al., 2009). These observations may mean that mutant *PIK3CA* results in a selective advantage for breast cancer cells at the primary site but not during metastatic progression and colonization of distant sites. Further analysis of *PIK3CA* status in a large number of metastatic lesions, circulating tumor cells and matched primary tumors should clarify this "*PIK3CA* paradox" (Meyer and Bentires-Alj, 2010). An alternative explanation is that patients with mutations in *PIK3CA* respond well to current standard of care resulting in this apparent paradox.

Constitutively active PI3K signaling, in association with further genomic alterations, induces mammary cancer in mice, which suggests a causative role for *PIK3CA* mutations in breast tumorigenesis. Additional gain- or loss-of-function genomic alterations may also contribute to metastasis in breast tumor progression. The identification of these synergistic oncogenic pathways is of paramount importance for the elucidation of the "wiring diagram" of tumors cells with *PIK3CA* mutations.

3.5.4 Therapeutic strategies and resistance in mutant *PIK3CA* mouse models

Several PI3K pathway-targeting compounds are currently being evaluated in cancer-related clinical trials (Leroy et al., 2014). *PIK3CA* mouse models also serve as a valuable tool for testing anti-cancer drugs (Courtney et al., 2010; Engelman et al., 2008). A test of the efficacy of the PI3K inhibitor GDC-0941 by Yuan *et al.* showed a decline in tumor growth in spindle cell tumors with *Pik3ca*[H1047R] and *P53* mutations (Yuan et al., 2013).

These mouse models have already proved useful for investigating resistance mechanisms. For example, *PIK3CA*[H1047R]-driven tumors were shown to recur after *PIK3CA*[H1047R] inactivation in a PI3K-dependent or -independent manner (Liu et al., 2011). Tumor survival in *c-MET* elevated tumors was shown to depend on an active endogenous PI3K pathway, whereas *c-MYC* elevation contributed to oncogene independence and GDC-0941 resistance (Liu et al., 2011). The PI3K-independent recurrence of *PIK3CA*[H1047R]-initiated mammary tumors shows how important it is to investigate associated pathways involved in tumor formation that may result in escape from treatment. Their delineation should pave the way for the development of mechanism-based combination therapies.

Parts of this introduction were published in

FEBS J. 2013 Jun;280(12):2758-65. doi: 10.1111/febs.12175. Epub 2013 Mar 1.
Mouse models of *PIK3CA* mutations: one mutation initiates heterogeneous mammary tumors.
Shany Koren and Mohamed Bentires-Alj.
The full published review can be found in **appendix 10.4**.

and are *in press* in

Molecular Cell. 2015
Breast tumor heterogeneity: Source of fitness, hurdle for therapy.
Shany Koren and Mohamed Bentires-Alj.

---

4 | Rationale of the work

Tumor heterogeneity impinges on prognosis, response to therapy, and metastasis. As such, heterogeneity is one of the most important and clinically relevant areas of cancer research. In-depth characterization and understanding of the origin of this phenotypic and molecular diversity is paramount to improving diagnosis, the definition of prognostic and predictive biomarkers, and the design of therapeutic strategies.

Breast cancer displays frequent intra- and intertumor heterogeneity as the result of genetic and non-genetic alterations that often enhance the vigor of cancer cells. Heterogeneity is found in histopathological subtypes, molecular profiles, clinical outcomes and responses to therapy. This heterogeneity suggests that distinct breast cancer subtypes derive from distinct epithelial cell subtypes in the mammary gland (Ince et al., 2007) and/or that cells acquire plasticity upon transformation (Brooks et al., 2015; Skibinski and Kuperwasser, 2015). As histology of tumors does not necessarily reflect the nature of the cell-of-origin (Lim et al., 2009; Molyneux et al., 2010), the key to understand pathogenesis of cancer is the identification of the target cells of transformation. Moreover, delineating mammary cell hierarchy and the consequences of transforming events on lineage organization is important, not only for understanding the cell-of-origin but also the heterogeneity and the aggressiveness of breast cancer.

As *PIK3CA* is the most frequent gain of function mutation in human breast cancer (Cancer Genome Atlas, 2012) and it can give rise to heterogeneous mammary tumors in animal models, it is of paramount importance to investigate the effect of oncogenic *PIK3CA* on tumor heterogeneity and aggressiveness. Several *PIK3CA*[H1047R]-expressing mouse models were developed (Adams et al., 2011; Liu et al., 2011; Meyer et al., 2011; Tikoo et al., 2012; Yuan et al., 2013). However, whether expression of *PIK3CA*[E545K] mutants also evokes mammary tumors *in vivo* and how they differ from *PIK3CA*[H1047R]-expressing tumors

remained elusive. Moreover, the target cell of transformation in mutant *PIK3CA*-evoked tumors has remained unknown. Expression of mutant *PIK3CA* in specific subpopulations of the mammary gland can reveal associations between targeted cell and tumor incidence, progression and phenotype. Additionally, the impact of mutant *PIK3CA*^{H1047R} on fate switching and cell plasticity during mammary tumor development has remained largely unclear. Its identification is essential for understanding tumor heterogeneity.

Specifically I wanted to elucidate the effect of:

i. *PIK3CA*^{E545K} expression on mammary tumorigenesis and tumor heterogeneity *in vivo*.

ii. *PIK3CA*^{H1047R} expression on the mammary cell hierarchy *in vivo*.

iii. *PIK3CA*^{H1047R} expression in basal Lgr5-positive and luminal keratin 8-positive mammary cells on tumor onset, heterogeneity and phenotype.

5 | Results Part I

Effects of *PIK3CA*[E545K] expression on mammary tumorigenesis

Research article published in Oncogenesis. 2013 Sep 30;2:e74. doi: 10.1038/oncsis.2013.38

(shared first authorship)

SHORT COMMUNICATION

Expression of *PIK3CA* mutant E545K in the mammary gland induces heterogeneous tumors but is less potent than mutant H1047R

DS Meyer[1,4,5], S Koren[1,4], C Leroy[1,2], H Brinkhaus[1], U Müller[1], I Klebba[1,6], M Müller[2], RD Cardiff[3] and M Bentires-Alj[1]

[1]Mechanisms of Cancer, Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland; [2]Developmental and Molecular Pathways, Novartis Institutes for Biomedical Research, Basel, Switzerland and [3]Department of Pathology, Center for Comparative Medicine, University of California Davis, Davis, CA, USA. Correspondence: Dr M Bentires-Alj, Mechanisms of Cancer, Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058, Basel, Switzerland.
E-mail: bentires@fmi.ch
[4]These authors contributed equally to this work.
[5]Current address: Helen Diller Family Comprehensive Cancer Center, School of Medicine, University of California San Francisco (UCSF), San Francisco, CA, USA.
[6]Current address: Department of Biosystems Science and Engineering (D-BSSE), ETH Zürich, Basel, Switzerland.
Received 17 June 2013; revised 13 August 2013; accepted 20 August 2013

The full published article can be found in **appendix 10.4**.
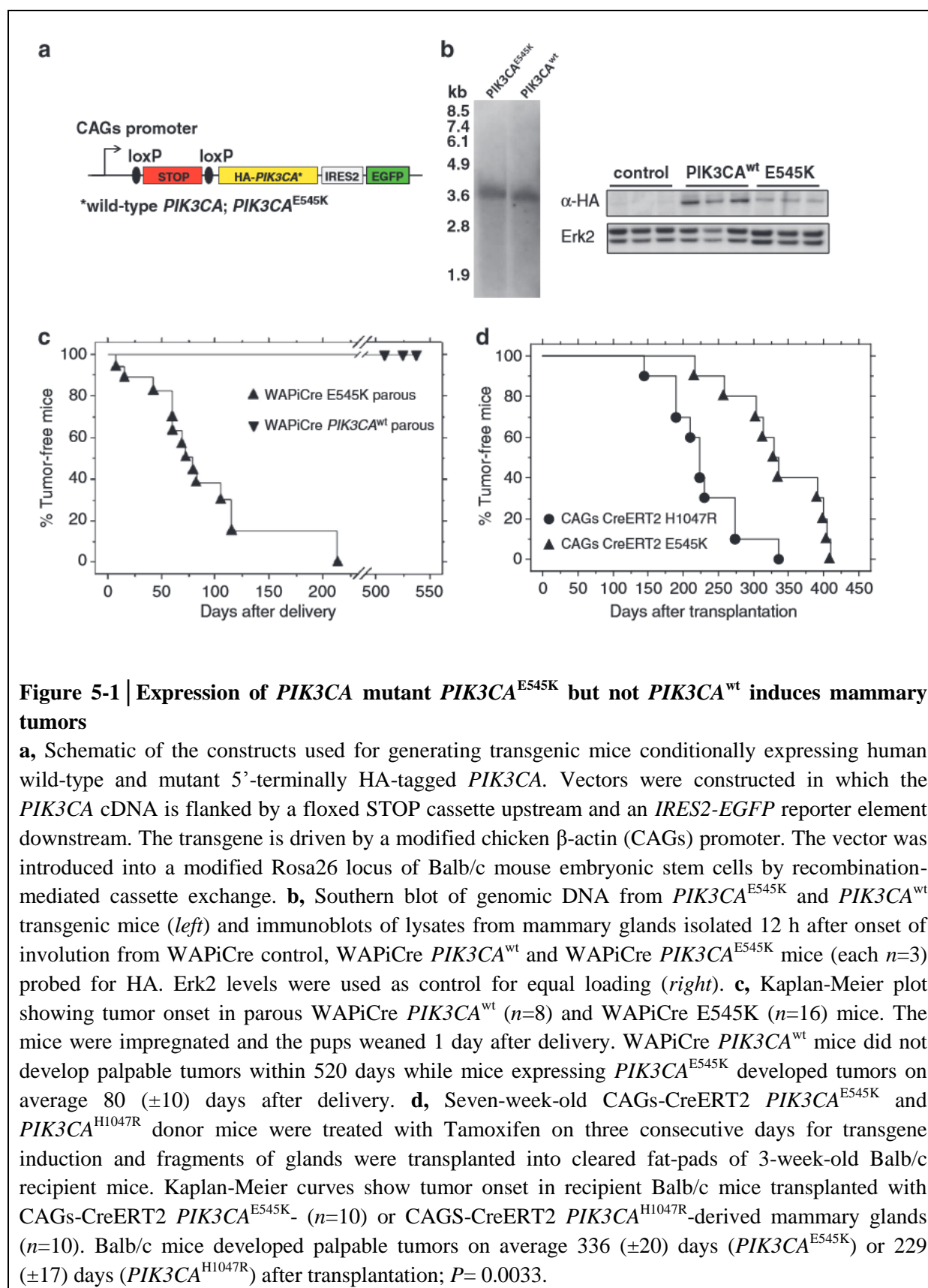
36

5.1 Summary

The PI3K signaling cascade is a key mediator of cellular growth, survival and metabolism and is frequently subverted in human cancer. The gene encoding for the alpha catalytic subunit of PI3K (*PIK3CA*) is mutated and/or amplified in ~30% of breast cancers. Mutations in either the kinase domain (H1047R) or the helical domain (E545K) are most common and result in a constitutively active enzyme with oncogenic capacity. *PIK3CA*[H1047R] was previously demonstrated to induce tumors in transgenic mouse models; however, it was not known whether overexpression of *PIK3CA*[E545K] is sufficient to induce mammary tumors and whether tumor initiation by these two types of mutants differs. Here, we demonstrate that expression of *PIK3CA*[E545K] in the mouse mammary gland induces heterogeneous mammary carcinomas but with a longer latency than *PIK3CA*[H1047R]-expressing mice. Our results suggest that the helical domain mutant *PIK3CA*[E545K] is a less potent inducer of mammary tumors due to less efficient activation of downstream Akt signaling.

5.2 Results

5.2.1 Expression of *PIK3CA*[E545K] but not wild-type *PIK3CA* induces mammary tumors

Previously it was shown that *PIK3CA*[H1047R] induces mouse mammary carcinomas (Adams et al., 2011; Liu et al., 2011; Meyer et al., 2011; Tikoo et al., 2012). To test whether expression of wild-type human *PIK3CA* (*PIK3CA*[wt]) or *PIK3CA*[E545K] also induces mammary tumors, we generated novel transgenic mice that conditionally express *PIK3CA*[wt] or *PIK3CA*[E545K] (**Fig. 5-1a**). To achieve equivalent transgene expression, we integrated *PIK3CA*[wt] or *PIK3CA*[E545K] into the ROSA26 locus using RMCE-mediated recombination (Tchorz et al., 2009). Correct integration of the target cassettes was confirmed in the resulting *PIK3CA*[wt] and *PIK3CA*[E545K] lines (**Fig. 5-1b,** *left*). *PIK3CA*[wt] and *PIK3CA*[E545K] animals were then crossed to WAPiCre mice in which expression of recombinase Cre is controlled by the whey acidic protein (WAP) promoter, which is mainly active in secretory mammary epithelial cells (Booth et al., 2007; Boulanger et al., 2005; Bruno and Smith, 2011; Wintermantel et al., 2002), and expression of the transgenes confirmed (**Fig. 5-1b,** *right*). This enabled us to directly compare the kinetics of tumor onset in *PIK3CA*[wt] and *PIK3CA*[E545K] mice, and the previously reported WAPiCre *PIK3CA*[H1047R] mice (Meyer et al., 2011).

The resulting bi-transgenic WAPiCre *PIK3CA*[wt] and WAPiCre *PIK3CA*[E545K] female mice were impregnated to achieve maximal Cre-mediated recombination and the pups removed 1 day after delivery. All WAPiCre *PIK3CA*[E545K] mice developed mammary tumors on average 80 (±10) days after delivery, while parous WAPiCre *PIK3CA*[wt] mice did not form tumors within 520 days (**Fig. 5-1c**). This indicates that overexpression of wild-type *PIK3CA* itself is insufficient to induce mammary tumors. Of note, the latency to tumor onset in WAPiCre *PIK3CA*[E545K] animals was significantly longer than that observed previously for WAPiCre *PIK3CA*[H1047R] mice (36 (± 4.9) days) (Meyer et al., 2011). We also crossed *PIK3CA*[E545K] and *PIK3CA*[H1047R] lines to CAGs-CreERT2 mice that express a tamoxifen-inducible Cre/estrogen receptor (ER) fusion protein under the control of a modified β-actin

Figure 5-1│Expression of *PIK3CA* mutant *PIK3CA*[E545K] but not *PIK3CA*[wt] induces mammary tumors

a, Schematic of the constructs used for generating transgenic mice conditionally expressing human wild-type and mutant 5'-terminally HA-tagged *PIK3CA*. Vectors were constructed in which the *PIK3CA* cDNA is flanked by a floxed STOP cassette upstream and an *IRES2-EGFP* reporter element downstream. The transgene is driven by a modified chicken β-actin (CAGs) promoter. The vector was introduced into a modified Rosa26 locus of Balb/c mouse embryonic stem cells by recombination-mediated cassette exchange. **b,** Southern blot of genomic DNA from *PIK3CA*[E545K] and *PIK3CA*[wt] transgenic mice (*left*) and immunoblots of lysates from mammary glands isolated 12 h after onset of involution from WAPiCre control, WAPiCre *PIK3CA*[wt] and WAPiCre *PIK3CA*[E545K] mice (each *n*=3) probed for HA. Erk2 levels were used as control for equal loading (*right*). **c,** Kaplan-Meier plot showing tumor onset in parous WAPiCre *PIK3CA*[wt] (*n*=8) and WAPiCre E545K (*n*=16) mice. The mice were impregnated and the pups weaned 1 day after delivery. WAPiCre *PIK3CA*[wt] mice did not develop palpable tumors within 520 days while mice expressing *PIK3CA*[E545K] developed tumors on average 80 (±10) days after delivery. **d,** Seven-week-old CAGs-CreERT2 *PIK3CA*[E545K] and *PIK3CA*[H1047R] donor mice were treated with Tamoxifen on three consecutive days for transgene induction and fragments of glands were transplanted into cleared fat-pads of 3-week-old Balb/c recipient mice. Kaplan-Meier curves show tumor onset in recipient Balb/c mice transplanted with CAGs-CreERT2 *PIK3CA*[E545K]- (*n*=10) or CAGS-CreERT2 *PIK3CA*[H1047R]-derived mammary glands (*n*=10). Balb/c mice developed palpable tumors on average 336 (±20) days (*PIK3CA*[E545K]) or 229 (±17) days (*PIK3CA*[H1047R]) after transplantation; *P*= 0.0033.

promoter; this results in the expression of Cre-ER in virtually all cells. Unexpectedly, bi-transgenic CAGs-Cre $PIK3CA^{E545K}$ and $PIK3CA^{H1047R}$ mice died by the age of 4 months even when no Tamoxifen was administered. Although we were unable to identify the exact cause of death, we concluded that leakiness of the CAGs-CreERT2 system caused premature and deleterious $PIK3CA^{E545K}$ or $PIK3CA^{H1047R}$ expression in various tissues of these mice (D.S.M. and M.B-A., unpublished observations).

To compare the tumor-initiating potential of the two different $PIK3CA$ mutants, we then transplanted pieces of mammary gland tissue from CAGs-CreERT2 $PIK3CA^{E545K}$ or $PIK3CA^{H1047R}$ donor mice previously treated with Tamoxifen into cleared fat pads of Balb/c mice. The mammary glands reconstituted by either CAGs-CreERT2 $PIK3CA^{E545K}$ or $PIK3CA^{H1047R}$-derived epithelium were hyperplastic (data not shown) and eventually formed tumors after 229 (±17, $PIK3CA^{H1047R}$) and 336 days (±20, $PIK3CA^{E545K}$), respectively (**Fig. 5-1d**). As observed in the WAPiCre mouse cohorts, $PIK3CA^{E545K}$ was significantly less potent than $PIK3CA^{H1047R}$ in the induction of mammary carcinomas, which is a possible explanation for the lower frequency of E542K/E545K mutations in human breast cancer (Cancer Genome Atlas, 2012; Saal et al., 2005).

5.2.2   WAPiCre $PIK3CA^{E545K}$-evoked mammary tumors are heterogeneous

Examination of 30 WAPiCre $PIK3CA^{E545K}$-derived tumors identified six distinct histotypes. By far the most prevalent tumor phenotype was adenosquamous carcinoma (60%) (**Fig. 5-2a,b**), which was also the most common histotype formed by WAPiCre $PIK3CA^{H1047R}$ mice (54.6%) (Meyer et al., 2011). Adenocarcinomas (23.3%) and carcinomas (6.7%) were also observed albeit at lower frequencies (**Fig. 5-2a,b**). An adenocarcinoma with squamous metaplasia (3.3%), an adenomyoepithelioma (3.3%) and spindle cell tumor (3.3%) were observed in one tumor only (**Fig. 5-2a,b**). The low frequency of adenomyoepithelioma in

WAPiCre *PIK3CA*^E545K mice is in stark contrast to the WAPiCre *PIK3CA*^H1047R animals, in which adenomyoepitheliomas accounted for ~23% of the tumors (Meyer et al., 2011). A further discrepancy between mice expressing *PIK3CA*^E545K or *PIK3CA*^H1047R was the complete absence of diffuse and invasive adenocarcinomatosis in WAPiCre *PIK3CA*^E545K-derived glands, a histological feature that was displayed by all tumor-surrounding tissue in WAPiCre *PIK3CA*^H1047R mice (Meyer et al., 2011).

The *PIK3CA*^E545K-induced tumors were stained for luminal keratin 8/18 (K8/18), basal/myoepithelial keratin 14 (K14), and myoepithelial α-smooth muscle actin (α-SMA) markers. The most frequent histotypes, adenosquamous carcinoma and adenocarcinoma, were positive for both luminal K8/18 and basal K14 **(Fig. 5-2c)**. In tumors of the adenosquamous carcinoma type, the relative tumor areas positive for K8/18 and K14 were ~35% and ~39%, respectively and largely negative for α-SMA (<1%) **(Fig. 5-2d)**. WAPiCre *PIK3CA*^E545K-evoked adenosquamous carcinomas also stained positive for ER (~8% of the tumor cells) and displayed a high proportion of Ki-67-positive cells (~35%) **(Fig. 5-2c,d)**. The relative tumor areas and cells positive for K14, K8/18, α-SMA, ER and Ki-67 were very similar to those observed in *PIK3CA*^H1047R-driven adenosquamous carcinomas (Meyer et al., 2011).

However, the number of apoptotic cells staining positively for cleaved caspase-3 was higher in *PIK3CA*^E545K-evoked adenosquamous carcinomas (~7%) **(Fig. 5-2c,d)** than those derived from WAPiCre *PIK3CA*^H1047R mice (~1%) (Meyer et al., 2011), indicating that *PIK3CA*^H1047R is a more potent suppressor of apoptosis than *PIK3CA*^E545K in mammary tumors.

In summary, both *PIK3CA*^E545K and *PIK3CA*^H1047R produced K14/K8/18-positive tumors of various histotypes, with the adenosquamous carcinoma type being the most common in both transgenic models. However, differences between the mouse models included low abundance of adenomyoepitheliomas and the absence of adenocarcinomatosis in WAPiCre *PIK3CA*^E545K mice.

Figure 5-2│WAPiCre *PIK3CA*^E545K^-evoked tumors are heterogeneous and express basal and luminal keratins

a, Diagram showing relative abundance of adenosquamous carcinoma (60%, *red*), adenocarcinoma (23.3%, *green*), carcinoma (6.7%, *purple*), adenocarcinoma with squamous metaplasia (3.3%, *dark blue*), adenomyoepithelioma (3.3%, *light blue*) and spindle cell tumor (3.3%, *orange*) among tumors (*n* =30) from parous WAPiCre *PIK3CA*^E545K^ mice. **b,** H&E-stained tumor sections of the indicated histotypes from parous WAPiCre *PIK3CA*^E545K^ mice. The top-left image shows a representative adenosquamous carcinoma with glands and squamous features. The top-right image shows an adenocarcinoma; the arrows indicate the gland lumen. The glands are lined by malignant epithelium. The center-left image shows an adenocarcinoma with squamous metaplasia; the asterisk shows an area with glands and the arrow indicates areas of metaplasia. The center-right image shows a spindle cell tumor with possible osseous metaplasia, intense pink stroma and large cells in the interstices. The bottom image shows an adenomyoepithelioma. Scale bar = 100 µm. **c,** Immunostaining for K14, K8/18, α-SMA, ER, Ki67 and cleaved caspase-3. Scale bars = 50 µm. **d,** Quantification of immunostaining for K8/18, K14, α-SMA, ER, Ki67 and cleaved caspase-3 isolated from parous WAPiCre *PIK3CA*^E545K^-evoked adenosquamous carcinomas (*n*=8). The data are presented as percentages of positive tumor area and tumor cells. Histological features of WAPiCre *PIK3CA*^E545K^-evoked adenosquamous carcinomas are compared to those of WAPiCre *PIK3CA*^H1047R^-evoked adenosquamous carcinomas previously reported (Meyer et al., 2011) (indicated by an asterisk).

5.2.3 Pregnancy accelerates *PIK3CA*-evoked tumorigenesis and *PIK3CA* mutants delay mammary gland involution

The variations in tumor histotypes and the discrepancy in tumor latency in WAPiCre *PIK3CA*^{E545K} and *PIK3CA*^{H1047R} mice suggest that different mechanisms underlie tumor initiation by these mutants. To gain a mechanistic insight that might explain these differences, we investigated whether pregnancy accelerates tumor onset in WAPiCre *PIK3CA*^{E545K} as it does in WAPiCre *PIK3CA*^{H1047R} mice (Meyer et al., 2011). Pregnancy accelerated tumor onset in WAPiCre *PIK3CA*^{E545K} mice, reducing latency from 228 ±15 days in nulliparous to 165 ±10 days in parous mice **(Fig. 5-3a)**. Interestingly, pregnancy appeared to accentuate the difference in tumor latency between WAPiCre *PIK3CA*^{E545K} and *PIK3CA*^{H1047R} mice, shown by 32 days difference in nulliparous versus 48 days difference in parous mice **(Fig. 5-3a)** (Meyer et al., 2011). We showed previously that a pregnancy-induced delay in mammary gland involution accounts, at least in part, for accelerated tumor kinetics in parous versus nulliparous *PIK3CA*^{H1047R} mice (Meyer et al., 2011). Thus, we hypothesized here that the longer tumor latency of parous WAPiCre *PIK3CA*^{E545K} compared with parous WAPiCre *PIK3CA*^{H1047R} animals is the result of a less-pronounced involution delay. Comparison of WAPiCre *PIK3CA*^{E545K} and WAPiCre *PIK3CA*^{H1047R} glands 15 days after weaning revealed a dramatic delay in involution compared with control animals **(Fig. 5-3b)**. The relative gland area occupied by epithelial cells was the same in WAPiCre *PIK3CA*^{E545K} and *PIK3CA*^{H1047R} mice and significantly larger than in WAPiCre control glands **(Fig. 5-3c)**. Similarly, there was no difference in the number of apoptotic or proliferating cells in glands expressing either of the *PIK3CA* mutations **(Fig. 5-3d)**. Interestingly, glands from WAPiCre *PIK3CA*^{wt} mice, which did not form tumors, displayed normal involution and numbers of apoptotic and proliferating cells similar to the controls **(Fig. 5-3)**, indicating that delay in involution is caused by mutant *PIK3CA* rather than by overexpression of the transgene. In summary,

Figure 5-3│Pregnancy accelerates *PIK3CA*-evoked tumorigenesis and *PIK3CA* mutants delay mammary gland involution.

a, Kaplan-Meier curves showing tumor onset in parous WAPiCre *PIK3CA*^{E545K} (*n*=16) and nulliparous WAPiCre *PIK3CA*^{E545K} (*n*=6) animals. Parous WAPiCre *PIK3CA*^{E545K} mice developed palpable tumors on average after 165 (±10) days while nulliparous mice developed tumors on average after 228 (±15) days (*PIK3CA*^{E545K}); *P*=0.0023. **b,** Representative images of whole mount (*top panel*), magnification of whole mount (*center panel*) and H&E (*lower panel*) staining of involuting glands from WAPiCre control, WAPiCre *PIK3CA*^{wt}, WAPiCre *PIK3CA*^{E545K}, and WAPiCre *PIK3CA*^{H1047R} mice as indicated. The glands were isolated 15 days after removal of the pups. Scale bar = 1cm (whole mounts). Scale bar = 100 μm (H&E sections). **c,** Bar graph showing relative epithelium to total gland area of involution at day 15 in whole mounts prepared from WAPiCre control (*n*=3), WAPiCre *PIK3CA*^{wt} (*n*=4), WAPiCre *PIK3CA*^{E545K} (*n*=4) and WAPiCre *PIK3CA*^{H1047R} (*n*=2). Shown are means ± s.d.; *P*=0.001 (WAPiCre vs. WAPiCre *PIK3CA*^{E545K}); *P*=0.01 (WAPiCre vs. WAPiCre *PIK3CA*^{H1047R}); *P*=0.46 (WAPiCre *PIK3CA*^{E545K} vs. WAPiCre *PIK3CA*^{H1047R}). **d,** Immunostaining for Ki67 and cleaved caspase-3 of day-15 involuting glands from WAPiCre control, WAPiCre *PIK3CA*^{E545K} and WAPiCre^{H1047R} mice (*upper panel*). Scale bar = 50 μm. Quantification of Ki67- and cleaved caspase-3-positive cells (*lower panel*). Shown are means ± s.e.m. *For Ki67-positive cells: *P*=2.18x10^{-5} (WAPiCre vs. WAPiCre *PIK3CA*^{E545K}), *P*=2.90x10^{-7} (WAPiCre vs. WAPiCre *PIK3CA*^{H1047R}). For cleaved caspase-3-positive cells: *P*=6.55x10^{-3} (WAPiCre vs. WAPiCre *PIK3CA*^{E545K}). NS=not significant.

PIK3CA[E545K] and *PIK3CA*[H1047R] transgene expression caused a dramatic but comparable delay in involution and, therefore, involution does not explain the different tumor kinetics observed in parous versus nulliparous mice expressing these mutations.

5.2.4 WAPiCre *PIK3CA*[E545K] involuting glands show reduced pAkt and increased pStat3

Comparison of lysates from WAPiCre *PIK3CA*[E545K]- and WAPiCre *PIK3CA*[H1047R]-derived mammary glands and tumors showed equal expression of p110α in tumors from both transgenic models **(Fig. 5-4a)**. Despite the enhanced oncogenic potential of the *PIK3CA*[H1047R] mutant, no differences in activation of the PI3K/Akt or the Erk pathways were observed in the tumors **(Fig. 5-4a)**. Similarly, a more detailed analysis of Akt1 and Akt2 isoform-specific phosphorylation revealed no difference between *PIK3CA*[E545K]- and WAPiCre *PIK3CA*[H1047R]-induced signaling **(Fig. 5-4b)**. Conceivably, by the time mammary tumors were established, numerous secondary mutations had resulted in a tumor heterogeneity that compromises the detection of potentially subtle differences in oncogenic signaling induced by either *PIK3CA* mutant. To circumvent this, we investigated molecular signaling events in mutant *PIK3CA*-expressing epithelial cells at an early pre-neoplastic stage. Protein lysates from mammary glands isolated 12 hours after the onset of involution revealed increased phosphorylation of Akt and decreased phosphorylation of the signal transducer and activator of transcription 3 (Stat3) in mutant relative to control glands. These observations were both more pronounced in *PIK3CA*[H1047R] than in *PIK3CA*[E545K] glands (for pAkt $P=2.5 \times 10^{-4}$; for pSTAT3 $P=0.047$) **(Fig. 5-4c)**.
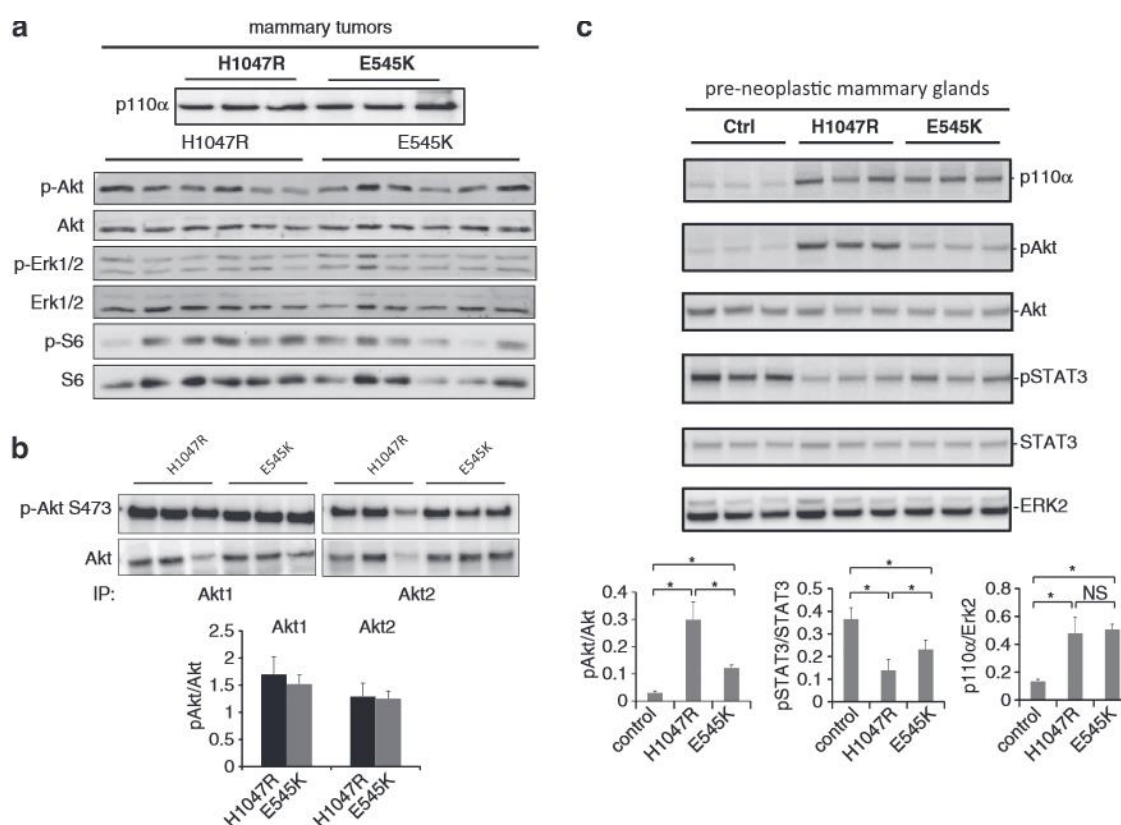
Figure 5-4 │ WAPiCre *PIK3CA*^{E545K} involuting glands show reduced pAkt and increased pSTAT3 compared with WAPiCre *PIK3CA*^{H1047R} at 12 h of involution.

a, Immunoblots of lysates from WAPiCre *PIK3CA*^{E545K} and WAPiCre *PIK3CA*^{H1047R} mammary tumors probed for the indicated proteins. **b,** Lysates of mammary tumors from WAPiCre *PIK3CA*^{E545K} and WAPiCre *PIK3CA*^{H1047R} mice were first immunoprecipitated with antibodies against either Akt1 or Akt2 and then probed for total Akt or S473 pAkt (*upper panel*). Bar graph showing relative levels of S473 pAkt normalized to total Akt in Akt1 or Akt2 immunoprecipitates (*lower panel*). **c,** Lysates of mammary glands isolated 12 h after onset of involution from WAPiCre control, WAPiCre *PIK3CA*^{H1047R} and WAPiCre *PIK3CA*^{E545K} mice (each n=3) probed for p110α, pAkt, Akt, pStat3, Stat3 and Erk2 as loading control (*upper panel*). Bar graphs showing relative amounts of pAkt (normalized to total Akt), pStat3 (normalized to total Stat3) and p110α (normalized to Erk2) in lysates of WAPiCre control, WAPiCre *PIK3CA*^{H1047R} and WAPiCre *PIK3CA*^{E545K} mammary glands (*lower panel*). *WAPiCre vs. WAPiCre *PIK3CA*^{H1047R}: For pAkt P=3.8x10⁻⁵; for pStat3 P=0.002; for p110α P=0.007. WAPiCre vs. WAPiCre *PIK3CA*^{E545K}: For pAkt P=1.8x10⁻⁴; for pStat3 P=0.02; for p110α P=8.8x10⁻⁵.WAPiCre *PIK3CA*^{H1047R} vs. WAPiCre *PIK3CA*^{E545K}: For pAkt P=2.5x10⁻⁴; for pStat3 P=0.047; for p110α P=0.7. NS= not significant.

5.3   Discussion

In summary, we found that expression of $PIK3CA^{E545K}$ in a transgenic mouse model potently induces heterogeneous mammary tumors while expression of wild-type $PIK3CA$ does not. Notably, although $PIK3CA^{E545K}$ evokes tumors with 100% penetrance it is a weaker inducer of mammary tumors than $PIK3CA^{H1047R}$ in two independent mouse models in which mutant $PIK3CA$ is either driven by the WAP or CAGs promoter. This may explain the lower frequency of helical versus kinase domain mutations in human breast cancer (Bader et al., 2005; Koren and Bentires-Alj, 2013).

It was proposed that $PIK3CA^{E545K}$ and $PIK3CA^{H1047R}$ mutants increase PI3K activity by distinct mechanisms. Biochemical assays suggest that $PIK3CA^{E545K}$ mutants mimic the activated wildtype p110α. $PIK3CA^{E545K}$ shows high basal activity and lipid binding and cannot be further activated by RTK phosphopeptides. In contrast, mutations in the catalytic domain (e.g., $PIK3CA^{H1047R}$) increase basal kinase activities and lipid binding, which can be further enhanced by RTK phosphopeptides (Hon et al., 2012). Moreover, mutations in the helical domain (e.g., $PIK3CA^{E545K}$) were shown to disrupt an inhibitory charge-charge interaction with the p85 N-terminal SH2 domain (Miled et al., 2007), while mutations in the catalytic domain were shown to increase the affinity for substrate-containing membranes (Mandelker et al., 2009). Another study showed that $PIK3CA^{E545K}$ mutants are independent of binding to the adaptor molecule p85 but require interaction with Ras-GTP. $PIK3CA^{H1047R}$ mutants are highly dependent on p85 for its oncogenic capacity but are independent of Ras-GTP (Zhao and Vogt, 2008b). Moreover, it was shown that in contrast to kinase mutants, the helical domain mutants directly associate with IRS1 without growth factor stimulation and without the interaction with p85 (Hao et al., 2013).

Here, we found differences in Akt and Stat3 activation in pre-neoplastic mammary glands from $PIK3CA^{E545K}$ and $PIK3CA^{H1047R}$ transgenic mice. This observation might explain the longer tumor latency observed in WAPiCre $PIK3CA^{E545K}$ compared with WAPiCre

PIK3CA^H1047R mice. Further investigation to understand the exact consequences of these mutations on downstream signaling and tumorigenesis is of profound interest.

5.4 Materials and methods

Transgenic Mice. We constructed a vector with a transcriptional STOP sequence flanked by *loxP* sites upstream of the 5'-terminally HA-tagged human *PIK3CA* cDNA (Addgene) and an *IRES2-EGFP* reporter element (pIRES2-EGFP vector, Clontech). The resulting construct was introduced into the modified Rosa26 locus of Balb/c mouse embryonic stem (ES) cells by RMCE as described earlier (Meyer et al., 2011). Chimeric mice were backcrossed to Balb/c mice and transgenic mice identified by genotyping (Meyer et al., 2011).

Immunoblotting. Protein lysates were extracted from inguinal mammary glands or tumors using LB buffer (50mM Tris-HCl pH8, 150mM NaCl, 1% NP-40) supplemented with 0.5mM sodium orthovanadate. Anti-p110α, anti-pAkt (Ser473), anti-Akt, anti-pErk1/2 (Thr202/Tyr204), anti-Erk1/2, anti-pS6 (Ser235/236), anti-S6, anti-Akt1, anti-Akt2, anti-pStat3 (Tyr705) and anti-Stat3 antibodies were purchased from Cell Signaling Technology.

Immunohistochemistry. The following antibodies were used: K14 (Thermo Scientific, RB-9020, 1:100), K8/18 (Fitzgerald, #GP11, 1:200), ER (Santa Cruz, SC-542, 1:1000), α-SMA (Thermo Scientific, RB-9010, 1:500), cleaved caspase-3 (Cell Signaling, #9661, 1:100) Ki-67 (Thermo Scientific, RB-9106, 1:50).

Southern Blotting. Genomic DNA from mouse tails was digested with 8 U of AvrII enzyme (New England BioLabs (NEB)) and separated on a 1% agarose gel. A DIG-labeled DNA probe targeting the neomycin resistance cassette was amplified using the PCR DIG Probe Synthesis Kit (Roche) and the primers 5'-ATGGGATCGGCCATTGAACAAGAT-3' and 5'-CGGCCATTTTCCACCATGATAT-3'.

6 | Results Part II

Effects of *PIK3CA*[H1047R] expression on mammary cell fate and cancer

LETTER
doi:10.1038/nature14669

PIK3CA[H1047R] induces multipotency and multi-lineage mammary tumours

Shany Koren[1], Linsey Reavie[1], Joana Pinto Couto[1], Duvini De Silva[1], Michael B. Stadler[1,2], Tim Roloff[1], Adrian Britschgi[1], Tobias Eichlisberger[1], Hubertus Kohler[1], Olulanu Aina[3], Robert D. Cardiff[3] & Mohamed Bentires-Alj[1]

[1]Friedrich Miescher Institute for Biomedical Research (FMI), 4058 Basel, Switzerland. [2]Swiss Institute of Bioinformatics, 4058 Basel, Switzerland. [3]Department of Pathology, Center for Comparative Medicine, University of California Davis, Davis, California 95616, USA.

The full published article can be found in **appendix 10.4**.

6.1   Summary

The adult mouse mammary epithelium contains self-sustained cell lineages that form the inner luminal and outer basal cell layers, with stem and progenitor cells contributing to its proliferative and regenerative potential (de Visser et al., 2012; Tao et al., 2014; van Amerongen et al., 2012; Van Keymeulen et al., 2011). A key issue in breast cancer biology is the effect of genomic lesions in specific mammary cell lineages on tumor heterogeneity and progression. The impact of transforming events on fate conversion in cancer cells-of-origins and thus their contribution to tumor heterogeneity remains largely elusive. Using *in situ* genetic lineage tracing and limiting dilution transplantation, we have unraveled the potential of *PIK3CA*$^{\text{H1047R}}$, one of the most frequent mutations occurring in human breast cancer (Cancer Genome Atlas, 2012), to induce multipotency during tumorigenesis in the mammary gland. We show that expression of *PIK3CA*$^{\text{H1047R}}$ in lineage-committed basal Lgr5-positive and luminal keratin 8-positive cells of the adult mouse mammary gland evokes cell dedifferentiation into a multipotent stem-like state, suggesting this to be a mechanism involved in the formation of heterogeneous, multi-lineage mammary tumors. Moreover, we show that the tumor cell-of-origin influences the frequency of malignant mammary tumors. Our results define a key effect of *PIK3CA*$^{\text{H1047R}}$ on mammary cell fate in the pre-neoplastic mammary gland and show that the cell-of-origin of *PIK3CA*$^{\text{H1047R}}$ tumors dictates their malignancy, thus revealing a mechanism underlying tumor heterogeneity and aggressiveness.

6.2 Results

6.2.1 Mutant *PIK3CA* induces mammary cell plasticity

To address the effects of mutant $PIK3CA^{H1047R}$ on basal or luminal lineage-restricted cells, we performed lineage tracing in adult Lgr5-creERT2/Tomato-reporter and K8-creERT2/Tomato-reporter mice with or without $PIK3CA^{H1047R}$ (Meyer et al., 2011) **(Extended Data Fig. 6-1a)**. Moreover, we used Lgr5- and K8-creERT2/$PIK3CA^{H1047R}$ or $PIK3CA^{WT}$ animals (Meyer et al., 2011; Meyer et al., 2013) for tracing of the GFP-reporter, and Lgr5-creERT2 and K8-creERT2 as controls **(Extended Data Fig. 6-1b)**. As previously reported (Plaks et al., 2013; Van Keymeulen et al., 2011), we found Lgr5-activity only in a subset of basal cells in the nipple area of the mammary gland **(Extended Data Fig. 6-2a-e)**. We further assessed the effects of $PIK3CA^{H1047R}$ expression on the distribution of mammary subpopulations by fluorescence activated cell sorting (FACS) on isolated mammary epithelial cells labeled with CD24 and Sca1, which were shown to enrich for luminal ($CD24^{high}Sca1^{negative}$, $CD24^{high}Sca1^{positive}$) and basal ($CD24^{low}Sca1^{negative}$) cells (Sleeman et al., 2006; Sleeman et al., 2007) **(Extended Data Fig. 6-3)**. Four days lineage-tracing confirmed Tomato-labeling of basal cells in Lgr5-creERT2/Tomato and Lgr5-creERT2/$PIK3CA^{H1047R}$/Tomato and of luminal cells in K8-creERT2/Tomato and K8-creERT2/$PIK3CA^{H1047R}$/Tomato animals **(Extended Data Fig. 6-2b,f and 6-4a,b)**. We further performed 4-, 8- and 13-week lineage-tracing. The distribution of Tomato-labeled subsets in Lgr5-creERT2/Tomato glands did not change with time where the progeny of Lgr5-positive cells was mostly of basal origin **(Fig. 6-1a,b and Extended Data Fig. 6-2g,h)**. In K8-creERT2/Tomato animals, labeling was restricted to the luminal subset, marking mostly mature $CD24^{high}Sca1^{positive}$ in 4 weeks and mostly $CD24^{high}Sca1^{negative}$ luminal progenitors in 13 weeks tracing, indicating the targeting of a long-term unipotent luminal subset **(Fig. 6-1a,c and Extended Data Fig. 6-4c,d)**. Further
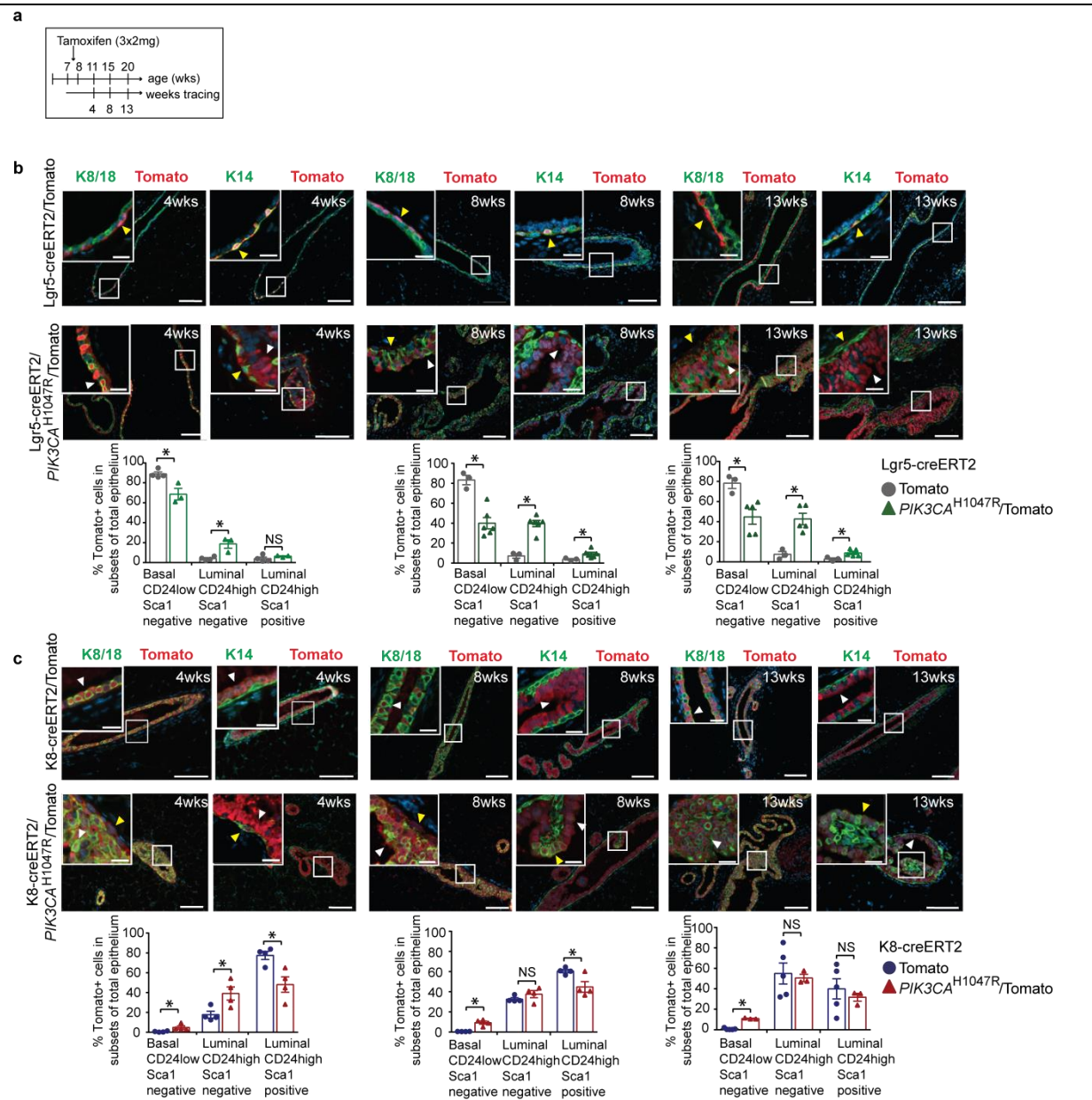
Figure 6-1 │ Mutant *PIK3CA* induces mammary cell plasticity

a, Timeline for lineage-tracing studies. **b,c,** Representative images of 4-, 8- and 13-week tracing and FACS quantification of Tomato-positive epithelial basal (CD24^{low}Sca1^{negative}) and luminal (CD24^{high}Sca1$^{negative/positive}$) subsets from Lgr5-creERT2/Tomato (**b,** immunofluorescence: left *n*=3, middle *n*=11, right *n*=6 mice; FACS: left *n*=4 technical replicates (each 1-3 pooled mice), middle *n*=3 technical replicates (each 1-2 pooled mice), right *n*=3 technical replicates (each 1 mouse)), Lgr5-creERT2/*PIK3CA*^{H1047R}/Tomato (**b,** immunofluorescence left *n*=3, middle *n*=10, right *n*=4 mice; FACS: left *n*=3 technical replicates (each 1-2 pooled mice), middle *n*=6 technical replicates (each 1 mouse), right *n*=5 technical replicates (each 1-2 pooled mice)), K8-creERT2/Tomato (**c,** immunofluorescence left *n*=4, middle *n*=8, right *n*=5 mice; FACS: left *n*=4 technical replicates (each 1-3 pooled mice), middle *n*=4 technical replicates (each 1-2 pooled mice), right *n*=5 technical replicates (each 1 mouse)) and K8-creERT2/*PIK3CA*^{H1047R}/Tomato animals (**c,** immunofluorescence left *n*=4, middle *n*=4, right *n*=3 mice; FACS: left *n*=4 technical replicates (each 1-2 pooled mice), middle *n*=4, right *n*=3 technical replicates (each 1 mouse)). White arrowheads indicate luminal and

53

yellow arrowheads basal Tomato-labeled cells. Scale bar 100 μm; magnification 20 μm. Bar graphs means ±s.e.m.; two-sided unpaired student`s t-test; *P<0.05; NS= not significant.

tracing of 20-28 weeks confirmed lineage-restriction in both models (**Unpublished Data Fig. 6-1**). In contrast, expression of *PIK3CA*^H1047R in Lgr5-creERT2/*PIK3CA*^H1047R/Tomato and K8-creERT2/*PIK3CA*^H1047R/Tomato mice resulted in labeling of both the luminal and the basal compartments (**Fig. 6-1b,c Extended Data Fig. 6-2g,h and 6-4c,d**). *PIK3CA*^H1047R-evoked multi-lineage labeling was not observed at 4-7 days post-induction (**Extended Data Fig. 6-2f,i and 6-4b**). Since *PIK3CA*^H1047R/WT constructs contain an IRES-GFP, we also used GFP as a readout of transgene expression. At 4 days post-Tamoxifen induction, K8-creERT2/*PIK3CA*^WT and K8-creERT2/*PIK3CA*^H1047R glands expressed similar levels of GFP indicating similar Cre recombination efficiency. In both models, 4- and 8-11-week lineage-tracing revealed an increase in GFP-labeled basal and luminal subsets in *PIK3CA*^H1047R compared to *PIK3CA*^WT animals that showed lineage-restricted GFP-labeling, consistent with higher PI3K-pathway activation (**Extended Data Fig. 6-5 and 6-6**). These results suggest that expression of *PIK3CA*^H1047R in basal- or luminal-restricted mammary cells triggers lineage plasticity and cell expansion.

6.2.2 Activation of *PIK3CA*^H1047R leads to expression of basal and luminal lineage-genes

Next, we assessed the effects of *PIK3CA*^H1047R on global gene expression in the pre-neoplastic gland. Microarray and QRT-PCR analyses were performed on mammary epithelial cell subpopulations from Lgr5-creERT2 control versus Lgr5-creERT2/*PIK3CA*^H1047R and K8-creERT2 control versus K8-creERT2/*PIK3CA*^H1047R animals. Genes expressed differentially between subpopulations of control and mutant mice were compared to subpopulation signatures from Meier-Abt *et al.* (Meier-Abt et al., 2013) and Lim *et al.* (Lim et al., 2010).

We found enrichment of luminal progenitor signature-genes in the basal Lgr5-creERT2/*PIK3CA*[H1047R] subset and in the newly formed basal K8-creERT2/*PIK3CA*[H1047R] subset. Enrichment of myoepithelial signature-genes was found in the newly formed luminal Lgr5-creERT2/*PIK3CA*[H1047R] and in the K8-creERT2/*PIK3CA*[H1047R] luminal subsets (**Fig. 6-2a,b and Extended Data Fig. 6-7**). Mammary cells co-expressing basal and luminal markers in neoplastic areas *in vivo* (**Fig. 6-2c**) and in freshly sorted mammary cells (**Unpublished Data Fig. 6-2**) confirmed that *PIK3CA*[H1047R] induces cell plasticity.

Figure 6-2 │ Activation of *PIK3CA*[H1047R] leads to expression of basal and luminal lineage-genes
Plots indicating enrichment of gene expression from FACS-sorted (**a**) *Lgr5-creERT2/PIK3CA*[H1047R] versus Lgr5-creERT2 control and (**b**) K8-creERT2/*PIK3CA*[H1047R] versus K8-creERT2 control subsets in signatures of mammary subpopulations from Lim *et al.*(Lim et al., 2010) and Meier-Abt *et al.*(Meier-Abt et al., 2013). Microarray was performed 4 weeks post Tamoxifen on basal CD24[low]Sca1[negative] and luminal CD24[high]Sca1[negative/positive] subsets of pooled mammary glands of 2-3 estrus-synchronized animals from three independent sortings. **c,** Representative images of immunostaining for basal (K14 red, K5 white) and luminal (K8/18 green) markers on mammary glands 4 weeks post Tamoxifen (*n*=3 mice). White arrowheads indicate double-positive cells. Scale bar 100 μm; magnification 20 μm.

6.2.3 Expression of *PIK3CA*[H1047R] evokes multipotent stem-like cells

We addressed how lineage plasticity is evoked by *PIK3CA*[H1047R] in functional assays. Limiting dilution transplantation revealed an increase in mammary-repopulating capacity of basal cells from Lgr5-creERT2/*PIK3CA*[H1047R] and K8-creERT2/*PIK3CA*[H1047R] compared with the respective controls **(Fig. 6-3)**. The GFP-negative basal CD24[low]Sca1[negative] population from K8-creERT2 mice comprises the bulk of basal cells (myoepithelial and mammary-repopulating cells) (Sleeman et al., 2006; Sleeman et al., 2007), explaining the lack of outgrowths at the number of cells transplanted. As previously reported (Sleeman et al., 2006), repopulating capacity can be found when 1,000 and more K8-creERT2 control basal cells are transplanted **(Unpublished Data Fig. 6-3)**. *PIK3CA*[H1047R]-expressing luminal cells also had repopulating capacity **(Extended Data Fig. 6-8a-c).** All outgrowths expressed luminal and basal markers. In colony formation assays, *PIK3CA*[H1047R] increased the percentage of double-positive (K14/K8/18) colonies derived from the newly formed luminal (Lgr5-creERT2/*PIK3CA*[H1047R]) and basal cells (K8-creERT2/*PIK3CA*[H1047R]). *PIK3CA*[H1047R]-expressing subsets overcame lineage-restriction, giving rise to both lineages albeit at low frequencies. Moreover, basal cells from Lgr5- and K8-creERT2/*PIK3CA*[H1047R] animals showed increased colony formation capacity **(Extended Data Fig. 6-8d-f)**. In mammosphere cultures, *PIK3CA*[H1047R] increased sphere-forming capacity of luminal cells. While control cells formed spheres with a hollow lumen after passaging, *PIK3CA*[H1047R]-expressing cells formed filled spheres, indicating the accumulation of less differentiated cells **(Extended Data Fig. 6-8g,h)**. Altogether, the data suggest that *PIK3CA*[H1047R] evokes cell dedifferentiation to a multipotent stem-like state, from which cells further differentiate into both cell lineages.

a Lgr5-creERT2 outgrowths

	cell #	Control	PIK3CA^{H1047R}
Basal CD24 low Sca1 negative	50	5/17 (29%)	10/19 (52%)
	100	4/15 (26%)	9/15 (60%)
	200	5/13 (38%)	9/13 (69%)
MRU frequency (95% CI)		1 in 289 (1/169 - 1/496)	1 in 111 (1/74 - 1/166)

P=0.003

K8-creERT2 outgrowths

	cell #	Control	PIK3CA^{H1047R}
Basal CD24 low Sca1 negative	100	0/6 (0%)	4/10 (40%)
	200	0/8 (0%)	7/11 (63%)
	400	0/5 (0%)	2/4 (50%)
MRU frequency (95% CI)		n.a.	1 in 253 (1/142-1/452)

P=4 x 10^{-6}

Figure 6-3│ Expression of *PIK3CA*^{H1047R} evokes multipotent stem-like cells
a, Number of outgrowths in cleared-fat pad transplantation of GFP-positive (Lgr5-positive) control and *PIK3CA*^{H1047R}-expressing basal subsets (upper panel). Outgrowths of GFP-negative control and GFP-positive *PIK3CA*^{H1047R}-expressing basal subsets (lower panel). The GFP-negative control basal subset comprises mammary repopulating cells and mostly myoepithelial cells. **b,** Representative immuno-stained sections (*n*=3 mice) and carmine-stained whole mounts of outgrowths from cleared-fat pad transplantation (100 cells condition). Scale bar 50 μm and 500 μm. Pooled data from 3 independent experiments (a-b).

6.2.4 The frequency of malignant tumor lesions is dictated by the cell-of-origin

Expression of *PIK3CA*^{H1047R} in Lgr5- and K8-positive cells induced mammary tumors on average after 108 and 78 days, respectively. Control and *PIK3CA*^{WT} animals developed no tumors (**Fig. 6-4a**). FACS analysis of GFP-positive tumor cells revealed a similar distribution of cancer subpopulations, with an accumulation of the CD24^{high}Sca1^{negative} subset. *PIK3CA*^{H1047R}-evoked mammary tumors expressed basal and luminal markers (**Extended Data Fig. 6-9a-d**). Additionally, cells double positive for basal and luminal markers were found (**Fig. 6-4b**). These results suggest that *PIK3CA*^{H1047R}-evoked cell plasticity results in multi-lineage tumors and that expression of basal and luminal markers is not an indicator of the origin of mammary cancers. It was proposed that basal-like mammary tumors may originate from luminal cells (Lim et al., 2009; Molyneux et al., 2010). Therefore, any inference of the cell-of-origin from the differentiation state of the tumor can be misleading.

Histological analysis showed heterogeneous phenotypes and differences in malignancy between both models. Lgr5-creERT2/*PIK3CA*[H1047R] formed unique benign multi-nodular rosette-type adenomyoepitheliomas, aggressive adenosquamous carcinomas with pilosebaceous differentiation and carcinosarcomas. K8-creERT2/*PIK3CA*[H1047R] led mainly to aggressive adenosquamous carcinoma, carcinosarcomas that infiltrated the surrounding tissue, adenocarcinomas and benign adenomyoepitheliomas. These results show that *PIK3CA*[H1047R] mainly evokes benign tumors (74%) with high intra-tumor heterogeneity when expressed in Lgr5-positive cells, in contrast to mostly aggressive mammary tumors (62%) with a distinctive infiltrative densely fibrotic phenotype when the cell-of-origin is K8-positive **(Fig. 6-4c and Extended Data Fig. 6-9e-f)**.

Microarray analysis, principle component analysis and hierarchical-clustering of tumors revealed a single cluster of Lgr5-creERT2/*PIK3CA*[H1047R] and 3 clusters of K8-creERT2/*PIK3CA*[H1047R] tumors. We found no correlation between tumor phenotype and clustering likely due to intra-tumor heterogeneity. We compared tumor expression profiles to different human breast cancer subtypes (Cancer Genome Atlas, 2012). The majority (7/10) of the K8-creERT2/*PIK3CA*[H1047R] tumors correlated best with the malignant Basal-like, HER2-enriched and Luminal-B profiles, whereas 3/10 clustered with benign Luminal-A and Normal-like breast cancers. Lgr5-creERT2/*PIK3CA*[H1047R] tumors resembled benign (2/10) but also malignant subtypes (3/10). However, 5/10 showed no similarity to a single subtype but were equidistant from all, representing the high intra-tumor heterogeneity in this model **(Fig. 6-4d and Extended Data Fig. 6-10)**. The fact that K8-creERT2/*PIK3CA*[H1047R] tumors but not Lgr5-creERT2/*PIK3CA*[H1047R] tumors clustered mostly with malignant breast cancers that have a poor prognosis is consistent with the histopathological results and suggests that, in the presence of the same initiating oncogenic mutation, the cell-of-origin dictates the frequency of aggressive tumors.
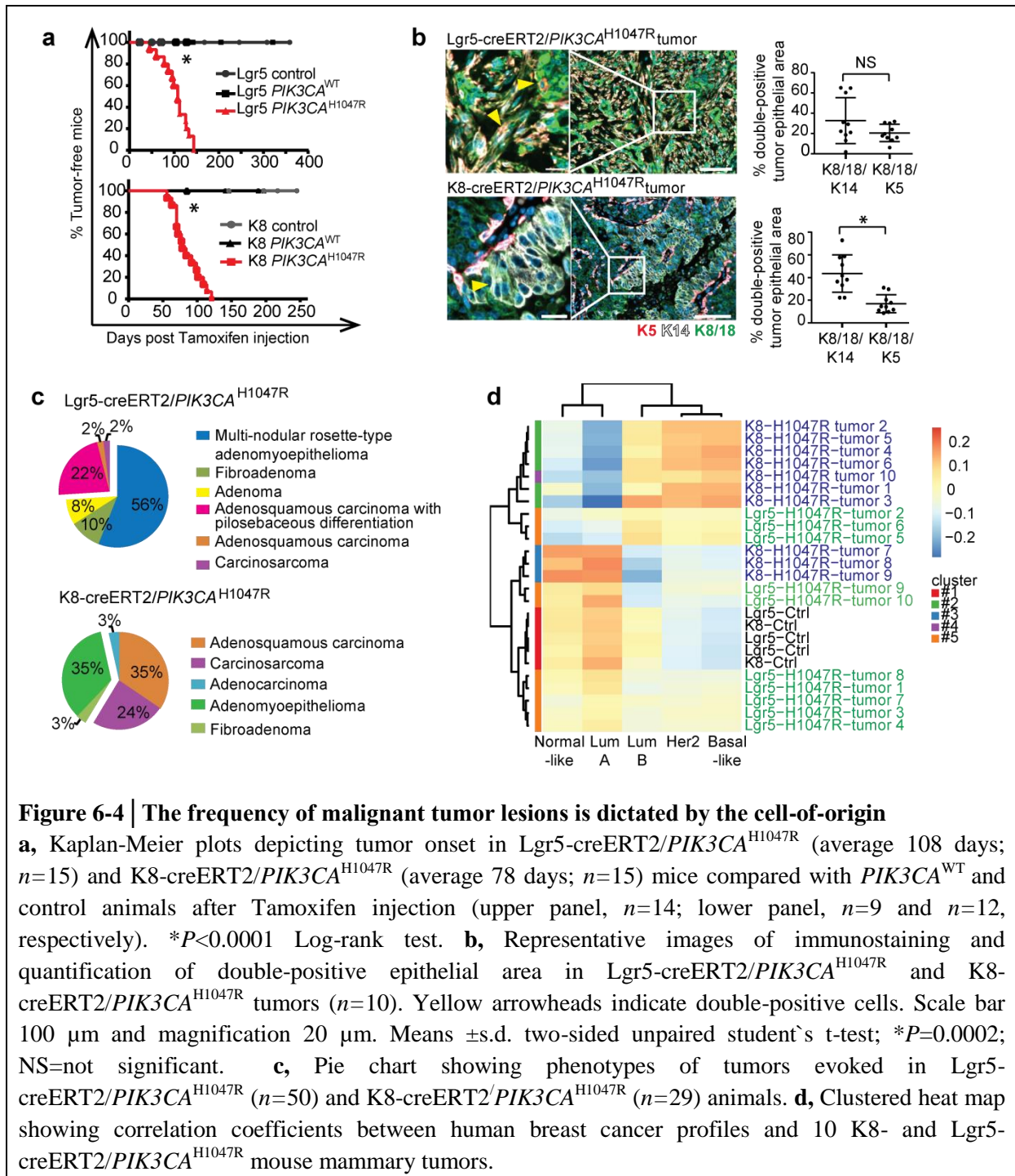
Figure 6-4 │ The frequency of malignant tumor lesions is dictated by the cell-of-origin
a, Kaplan-Meier plots depicting tumor onset in Lgr5-creERT2/*PIK3CA*[H1047R] (average 108 days; *n*=15) and K8-creERT2/*PIK3CA*[H1047R] (average 78 days; *n*=15) mice compared with *PIK3CA*[WT] and control animals after Tamoxifen injection (upper panel, *n*=14; lower panel, *n*=9 and *n*=12, respectively). *P<0.0001 Log-rank test. **b,** Representative images of immunostaining and quantification of double-positive epithelial area in Lgr5-creERT2/*PIK3CA*[H1047R] and K8-creERT2/*PIK3CA*[H1047R] tumors (*n*=10). Yellow arrowheads indicate double-positive cells. Scale bar 100 μm and magnification 20 μm. Means ±s.d. two-sided unpaired student`s t-test; *P=0.0002; NS=not significant. **c,** Pie chart showing phenotypes of tumors evoked in Lgr5-creERT2/*PIK3CA*[H1047R] (*n*=50) and K8-creERT2/*PIK3CA*[H1047R] (*n*=29) animals. **d,** Clustered heat map showing correlation coefficients between human breast cancer profiles and 10 K8- and Lgr5-creERT2/*PIK3CA*[H1047R] mouse mammary tumors.

6.3 Extended Data Figures and Unpublished Data Figures
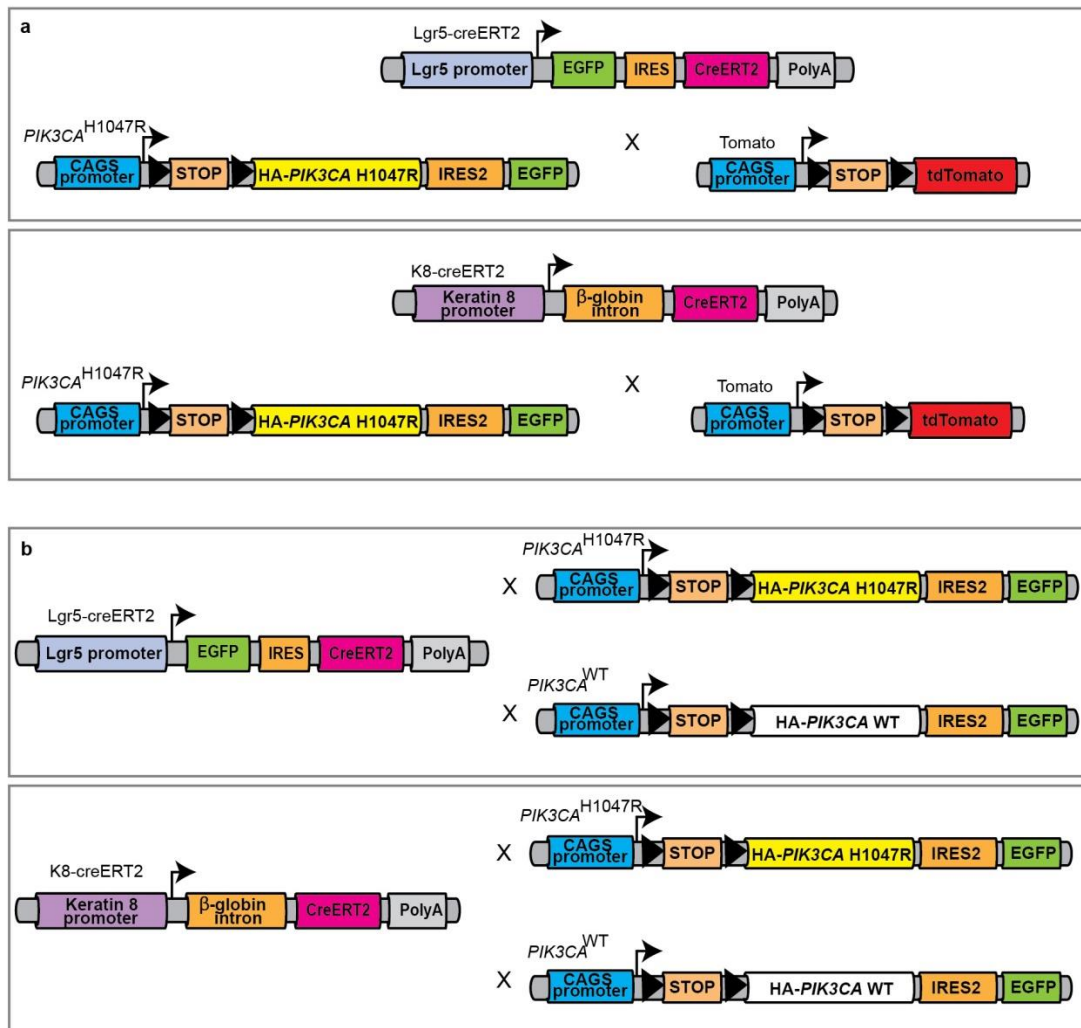
Extended Data Figure 6-1│Scheme depicting mouse lines generated for lineage tracing studies
a, Lgr5-creERT2 (Kinzel et al., 2014) or K8-creERT2 (Van Keymeulen et al., 2011) animals were crossed to transgenic lox-STOP-lox *PIK3CA*^H1047R (Meyer et al., 2011) and/or Tomato-reporter mice, generating Lgr5-creERT2/Tomato, K8-creERT2/Tomato, Lgr5-creERT2/*PIK3CA*^H1047R/Tomato and K8-creERT2/*PIK3CA*^H1047R/Tomato animals for lineage-tracing studies. Lgr5-creERT2/Tomato and K8-creERT2/Tomato animals were used as controls. **b,** Lgr5-creERT2 (Kinzel et al., 2014) and K8-creERT2 (Van Keymeulen et al., 2011) animals were crossed to lox-STOP-lox *PIK3CA*^H1047R (Meyer et al., 2011) or *PIK3CA*^WT (Meyer et al., 2013) animals. Lgr5-creERT2 and K8-creERT2 animals were used as controls. Tamoxifen injection induces *PIK3CA*^H1047R, *PIK3CA*^WT, and/or Tomato expression.
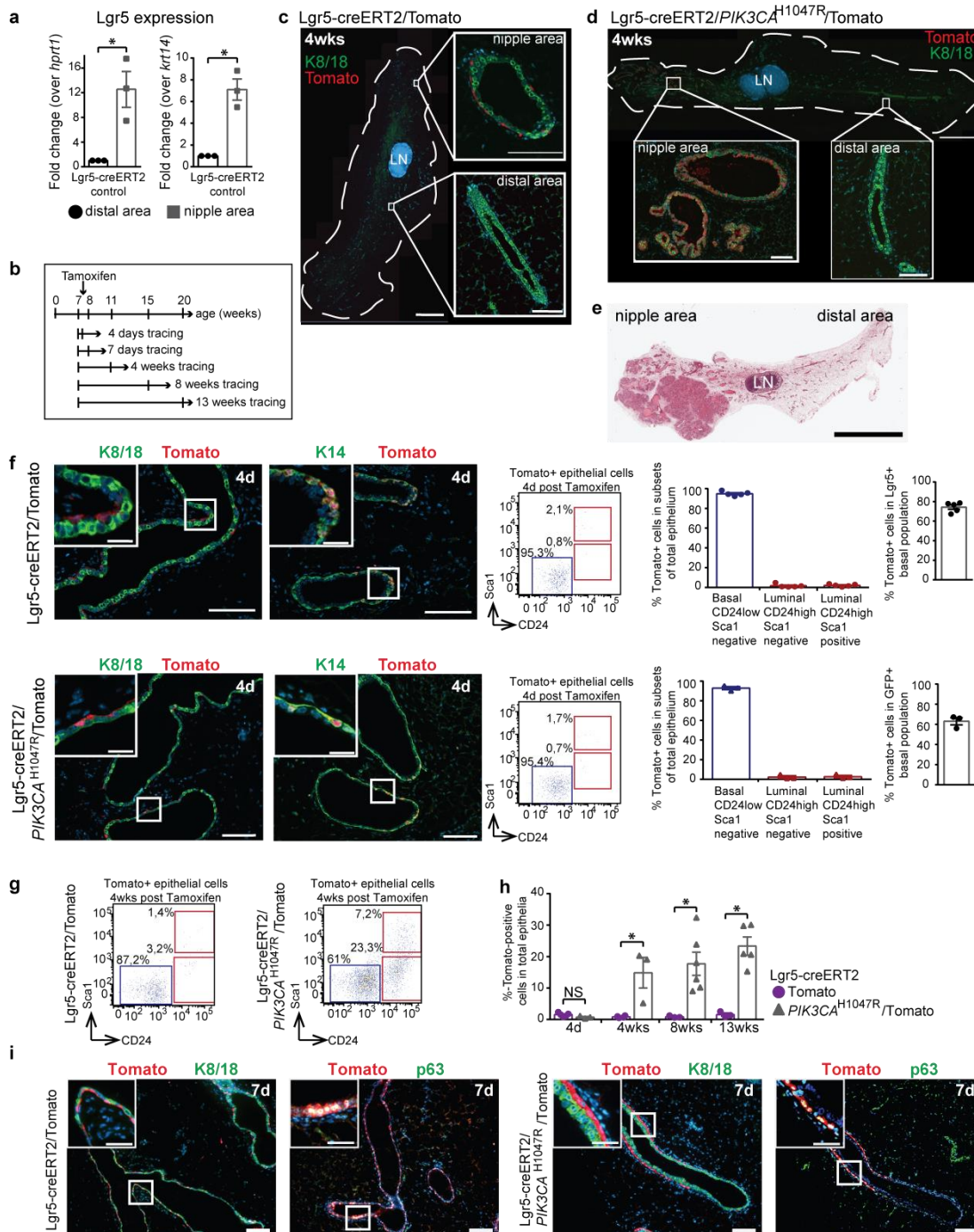
Extended Data Figure 6-2│Lgr5-creERT2/Tomato and *PIK3CA*[H1047R]/Tomato labeling in the mammary nipple area

a, Lgr5-expression in the nipple and distal area of Lgr5-creERT2 glands (*n*=3 mice). **b,** Tracing scheme. **c,d,** Representative images of mammary glands after 4 weeks tracing (*n*=3 mice for each genotype). Scale bar 2 mm; magnification 100 μm. **e,** Representative H&E-staining of a Lgr5-creERT2/*PIK3CA*[H1047R] mammary gland with a tumor. Scale bar 500 μm. LN: lymph node. **f,** Representative images, FACS plots and quantification of 4 days tracing (24 h after the last Tamoxifen injection) (upper panel: immunofluorescence *n*=3 mice, FACS: *n*=5 technical replicates (each 1-2 pooled mice); lower panel: immunofluorescence *n*=3 mice, FACS: *n*=3 technical replicates (each 1 mouse)). Scale bar 100 μm; magnification 20 μm. **g,** Representative FACS plots of 4-week tracing. **h,**

61

Percentage of total Tomato-positive cells in the tracing experiments (Lgr5-creERT2/Tomato: 4 days *n=5*, 4 weeks *n=4*, 8 and 13 weeks *n=3* technical replicates (each 1-2 pooled mice); Lgr5-creERT2/*PIK3CA*[H1047R]/Tomato: 4 days *n=3*, 4 weeks *n=3*, 8 weeks *n=6* and 13 weeks *n=5* technical replicates (each 1-2 pooled mice)). **i,** Representative images of 7 days tracing (left *n=4* mice; right *n=2* mice). Scale bar 100 μm, magnification 50 μm, Bar graphs means ±s.e.m. two-sided unpaired student`s t-test, *P<0.05. NS=not significant.

Extended Data Figure 6-3 │ Gating scheme for FACS experiments
Representative FACS plots of **a,** K8-creERT2/Tomato, **b,** K8-creERT2/*PIK3CA*[H1047R]/Tomato, **c,** Lgr5-creERT2/*PIK3CA*[WT], and **d,** Lgr5-creERT2/*PIK3CA*[H1047R] animals 4 weeks after Tamoxifen injection. The gating strategy shown illustrates the elimination of doublets, dead cells (DAPI[high]), and white blood cells (CD45[positive]) and the sorting of Tomato- or GFP-positive mammary epithelial subsets (basal CD24[low]Sca1[negative], luminal CD24[high]Sca1[negative]; luminal CD24[high]Sca1[positive]).

Extended Data Figure 6-4│K8-creERT2/Tomato and *PIK3CA*$^{\text{H1047R}}$/Tomato labeling in the mammary gland

a, Scheme depicting time line of tracing experiments. **b,** Representative images and FACS quantifications of K8-creERT2/Tomato and K8-creERT2/*PIK3CA*$^{\text{H1047R}}$/Tomato mammary glands 4 days post Tamoxifen (24 h after the last Tamoxifen injection) (upper panel: immunofluorescence $n=5$ mice, FACS: $n=7$ technical replicates (each 1-2 pooled mice); lower panel: immunofluorescence $n=3$ mice, FACS: $n=3$ technical replicates (each 1 mouse)). Scale bar 100 μm; magnifications 20 μm. **c,** Representative FACS plots of 4-week Tomato tracing. **d,** Percentage of total Tomato-positive cells in mammary glands. (K8-creERT2/Tomato: 4 days $n=7$, 4 and 8 weeks $n=4$, 13 weeks $n=5$ technical replicates (each 1-3 pooled mice); K8-creERT2/*PIK3CA*$^{\text{H1047R}}$/Tomato: 4 days $n=3$, 4 and 8 weeks $n=4$ and 13 weeks $n=3$ technical replicates (each 1-2 pooled mice)). Bar graphs means ±s.e.m. *$P<0.05$; two-sided unpaired student`s t-test.

Extended Data Figure 6-5 │ Tracing of GFP-positive mammary subsets
a, Percentage of GFP-labeled cells in K8-creERT2/*PIK3CA*[H1047R] versus K8-creERT2/*PIK3CA*[WT]
animals 4 days post Tamoxifen (24 h after the last Tamoxifen injection) (*n*=3 technical replicates, 2
mice per genotype). **b,d,** Representative FACS plots and percentage of GFP-positive cells in
mammary gland subsets and total mammary epithelial cells 4 and 8-11 weeks post Tamoxifen. **c,e,** Bar

graphs showing total numbers of GFP-positive cells and numbers of GFP-positive cells in basal (CD24^{low}Sca1^{negative}) and luminal (CD24^{high}Sca1^{negative}; CD24^{high}Sca1^{positive}) subsets of Lgr5-creERT2/*PIK3CA*^{H1047R} **(c)** and K8-creERT2/*PIK3CA*^{H1047R} **(e)** mammary epithelial cells (b,c: 4 weeks: non-induced control *n=3*, control *n=9*, *PIK3CA*^{WT} *n=3*, *PIK3CA*^{H1047R} *n=9* sortings with each 1-4 pooled mice, 8-11 weeks: non-induced control *n=3*, control *n=3*, *PIK3CA*^{WT} *n=3*, *PIK3CA*^{H1047R} *n=4* sortings with each 1-4 pooled mice; d,e: 4 weeks: *PIK3CA*^{WT} and *PIK3CA*^{H1047R} *n=3* sortings with each 1-5 pooled mice, 8-11 weeks: *PIK3CA*^{WT} *n=4*, *PIK3CA*^{H1047R} *n=5-6* sortings with each 1-4 pooled mice). Bar graphs, means ±s.e.m; two-sided unpaired student`s t-test, *P<0.05; NS=not significant.

Extended Data Figure 6-6 │ Expression of *PIK3CA*^{H1047R} induces Akt phosphorylation
Immunoblot and quantification of lysates from K8-creERT2 control, *PIK3CA*^{WT} and *PIK3CA*^{H1047R} mammary glands 4 weeks post Tamoxifen for p110α, pAkt, Akt, pan Keratin and Erk2 (loading control). Protein levels were normalized to pan Keratin for normalization of epithelial content. Bar graphs depict fold change over control lysate. Bar graph means ±s.d. two-sided unpaired student`s t-test, *P<0.006; NS= not significant.
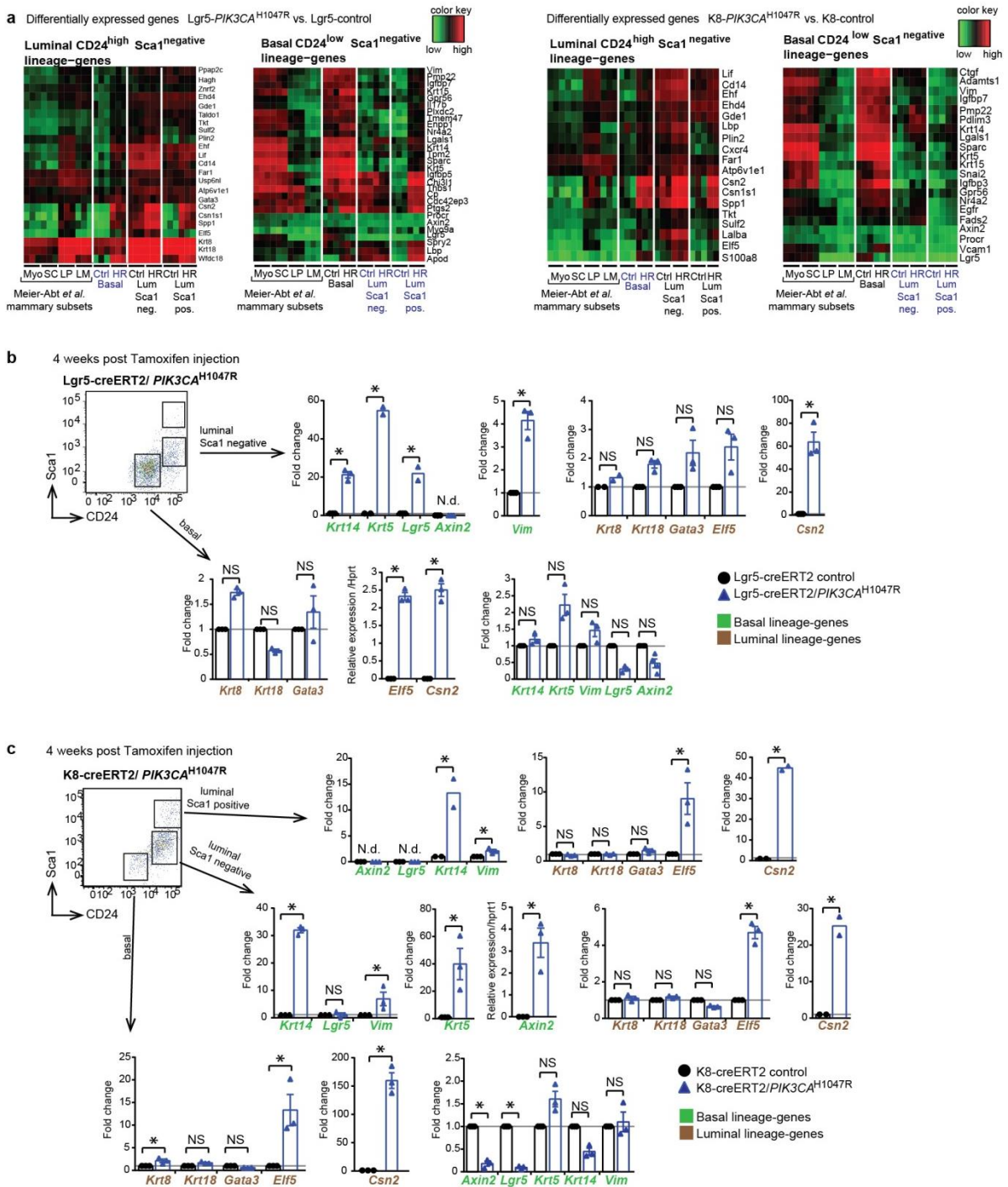
Extended Data Figure 6-7│Expression of basal and luminal lineage-genes in *PIK3CA*[H1047R] subsets

a, Expression heat maps of selected luminal and basal genes (left: Lgr5-creERT2/*PIK3CA*[H1047R] versus control; right K8-creERT2/*PIK3CA*[H1047R] versus control). Myo: myoepithelial; SC: stem cell-enriched; LP: luminal progenitors; LM: mature luminal cells. **b,c** Expression profiles of basal and luminal lineage-genes in mammary subsets of (**b**) Lgr5-creERT2/*PIK3CA*[H1047R] compared to Lgr5-creERT2 control and (**c**) K8-creERT2/*PIK3CA*[H1047R] compared to K8-creERT2 control animals. The QRT-PCR results are representative of 2-3 experiments of 4 pooled animals of each genotype. Bar graphs means ±s.e.m. two-sided unpaired student`s t-test, *P<0.05,. NS=not significant. N.d.=not detected.
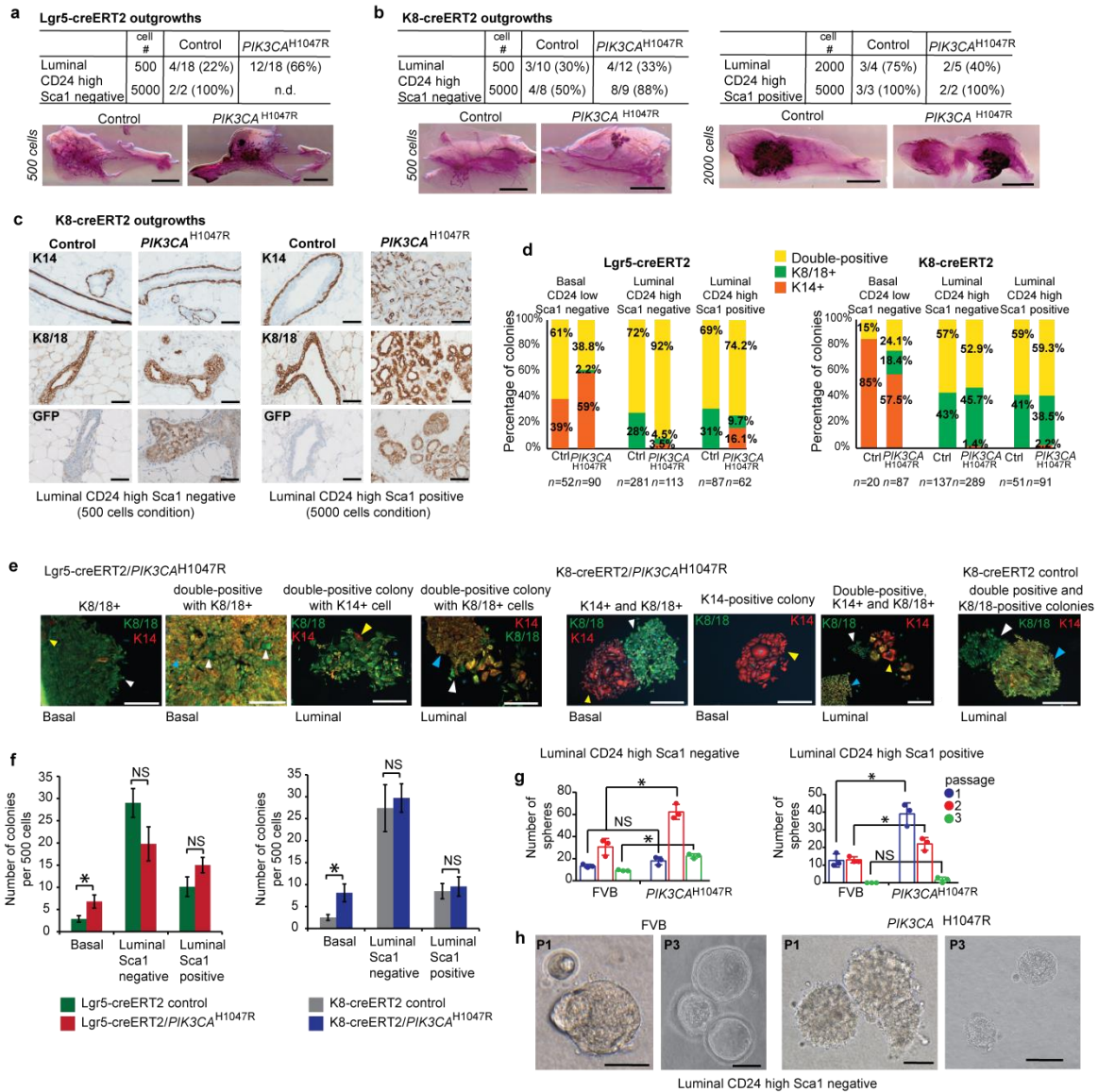
Extended Data Figure 6-8 │ Luminal *PIK3CA*[H1047R] cells repopulate a mammary gland

Number of outgrowths in cleared-fat pad transplantation of (**a**) GFP-negative Lgr5-creERT2 control and GFP-positive Lgr5-creERT2/*PIK3CA*[H1047R]-expressing luminal subsets (CD24[high]Sca1[negative]) and GFP-negative K8-creERT2 control and GFP-positive K8-creERT2/*PIK3CA*[H1047R]-expressing luminal subsets (left, CD24[high]Sca1[negative] and right, CD24[high]Sca1[positive]) (**b**). Representative carmine-stained whole mounts (bottom). Scale bar 500 μm. **c,** Representative immuno-stained sections. Scale bar 50 μm. Data from 3 independent experiments (a-c). **d,** Percentage of K14-, K8/18- and double-positive (K14/K8/18) colonies derived from Lgr5-creERT2/*PIK3CA*[H1047R], Lgr5-creERT2 control (left; pooled data from *n*=4 independent experiments (1-5 pooled mice)), K8-creERT2/*PIK3CA*[H1047R] and K8-creERT2 control subsets (right; pooled data from *n*=3 independent experiments (1-5 pooled mice)). Total number of quantified colonies is shown. **e,** Representative images of colonies. Arrowheads indicate K8/18- (white), K14- (yellow) and double-positive (blue) colonies. Scale bar 500 μm. **f,** Number of colonies derived from basal and luminal cells from Lgr5- and K8-creERT2/*PIK3CA*[H1047R] and control mice. (left: pooled data from 3 independent sortings (each 1-5 pooled animals), total *n*=8 (control), *n*=10 (mutant) technical replicates for basal subset, *n*=9 (control), *n*=5 (mutant) technical replicates for luminal CD24[high]Sca1[negative] subset and *n*=8 (control), *n*=4 (mutant) technical replicates

67

for luminal CD24^{high}Sca1^{positive} subset. Right: pooled data from 2 independent sortings (each 1-5 pooled animals), total $n=8$ (control), $n=10$ (mutant) technical replicates for basal subset, $n=5$ (control), $n=10$ (mutant) technical replicates for luminal CD24^{high}Sca1^{negative} subset and $n=6$ (control), $n=9$ (mutant) technical replicates for luminal CD24^{high}Sca1^{positive} subset). 500 cells were seeded for each replicate. A colony was defined as a cell cluster of >5 cells. Bar graphs means ±s.e.m. Two-sided unpaired student's t-test. *$P<0.05$. **g,** Bar graph showing number of spheres derived from FVB-control and *PIK3CA*^{H1047R}-expressing luminal (CD24^{high}Sca1$^{negative/positive}$) mammary cells over three passages (p1,2,3). Representative data (3 replicates, $n=4$ mice per genotype) from two independent experiments. Bar graphs means ±s.d. *$P<0.02$, two-sided unpaired student's t-test. **h,** Representative images of spheres derived from CD24^{high}Sca1^{negative} cells in passage one (P1) and three (P3). Scale bar 100 µm. NS=not significant. N.d.=not determined.
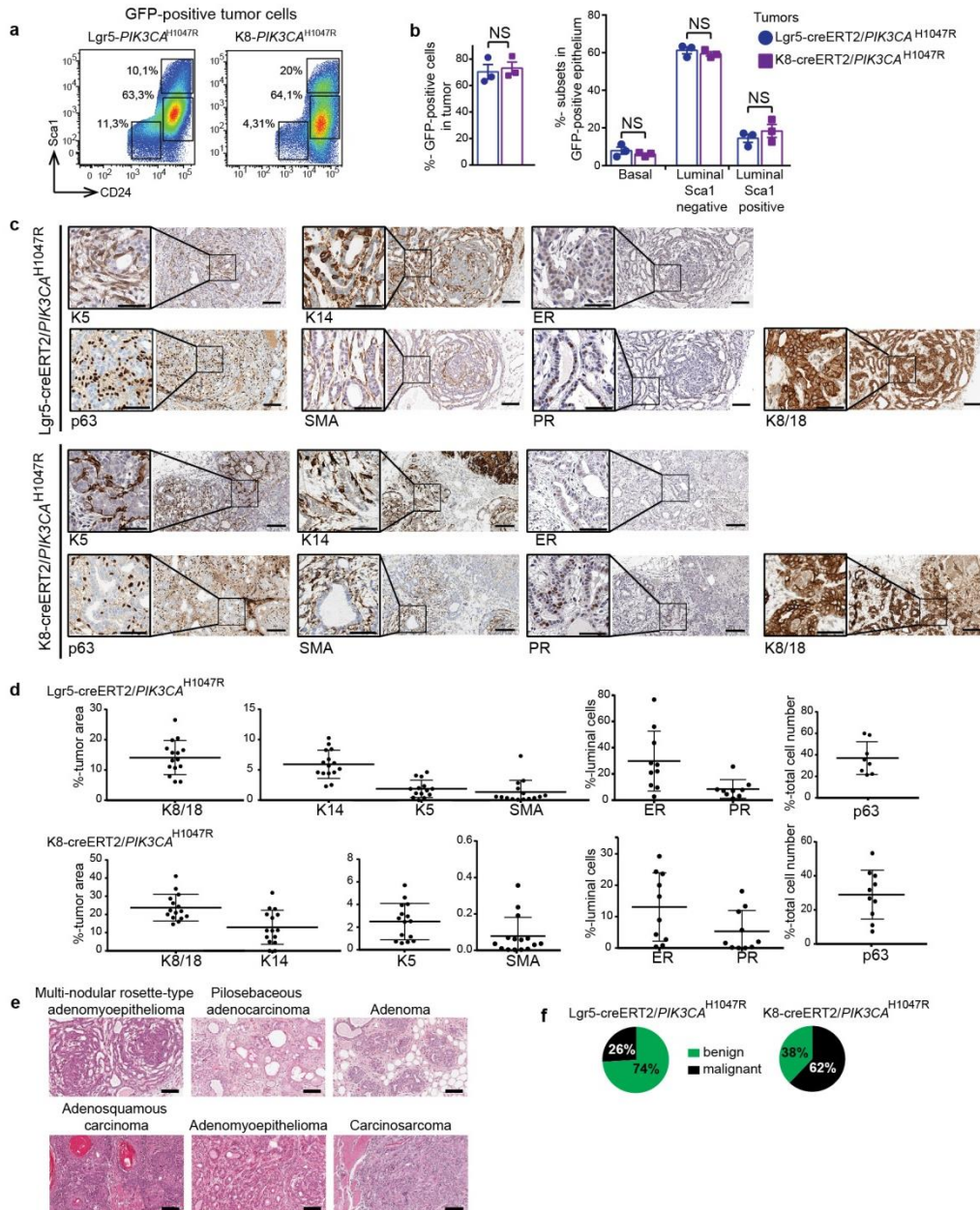
Extended Data Figure 6-9 │ $PIK3CA^{H1047R}$-evoked tumors express basal and luminal markers
a, Representative FACS plots of Lgr5-creERT2/$PIK3CA^{H1047R}$ and K8-creERT2/$PIK3CA^{H1047R}$ tumors ($n=3$). **b,** Percentages of total GFP-positive cells and GFP-positive basal (CD24^{low}Sca1^{negative}) and luminal (CD24^{high}Sca1$^{negative/positive}$) subsets of Lgr5-creERT2 and K8-creERT2/$PIK3CA^{H1047R}$-tumors ($n=3$). Bar graph means ±s.e.m. NS= not significant; two-sided unpaired student`s t-test **c,** Immunostaining for basal and luminal markers on serial sections of a multi-nodular rosette-type adenomyoepithelioma (Lgr5-creERT2/$PIK3CA^{H1047R}$) and adenomyoepithelioma (K8-creERT2/$PIK3CA^{H1047R}$). Scale bar 100 μm; magnification 50 μm. **d,** Quantification of basal- and luminal-lineage markers of Lgr5-creERT2 and K8-creERT2/$PIK3CA^{H1047R}$ tumors. Each dot represents one tumor (upper panel: K8/18, K14 and SMA $n=15$, K5 $n=14$, ER $n=10$, PR $n=9$, p63 $n=8$; lower panel: K8/18, K14, SMA and K5 $n=15$, ER, PR and p63 $n=10$). All Lgr5-creERT2/$PIK3CA^{H1047R}$ tumors and 8/10 and 6/10 of K8-creERT2/$PIK3CA^{H1047R}$ tumors show more than 1% of ER- and/or PR-positive cells, respectively. Bar graph means ±s.d. **e,** Representative H&E-stainings of tumor phenotypes. Scale bar 100 μm. **f,** Percentage of benign and malignant mammary tumors.
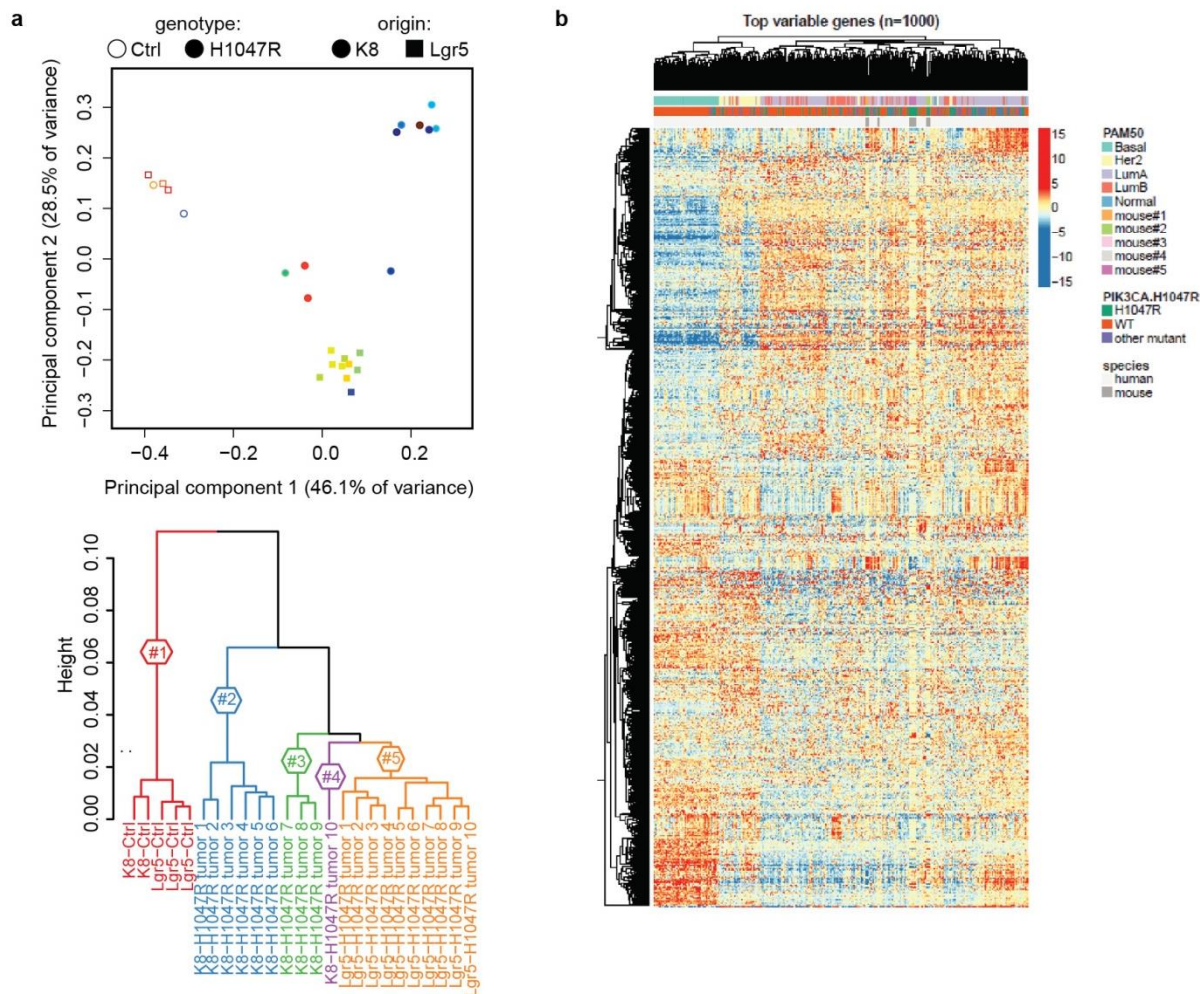
Extended Data Figure 6-10 │ Expression profiling of K8- and Lgr5-creERT2/*PIK3CA*[H1047R] mammary tumors

a, Principle component analysis and dendogram of a hierarchical clustering of gene expression profiles from 10 K8- and 10 Lgr5-CreER[T2]/*PIK3CA*[H1047R] tumors and 2–3 reference mammary glands. Each dot indicates one sample. Circles represent K8-CreER[T2] and squares represent Lgr5-CreER[T2] animals expressing *PIK3CA*[H1047R] (filled symbols) or not (open symbols). **b**, Heat map of the top 1,000 genes that vary between K8-CreER[T2]/*PIK3CA*[H1047R] and Lgr5-CreER[T2]/*PIK3CA*[H1047R] tumors and The Cancer Genome Atlas (TCGA) human breast cancer gene signatures.
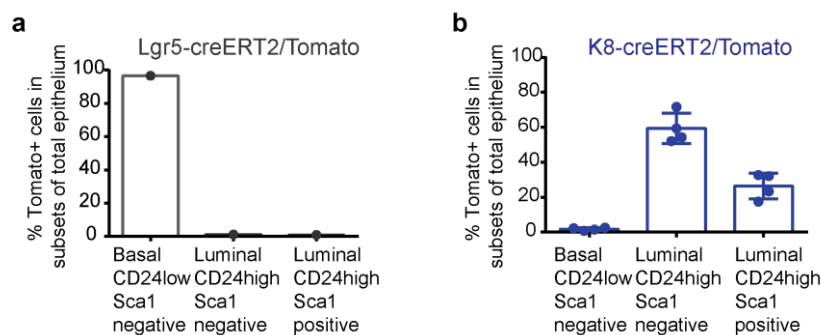
Unpublished Data Figure 6-1 │Long-term tracing of control mammary glands

FACS quantifications of Tomato-positive epithelial basal (CD24^{low}Sca1^{negative}) and luminal (CD24^{high}Sca1$^{negative/positive}$) subsets from **a,** Lgr5-creERT2/Tomato control mice (*n=1* technical replicate with 3 pooled mice) and **b,** K8-creERT2/Tomato (*n=4* technical replicates with each 2-3 pooled mice) 20-28 weeks after Tamoxifen injection.
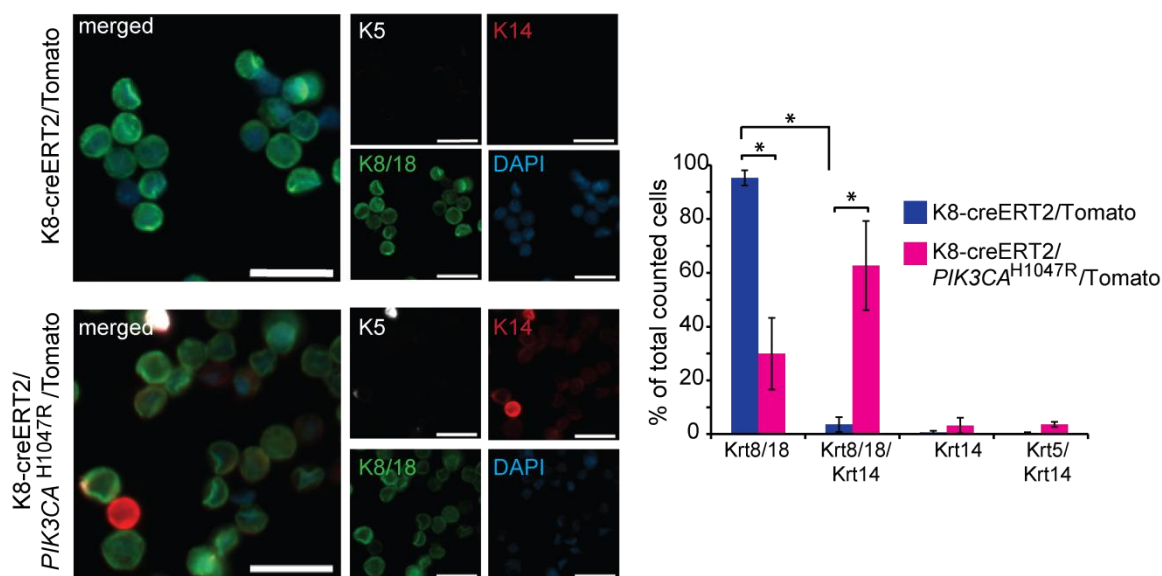
Unpublished Data Figure 6-2 │Expression of *PIK3CA*^{H1047R} leads to double-positive mammary cells

Representative images (*left*) and quantifications (*right*) of Tomato-positive sorted cells from K8-creERT2/Tomato and K8-creERT2/*PIK3CA*^{H1047R}/Tomato mammary glands 4 weeks post Tamoxifen injection stained for luminal K8/18 (green) and basal K14 (red) and K5 (white) markers. Merged (*left*) and single channels (*right*) are shown. K8-creERT2/Tomato *n=1545* and K8-creERT2/*PIK3CA*^{H1047R}/Tomato *n=872* counted cells from 3 independent FACS-sortings. Scale bar 20 μm. Unpaired student`s t-test *P<0.02*.

K8-creERT2 outgrowths

	cell #	Control
Basal CD24 low Sca1 negative	1000	1/10 (10%)
	2500	3/10 (30%)
	5000	5/8 (62.5%)
MRU frequency (95% CI)	1 in 6250 (1/3245 - 1/12077)	
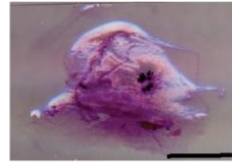
Unpublished Data Figure 6-3 │ K8-creERT2 control basal cells repopulate a mammary gland
Number of outgrowths in cleared-fat pad transplantation of GFP-negative K8-creERT2 control basal subsets (CD24^{low}Sca1^{negative}) (*left*) and representative carmine-stained whole mount (1000 cells condition) (*right*). Scale bar 500 μm. Data from one transplantation experiment with FACS-sorted cells from 4 pooled mice.

6.4 Discussion

The PI3K pathway is frequently deregulated in cancer but the effect on cell fate and lineage commitment in pre-neoplastic cells and the effect of its expression in subpopulations of the mammary gland remained ill-defined. Here, we show that expression of *PIK3CA*[H1047R] dedifferentiates lineage-restricted epithelial cells into a multipotent stem-like state from which cells further differentiate, revealing a mechanism by which heterogeneous mixed-lineage tumors may develop **(Figure 6-5)**. Evidence comes from *in vivo* genetic lineage tracing, gene expression analyses and functional assays showing that expression of *PIK3CA*[H1047R] in lineage-restricted basal or luminal mammary cells results in labeling of both basal and luminal cell lineages in pre-neoplastic mammary glands, in an increase of mammary repopulating capacity and in the formation of heterogeneous mammary tumors expressing basal and luminal markers. Furthermore, our study shows that the tumor cell-of-origin influences the frequency of malignant mammary tumors.

Figure 6-5 │ Model of the effect of *PIK3CA*[H1047R] on cell fate in pre-neoplastic mammary glands
a, Under physiological conditions Lgr5-positive basal cells contribute to the basal lineage in the adult mammary gland (black arrows). Upon expression of *PIK3CA*[H1047R] cells dedifferentiate into multipotent stem-like cells from which cells further differentiate to basal and luminal mammary epithelial cells, contributing to mostly mixed-lineage benign tumors (blue arrows). **b,** K8-positive luminal cells contribute to the homeostasis of luminal cells in the adult mammary gland (black arrows). Upon expression of *PIK3CA*[H1047R] cells dedifferentiate into multipotent stem-like cells from which cells further differentiate to basal and luminal mammary epithelial cells, contributing to mostly mixed-lineage malignant tumors (red arrows).

The observation that $PIK3CA^{H1047R}$ expression evokes luminal and basal progeny derived from Lgr5-positive basal and from K8-positive luminal cells, respectively, suggests that cells undergo dedifferentiation, a process in which cells revert back in hierarchy, or transdifferention, a process in which cells differentiate into another lineage without going through early cellular states in hierarchy. The increase of mammary repopulating and sphere-forming capacity and the increase of double-positive cells in $PIK3CA^{H1047R}$-mutant cells suggest that cells dedifferentiate.

GFP driven by the Lgr5-promoter was found in a rare luminal subset of the control gland **(Extended Data Fig. 6-5b,c)**, as shown previously (Van Keymeulen et al., 2011). Conceivably, such rare cells expand due to $PIK3CA^{H1047R}$ and form the luminal GFP-positive population. Though, we observed multi-lineage labeling upon $PIK3CA^{H1047R}$ expression in a basal as well as a luminal cell-driven model. An alternative possibility that we cannot firmly exclude is that Lgr5- and K8-positive populations contain rare bipotent subsets that are quiescent or not efficiently labeled in physiological conditions and are, therefore, not detected by lineage tracing but may expand upon $PIK3CA^{H1047R}$ expression.

Lgr5 marks stem cells in small intestine, colon (Barker et al., 2007), hair follicle (Jaks et al., 2008), and stomach (Barker et al., 2010). Moreover, overexpression of LGR5 has been found in several cancers (Uchida et al., 2010), indicating an important effect of Lgr5 on tumorigenesis. Surprisingly, expression of $PIK3CA^{H1047R}$ under control of the Lgr5-creERT2 promoter caused exclusively mammary tumors with 100% penetrance and frequent benign histotypes. Lgr5-positive cells were identified previously as multipotent stem cells fueling tumor growth in APC-deleted intestinal adenomas (Schepers et al., 2012). Remarkably, we found no evidence of gastrointestinal cancers in $PIK3CA$-mutant mice, suggesting that $PIK3CA^{H1047R}$ is a tumor-initiating mutation in breast rather than in intestine. Alternatively, cancer initiation and progression triggered by mutant $PIK3CA$ might be slower in the intestine compared to the breast and therefore, could not be identified before the end point. Different

74

wiring diagrams in adult stem cells of distinct tissues may make some cells more prone to tumorigenesis than others upon mutation of *PIK3CA*. Further studies of the nature of Lgr5-positive cells in the breast compared to other tissues are needed.

The fundamental questions of which mammary cells are susceptible to which combination of oncogenes and how this impinges on tumor progression and aggressiveness warrant further investigation. Understanding these dynamic relationships is paramount for understanding tumor heterogeneity and for identifying prognostic and predictive biomarkers.

6.5   Materials and methods

Mice. Lgr5-creERT2 mice (C57BL/6) were generated and provided by B. Kinzel and J. Tchorz (Kinzel et al., 2014). Mice were backcrossed to the FVB background and used in this study. K8-creERT2 mice (CD-1) were provided by C. Blanpain (Van Keymeulen et al., 2011). The generation of *PIK3CA*^H1047R and *PIK3CA*^WT (pure FVB) was described previously (Meyer et al., 2011; Meyer et al., 2013). Tomato-reporter animals (C57BL/6) were provided by B. Roska. For lineage-tracing studies, mice of a mixed background (FVB/CD-1/C57BL/6 and FVB/C57BL/6) were used. For studies without a Tomato-reporter, mice with a pure FVB background (Lgr5-creERT2 model) or mixed FVB/CD-1 background (K8-creERT2 model) were used. Mouse colonies were maintained in the animal facility of the Friedrich Miescher Institute for Biomedical Research and experiments were carried out in accordance with Swiss national guidelines on animal welfare and the regulations of the Canton of Basel Stadt, Switzerland.

Targeting Tomato and/or GFP expression. Adult 7- to 8-week-old mice were induced by intraperitoneal injection of Tamoxifen (2 mg/25 g of body weight) for three consecutive days (Sigma, diluted in sunflower seed oil, Sigma) to activate cell-specific expression of the Cre recombinase and thus expression of Tomato and/or the *PIK3CA*^H1047R or /WT transgene. Tomato and/or GFP were expressed in recombined cells derived from Lgr5-positive or K8-positive cells, respectively. For the Lgr5 model, after Cre induction, GFP expression was derived from both Lgr5 promoter activity and *PIK3CA*^H1047R expression.

Histology and immunostaining. Tumors and dissected mammary glands were spread on a glass slide and fixed in 20% formalin for 24 h at 4°C. Samples were then processed, embedded in paraffin and sectioned (3 μm). Immunofluorescence studies were performed using a Ventana DiscoveryUltra instrument (Roche Diagnostics) following the RUO

Discovery Universal method. Briefly, slides were pretreated with CC1 for 40 min and then incubated with primary antibodies for 1 h at 37°C. After washing, secondary antibodies were incubated for 32 min at 37°C. Slides were then washed with reaction buffer (three times), PBS (two times) and then incubated with DAPI (1 µg/ml) for 5 min. Finally, slides were rinsed with PBS (three times) and mounted in Mount Fluor (ProTaqs). The following antibodies were used: anti-keratin8/18 (guinea-pig, 1:200, Fitzgerald, 20R-CP004), anti-RFP (rabbit, 1:500, Rockland, 600-401379), anti-keratin14 (chicken, 1:500, Covance, Sig3476), anti-p63 (mouse, 1:500, Thermo Scientific, ma121871), anti-keratin5 (rabbit, 1:500, abcam, ab52635), anti-chicken Alexa Fluor 488, anti-chicken Alexa Fluor 568, anti-guinea pig Alexa Fluor 488, anti-mouse Alexa Fluor 488, anti-rabbit Alexa Fluor 647. All Alexa secondary antibodies were obtained from Molecular Probes (Invitrogen). Immunohistochemistry experiments were performed for keratin 14, keratin 8/18, GFP, and p63 using a Ventana DiscoveryXT instrument (Roche Diagnostics) following the Research IHC DAB Map XT procedure. Slides were treated with a mild CC1 and incubated with the primary antibodies for 1 h at 37°C. Following brief washes, biotinylated donkey-anti-rabbit and biotinylated anti-guinea-pig, respectively, were applied for 32 min at 37°C. For mouse-anti p63 detection, a monoclonal rabbit-anti-mouse antibody was applied for 32 min at 37°C, followed by incubation with a polymer anti-rabbit conjugated with HRP (ImmPRESS anti-rabbit peroxidase, Vector Laboratories) for 32 min at 37°C. Finally, sections were counterstained with hematoxylin II and bluing reagent (4 min). Staining against ERα-, SMA- and keratin 5 was performed by deparaffinization followed by antigen retrieval with citrate buffer and quenching with PBS+3% H_2O_2. Slides were then blocked with PBS+2,5% normal goat serum and primary antibodies incubated overnight at 4°C in PBS+1% BSA+ 0.5% Tween-20. After brief washes, secondary antibodies were then incubated in PBS+1% BSA for 30 min at RT. Signals were enhanced using the Vectastain ABC system and visualized with 3,3`-diaminobenzidine (DAB; Sigma). Hematoxylin was used as counterstain. Anti-PR-staining

was performed without antigen retrieval. The following antibodies were used: anti-keratin 8/18 (guinea-pig, 1:500, Fitzgerald, 20R-CP004), anti-keratin 14 (rabbit, 1:500, Thermo Scientific, Rb9020), anti-p63 (mouse, 1:1000, Thermo Scientific, ma121871), anti-keratin 5 (rabbit, 1:1000, abcam, ab52635), anti-SMA (rabbit, 1:500, Thermo Scientific, Rb9010), anti-ERα (rabbit, 1:1000, Santa Cruz, sc-542), anti-PR (rabbit, 1:200, Thermo Scientific, Rm9102), anti-GFP (rabbit, 1:50, Invitrogen, A11122), Secondary antibodies were biotinylated anti-rabbit IgG (1:200, Jackson Immunoresearch), biotinylated anti-guinea pig IgG (1:200; Vector Laboratories), and biotinylated anti-mouse IgG (1:200; Abcam). Hematoxylin and eosin staining was performed using standard protocols.

Microscopy image acquisition. For immunofluorescence, images of stained sections were captured using a Zeiss Z1 wide-field fluorescent microscope, 5x/0.13, 10x/0.45 or 20x/0.8 (Plan-APOCHROMAT) objectives, an AxioCamMRc camera (1'024 x 1'024, pixel size 6.45 μm) and an Axiocam506 camera (2'752 x 2'208, pixel size 4.54 μm). Whole-mount fluorescent mammary gland and tumors sections were imaged on the Zeiss Axio Scan.Z1 slide scanner, (ORCA-Flash4.0 camera, 2'048 x 2'048, pixel size 6.5 μm). Representative images were cropped and processed using the ZenBlue software. For immunohistochemistry, stained sections were examined using a Nikon E600 Eclipse brightfield microscope (20x/0.5 and 40x/0.75 objectives) and images captured with a Nikon DXM1200 camera (2'592 x 1'944, pixel size 6.7 μm) using the IMS acquisition software. All images were scaled appropriately.

Quantification of immunohistochemistry. Five representative images of 8-15 tumors of each genotype were captured with a Nikon E600 Eclipse brightfield microscope (20x/0.5 objective) and positively stained tumor areas, luminal cells, and total epithelial cells

quantified with Image J (Fiji). Evaluation of tissue sections was performed blindly by two independent investigators.

Quantification of double-positive tumor epithelial area. 10 tumors from each model were stained with anti-keratin14, anti-keratin5, anti-keratin8/18, anti-rabbit Alexa Fluor 647, anti-chicken Alexa Fluor 568, and anti-guinea pig Alexa Fluor 488. Whole-mount tumor sections were scanned using the Zeiss Axio Scan Z1 slide scanner. For computational reasons, the whole images of tumors were first tiled and the channels split using the ZenBlue software and Matlab. Images were then processed using Ilastik (Sommer et al., 2011), an interactive supervised Machine Learning toolkit, and the subsequent prediction maps were treated with batch functions written with the Matlab programing language. For the statistical analysis shown in **Fig. 6-4b**, a total of 23 regions of tumors were subsequently tiled into 14'872 squared images of 1'024x1'024 pixels. In a second step, data from a set of 9 selected tiles were generated by annotating a few regions representative of the predefined classes. For each of the three channels the training data were processed using the four phenotype classes: '1.background', '2.fluorescent marker (Alexa Fluor 488, 568 or 647)', '3.blood cells', and '4.Stroma'. After annotation of the different pixels of the different classes by brush stroke, features of the labeled pixels and their local neighborhood are used to train a Random Forest classifier. In a third step, the inferred classifier was used to predict all the tiles in a batch process. The last step implemented in Matlab determined the masks for all the channels and tiles. In order to perform the analysis on relevant tiles, i.e. tiles sufficiently covered by tissue, we performed a k-mean clustering on a set of statistical features extracted from the tiles. This enabled us to suppress tiles that were merely in the background or those that contained non-epithelial structures. For each tile, the fluorescent marker mask corresponded to pixels having a probability above 50% in the class 'fluorescent marker (Alexa)'. In addition, a non-tissue mask was calculated by the mean of morphological filters on the union of background, blood

cells and stroma classes. The mask for tissue represented the complement image of the non-tissue mask. We calculated on each tile the ratio between the areas of pixels in the mask of the fluorescent marker and those in the tissue mask. Finally, the distribution of double-labelled fluorescent marker within the whole tumor epithelial area was represented in **Fig. 6-4b**.

Preparation of mammary single-cell suspensions and labeling. Mammary glands were dissected and intra-mammary lymph nodes removed. To obtain mammary organoids, mammary glands were processed as described previously (Meier-Abt et al., 2013; Sleeman et al., 2006). To obtain single mammary epithelial cells, organoids were washed in serum-free Leibowitz L15-medium (Gibco) and digested with Hyclone HyQTase (Thermo Scientific). Single cells were washed and filtered through a 40-µm cell strainer (BD Falcon) and counted; 10^6 cells/ml were stained with the following antibodies: PE-Cy7-CD45 (Biolegend; clone 30-F11), APC-Sca1 (Biolegend; clone E13-161.7), PerCP-Cy5.5-CD24 (Biolegend; clone M1/69), PE-CD49f (BD-Pharmingen), Alexa700-CD24 (Novus Biologicals; clone M1/69), and DAPI (2 µg/ml, Invitrogen).

Flow cytometry. FACS was carried out with a BD FACSAria III (Becton Dickinson) using a 100-µm nozzle. Cells were gated based on their forward- and sideward-scatter. Pulse-width was used to exclude doublets. DAPI-negative/CD45-negative cells were gated for Tomato or GFP. CD24, Sca1, and CD49f subsets were then gated on GFP-positive epithelium. Tomato-positive cells were gated using only CD24, CD45, and Sca1 antigens. The same numbers of living cells were recorded in each condition. FACS data were analyzed using FlowJo (Tree Star). Total cell numbers were determined by enumerating the total epithelial content after single-cell isolation and calculating back based on the percentages obtained from antibody-staining, sorting, and FlowJo analysis. Cell numbers were subsequently normalized to one

animal (3-9 independent FACS-experiments of 1-5 pooled animals per time point were quantified; ±s.e.m.; $P<0.05$, student`s t-test).

***In vitro* colony formation assay and quantification.** Freshly sorted cells of each subpopulation (500 cells) were plated as previously described (Meier-Abt et al., 2013). Seven days later, the colonies were fixed with acetone/methanol (1:1), washed, blocked with 2.5% normal goat serum and stained with anti-keratin8/18 (guinea-pig, 1:500, Fitzgerald, 20R-CP004), anti-keratin14 (rabbit, 1:500, Thermo Scientific, Rb9020), DAPI (2 µg/ml, Invitrogen), anti-guinea-pig Alexa Fluor 488 and anti-rabbit Alexa Fluor 647 (1:1000, Invitrogen). Colonies were imaged using the Zeiss Z1 wide-field fluorescent microscope (5x/0.13 DIC). Colonies were defined as a cluster of more than five cells. The number of colonies per well was determined manually. Colonies containing more than 20% of keratin8/18/keratin14 double-positive cells were defined as "double-positive".

***In vitro* mammosphere culture.** Mammosphere cultures were performed as described previously (Cicalese et al., 2009). Freshly sorted subsets (luminal CD24^high^Sca1^negative^ and CD24^high^Sca1^positive^) cells from adult FVB and uninduced *PIK3CA*^H1047R^ females were plated at 20'000 cells/ml in 6-well ultra-low attachment plates (Falcon) in DMEM/F12 medium (Gibco) supplemented with 5 µg/ml insulin, 0.5 µg/ml hydrocortisone, 2% B27 (Invitrogen), 20 ng/ml EGF and bFGF (BD Biosciences), and cholera toxin (Sigma) and cultured at 37°C in 5% CO_2. After 5 days, mammospheres were collected and dissociated into single cells using HYQtase (Gibco) and counted. Luminal subsets from FVB control and *PIK3CA*^H1047R^ animals were subsequently treated with TAT-Cre (Millipore) (0.5 µM) at a density of 20'000 cells/ml in 4 ml overnight at 37°C in 5% CO_2. The next morning, the medium was replaced and cells were cultured for 72 h at 37°C in 5% CO_2 to allow for maximal recombination and expression of *PIK3CA*^H1047R^. After 72 h, *PIK3CA*^H1047R^-GFP^positive/negative^ cells were sorted and plated into

24-well ultralow attachment plates in medium (described above) supplemented with 2% Matrigel (growth factor reduced; BD; 356230) at a density of 1'000 cells/well. Control and *PIK3CA*^H1047R^ spheres from each subset were enumerated every 7 days, at which point spheres were dissociated and re-plated at a density of 1'000 cells/well.

Mammary fat pad transplantation. For limiting dilution transplantation, freshly sorted cells from Lgr5-creERT2 control and Lgr5-creERT2/*PIK3CA*^H1047R^ animals (pure FVB background) and from K8-creERT2 control and K8-creERT2/*PIK3CA*^H1047R^ F1-hybrid (FVB/CD-1) littermates were used. Sorted cells were resuspended in limiting dilution numbers in PBS+2% FCS with 25% Matrigel (growth factor reduced; BD; 356230) and injected in 20-µl volumes into inguinal glands of 3-week-old FVB females that had been cleared of endogenous mammary epithelium (Deome et al., 1959). Cells from control and from *PIK3CA*^H1047R^ animals were injected in the same animal on opposite sides. After 8 weeks, glands of the recipients were removed for evaluation. Glands were spread on a glass slide and fixed in Carnoy`s fixative overnight. Whole mount staining with carmine alum was performed as previously described (Meier-Abt et al., 2013) and scanned with an Epson 1600 Pro scanner. An outgrowth was defined as an epithelial structure composed of ducts arising from a central point with lobules and/or terminal end buds (Shackleton et al., 2006). Frequencies of mammary-repopulating-units between different cell populations were calculated and statistically compared using the Extreme Limiting Dilution Analysis (ELDA) (Hu and Smyth, 2009) online tool (http://bioinf.wehi.edu.au/software/elda/).

Immunoblotting. Lysates from mammary glands were prepared by lysing cryo-homogenized mammary gland powder in RIPA buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) supplemented with 1x protease inhibitor cocktail (Complete Mini, Roche), 0.2 mMol/L sodium orthovanadate, 20 mM sodium fluoride and 1

mMol/L phenylmethylsulfonyl fluoride. Lysates (30 – 80 µg) were subjected to SDS-PAGE, transferred to PVDF membranes (Immobilon-P, Millipore) and blocked for 1 h at room temperature with 5% milk or BSA in PBS-0.05% Tween 20. Membranes were then incubated overnight with primary antibodies (1:200 – 1:3000) and exposed to secondary HRP-coupled anti-mouse or anti-rabbit antibodies at 1:5'000–10'000 for 1 h at room temperature. Results are representative of at least three different experiments. The following antibodies were used: anti-AKT pan (Cell Signaling), anti-pAKT (Ser473, Cell Signaling), anti-p110α (Cell Signaling), anti-ERK2 (Santa Cruz) and anti-Keratin pan (SantaCruz). Blot densities were quantified using ImageJ and normalized to pan-Keratin for epithelial content.

RNA isolation. For pre-neoplastic gene expression profiling, mammary epithelial subsets were FACS sorted as described above into extraction buffer and total RNA from 250 or 2'000 sorted cells isolated using an Arcturus PicoPure RNA Isolation Kit (Life Technologies). Subsets of pooled mammary glands of 2-3 estrus-synchronized animals (confirmed by vaginal smear) per genotype in three independent sortings were collected for microarray analysis. RNA from mammary subsets of four animals per genotype was collected for QRT-PCR. For tumor gene expression profiling and expression of Lgr5 in the mammary gland, total RNA from 50 mg cryo-homogenized tumor tissue or nipple and distal area of mammary glands were extracted using the TRIzol method (Life Technologies) according to the manufacturer's instructions. Genomic DNA was removed by DNAse I digestion (Qiagen) and RNA purified using an RNeasy Plus Mini Kit (Qiagen). RNA concentration was measured with a Nanodrop 1000 machine and RNA quality assessed using an Agilent 2100 bioanalyzer and RNA Pico Chips or RNA Nano Chips.

Microarray. Total RNA from 250 sorted cells (pre-neoplastic Lgr5), 2'000 sorted cells (pre-neoplastic K8) or 100 ng from tumor tissue was used as the input for synthesis of amplified

cDNA with the NuGen Ovation Pico WTA System (NuGen Inc.). The resulting double-stranded cDNA was fragmented and labeled using the Affymetrix GeneChip WT Terminal Labeling kit (Affymetrix). Affymetrix Gene Chip Mouse gene 1.0 ST microarrays were hybridized according to the GeneChip Whole Transcript (WT) Sense Target Labeling Assay Manual (Affymetrix) with a hybridization time of 16 h. Scanning was performed with Affymetrix GCC Scan Control Software v. 3.0.0.1214 on a GeneChip Scanner 3000 7 G with autoloader.

Normalization and analysis of mammary epithelial subset microarray data. The arrays for GEO GSE40875 (Meier-Abt *et al.*), GSE59870 (**Fig. 6-2a**), and GSE65411 (**Fig. 6-2b**) datasets were RMA normalized using the bioconductor package affy (R3.0.1/Bioconductor 2.13). Heatmaps with unscaled normalized expression values for selected genes were plotted with the heatmap.2 function of the gplots package. Differential gene expression for the comparison of mature luminal (LM) or luminal progenitor (LP) versus myoepithelial (MYO) cells was calculated with limma. The function topTable was used to select the top 300 upregulated genes as luminal mature versus myoepithelial UP (Luminal_Mature UP, MEIER-ABT) and luminal progenitors versus myoepithelial UP (Luminal_progenitors UP, MEIER-ABT), and the top 300 upregulated genes as myoepithelial versus luminal mature UP (Myoepithelial UP, MEIER-ABT). Signatures LIM_MAMMARY_LUMINAL_MATURE_DN,

LIM_MAMMARY_LUMINAL_MATURE_UP,

LIM_MAMMARY_LUMINAL_PROGENITOR_DN,

LIM_MAMMARY_LUMINAL_PROGENITOR_UP,

LIM_MAMMARY_STEM_CELL_DN and LIM_MAMMARY_STEM_CELL_UP were downloaded from http://www.broadinstitute.org/gsea/msigdb/search.jsp in gmt format,

combined with the signatures described above, and used in the PGSEA and smcPlot functions of the Bioconductor PGSEA package.

QRT-PCR. RNA was converted into cDNA using SuperScript III Reverse Transcriptase (Invitrogen). Quantitative real-time PCR was performed on unamplified cDNA normalized to the number of sorted cells for each subset (6'000-10'000 sorted cells). Taqman-probes (Life technologies) and Taqman Universal PCR Mastermix (Applied Biosystems) were applied. The following Taqman-probe IDs were used: Krt14 Mm00516879_m1, Krt5 Mm01305291_g1, Lgr5 Mm00438890_m1, Vim Mm01333430_m1, Krt8 Mm00835759_m1, Krt18 Mm01601702_g1, Gata3 Mm00484683_m1, Elf5 Mm00468732_m1, Csn2 Mm04207885_m1, Axin2 Mm00443610_m1, and PIK3CA Hs00907957_m1. Cycling was performed with StepOne Plus Real-time PCR Systems (Applied Biosystems). The results are representative of three QRT-PCR experiments of pooled mammary subsets from four animals of each genotype. Results of *Csn2* (CD24^{high}Sca1$^{negative/positive}$) and *Krt14* (CD24^{high}Sca1^{positive}) from K8-CreERT2/*PIK3CA*^{H1047R} cells and of *Krt5* and *Lgr5* (CD24^{high}Sca1^{negative}) from Lgr5-CreERT2/*PIK3CA*^{H1047R} cells are representative of two experiments. Statistical data analysis was performed using deltaCT-values.

Normalization and analysis of tumor mouse microarray data. Mouse Affymetrix microarrays (GSE59872) were background corrected, quantile normalized, and log2 transcript cluster expression values calculated using the rma() function of the Bioconductor package "oligo" (Carvalho and Irizarry, 2010). Transcript cluster identifiers were mapped to Entrez Gene identifiers using the Bioconductor package "mogene10sttranscriptcluster.db", and transcript clusters associated with none or multiple genes were removed. Where multiple transcript clusters were associated with a single e, that with the maximal variance across samples was selected; this resulted in a total number of 20'365 transcript clusters with unique

gene assignments. Principal component analysis was performed on log2 expression values from which the mean over samples had been subtracted for each gene. Hierarchical clustering of samples was performed using (1-r) as distance metric, with r being the Pearson's correlation coefficient for log2 expression values between each pair of samples. The clustering dendrogram was visualized using the "dendextend" R package.

Analysis of human TCGA breast cancer data and comparison to mouse. Human breast cancer expression data and corresponding clinical data (Cancer Genome Atlas, 2012) were obtained from https://tcga-data.nci.nih.gov/docs/publications/brca_2012/, corresponding to the November 11, 2011 data freeze that contains 522 tumor samples with clinical annotation. Human gene symbols were mapped to Entrez Gene identifiers and genes selected that are one-to-one homologues between human and mouse according to Homologene (build 68, downloaded from ftp://ftp.ncbi.nlm.nih.gov/pub/HomoloGene/build68/ (Coordinators, 2014)) and that were measured on both human and mouse experimental platforms; this resulted in 13'969 genes. In order to reduce technical differences between human and mouse samples, all samples were scaled to a standard deviation of 1, and the within-special mean was subtracted from each gene. Heatmap and clustering of the combined human and mouse expression data were performed using the function aheatmap() from the R package "NMF" (Gaujoux and Seoighe, 2010) on the top 1'000 genes ranked by variance across human samples. Mouse samples were compared to human PAM50 tumor subtypes (Normal, LumA, LumB, Her2 and Basal) by calculating Pearson's correlation coefficients for each mouse sample against the averages of all human samples within each subtype.

Survival analysis. Tumor incidence was determined by palpation. Kaplan-Meier plots were generated using the survival calculation tool from Graphpad Prism and significance calculated using the Log-rank test.

Statistical data analysis. The number of mice was calculated by performing power analysis using data from small pilot experiments. Values represent the means ±s.e.m (standard error of the mean) or ±s.d. (standard deviation). Depending on the type of experiments, data were tested using unpaired student's t-test or Log-rank test. *P<0.05 was considered statistically significant. The experiments were not randomized.

7 | Concluding remarks and perspectives

Current mouse models of mutant *PIK3CA* are contributing to a better understanding of breast cancer pathogenesis associated with alterations in the PI3K pathway. By using transgenic animals expressing *PIK3CA*[H1047R] and *PIK3CA*[E545K] mutations, we unraveled that *PIK3CA*[H1047R] is a more potent inducer of mammary tumors than *PIK3CA*[E545K], what might explain the higher frequency of *PIK3CA*[H1047R] mutations in human breast cancer. Future investigations are needed to decipher further potential distinct downstream signaling in E545K and H1047R-expressing mutants to better understand and explain differences in potency.

Moreover, by using *in situ* genetic lineage tracing, gene expression analyses and functional assays, we show that *PIK3CA*[H1047R] causes dedifferentiation of mammary cells to a multipotent stem-like state, which may be a mechanism by which heterogeneity develops in breast cancer. Additionally, we revealed that the cell-of-origin dictates the frequency of malignant tumor lesions, as the expression of *PIK3CA*[H1047R] in luminal mammary epithelial cells causes more frequently aggressive mammary tumors compared to expression in basal mammary epithelial cells.

These findings lead to further questions. What is the mechanism underlying the observed dedifferentiation to a multipotent stem-like state? Expression of *PIK3CA*[H1047R] modulates gene expression. This might be caused by regulation or activation of transcriptional factors, co-factors and/or epigenetic remodelers which may impinge on chromatin structure and eventually unlock multipotency programs. During mammary gland development dynamic changes occur in the mammary epigenome involving polycomb-group-complexes (PRC) and DNA methylation (Pal et al., 2013; Rijnkels et al., 2010). Also cancer hallmarks were described to be controlled by posttranslational modifications of histones and methylation of DNA. Epigenetic remodeling and genetic alterations may work cooperatively to promote tumor initiation and progression in breast cancer (Shen and Laird, 2013; Stefansson and

Esteller, 2013). It was previously shown that activated Akt phosphorylates the PRC2 member EZH2 at an inhibitory site leading to reduction of methylated H3K27 sites (Cha et al., 2005). Moreover, epigenetic silencing was found to be mediated by the PI3K-pathway in breast cancer (Zuo et al., 2011). Therefore, it would be of profound interest to investigate whether *PIK3CA*[H1047R] oncogene-induced dedifferentiation is triggered by epigenetic events.

We did not find metastasis in animals expressing *PIK3CA*[H1047R] or *PIK3CA*[E545K] under various promoters (WAPiCre, CAGs-creERT2, Lgr5-creERT2, K8-creERT2, K14-cre). Similarly, other studies rarely report metastasis in these mouse models (please see section 3.5.3). It would be of interest to investigate whether additional genetic alterations are required for metastasis in a mutant *PIK3CA* context. Conceivably, metastatic spread might be a slow process in these models and therefore could not be found before the end point. Transplantation studies would be helpful to investigate the role of mutant *PIK3CA* in metastasis. Alternatively, expression of mutant *PIK3CA* in other mammary epithelial cell subtypes, which were not tested yet, may cause metastasis.

Oncogene-induced multipotency and the importance of the cell-of-origin on tumor phenotype were shown in this study by using *PIK3CA*[H1047R]. *PTEN*, a negative regulator of the PI3K-pathway, was shown to be essential in lineage fate determination and restriction of activation of hematopoietic stem cells (Zhang et al., 2006). Moreover, distinct molecular signatures that are predictive of human patient outcome were reported in *Pten* inactivated prostate tumors from basal or luminal origins. *Pten* deletion in luminal prostate cells evokes more aggressive tumors compared to its expression in basal cells (Wang et al., 2013). Furthermore, previous lineage-tracing approaches show that basal prostate cells differentiate into luminal cells upon *Pten* loss suggesting that deregulation of epithelial differentiation is a critical step for tumor initiation (Choi et al., 2012). Little is known whether other *PIK3CA* mutations (e.g. E545K) or oncogenes and tumor suppressors also influence cell fate decisions.

Further lineage-tracing studies in distinct mammary epithelial subpopulations expressing different tumor-initiating lesions could answer this question.

We and others now show that the cell-of-origin dictates the tumor phenotype, molecular subtype and clinical outcome. Mammary tumors arising from luminal epithelial cells cluster more frequently with basal-like cancers. Moreover, luminal to basal cell fate changes were shown to have a poorer disease-free survival (Koren et al., 2015; Van Keymeulen et al., 2015). It is yet unknown whether the cellular origin of breast cancer could predict treatment success. The influence of the cell-of-origin on response to treatment should be the focus of future investigations.

8 | References

Adams, J.R., Xu, K., Liu, J.C., Agamez, N.M., Loch, A.J., Wong, R.G., Wang, W., Wright, K.L., Lane, T.F., Zacksenhaus, E.*, et al.* (2011). Cooperation between Pik3ca and p53 mutations in mouse mammary tumor formation. Cancer research *71*, 2706-2717.

Allinen, M., Beroukhim, R., Cai, L., Brennan, C., Lahti-Domenici, J., Huang, H., Porter, D., Hu, M., Chin, L., Richardson, A.*, et al.* (2004). Molecular characterization of the tumor microenvironment in breast cancer. Cancer cell *6*, 17-32.

Andrechek, E.R., White, D., and Muller, W.J. (2005). Targeted disruption of ErbB2/Neu in the mammary epithelium results in impaired ductal outgrowth. Oncogene *24*, 932-937.

Bachman, K.E., Argani, P., Samuels, Y., Silliman, N., Ptak, J., Szabo, S., Konishi, H., Karakas, B., Blair, B.G., Lin, C.*, et al.* (2004). The PIK3CA gene is mutated with high frequency in human breast cancers. Cancer biology & therapy *3*, 772-775.

Bader, A.G., Kang, S., and Vogt, P.K. (2006). Cancer-specific mutations in PIK3CA are oncogenic in vivo. Proceedings of the National Academy of Sciences of the United States of America *103*, 1475-1479.

Bader, A.G., Kang, S., Zhao, L., and Vogt, P.K. (2005). Oncogenic PI3K deregulates transcription and translation. Nature reviews. Cancer *5*, 921-929.

Banerji, S., Cibulskis, K., Rangel-Escareno, C., Brown, K.K., Carter, S.L., Frederick, A.M., Lawrence, M.S., Sivachenko, A.Y., Sougnez, C., Zou, L.*, et al.* (2012). Sequence analysis of mutations and translocations across breast cancer subtypes. Nature *486*, 405-409.

Barbareschi, M., Buttitta, F., Felicioni, L., Cotrupi, S., Barassi, F., Del Grammastro, M., Ferro, A., Dalla Palma, P., Galligioni, E., and Marchetti, A. (2007). Different prognostic roles of mutations in the helical and kinase domains of the PIK3CA gene in breast carcinomas. Clinical cancer research : an official journal of the American Association for Cancer Research *13*, 6064-6069.

Barker, N., Huch, M., Kujala, P., van de Wetering, M., Snippert, H.J., van Es, J.H., Sato, T., Stange, D.E., Begthel, H., van den Born, M.*, et al.* (2010). Lgr5(+ve) stem cells drive self-renewal in the stomach and build long-lived gastric units in vitro. Cell stem cell *6*, 25-36.

Barker, N., van Es, J.H., Kuipers, J., Kujala, P., van den Born, M., Cozijnsen, M., Haegebarth, A., Korving, J., Begthel, H., Peters, P.J.*, et al.* (2007). Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature *449*, 1003-1007.

Bhang, H.E., Ruddy, D.A., Krishnamurthy Radhakrishna, V., Caushi, J.X., Zhao, R., Hims, M.M., Singh, A.P., Kao, I., Rakiec, D., Shaw, P.*, et al.* (2015). Studying clonal dynamics in response to cancer therapy using high-complexity barcoding. Nature medicine *21*, 440-448.

Bissell, M.J., and Hines, W.C. (2011). Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression. Nature medicine *17*, 320-329.

References

Bissell, M.J., Rizki, A., and Mian, I.S. (2003). Tissue architecture: the ultimate regulator of breast epithelial function. Current opinion in cell biology *15*, 753-762.

Booth, B.W., Boulanger, C.A., and Smith, G.H. (2007). Alveolar progenitor cells develop in mouse mammary glands independent of pregnancy and lactation. Journal of cellular physiology *212*, 729-736.

Borowsky, A.D. (2011). Choosing a mouse model: experimental biology in context--the utility and limitations of mouse models of breast cancer. Cold Spring Harbor perspectives in biology *3*, a009670.

Boulanger, C.A., Wagner, K.U., and Smith, G.H. (2005). Parity-induced mouse mammary epithelial cells are pluripotent, self-renewing and sensitive to TGF-beta1 expression. Oncogene *24*, 552-560.

Bousquet, G., Feugeas, J.P., Ferreira, I., Vercellino, L., Jourdan, N., Bertheau, P., de Bazelaire, C., Barranger, E., and Janin, A. (2014). Individual xenograft as a personalized therapeutic resort for women with metastatic triple-negative breast carcinoma. Breast cancer research : BCR *16*, 401.

Boyault, S., Drouet, Y., Navarro, C., Bachelot, T., Lasset, C., Treilleux, I., Tabone, E., Puisieux, A., and Wang, Q. (2012). Mutational characterization of individual breast tumors: TP53 and PI3K pathway genes are frequently and distinctively mutated in different subtypes. Breast cancer research and treatment *132*, 29-39.

Brooks, M.D., Burness, M.L., and Wicha, M.S. (2015). Therapeutic Implications of Cellular Heterogeneity and Plasticity in Breast Cancer. Cell stem cell *17*, 260-271.

Bruno, R.D., and Smith, G.H. (2011). Functional characterization of stem cell activity in the mouse mammary gland. Stem cell reviews *7*, 238-247.

Buttitta, F., Felicioni, L., Barassi, F., Martella, C., Paolizzi, D., Fresu, G., Salvatore, S., Cuccurullo, F., Mezzetti, A., Campani, D.*, et al.* (2006). PIK3CA mutation and histological type in breast carcinoma: high frequency of mutations in lobular carcinoma. The Journal of pathology *208*, 350-355.

Campbell, I.G., Russell, S.E., Choong, D.Y., Montgomery, K.G., Ciavarella, M.L., Hooi, C.S., Cristiano, B.E., Pearson, R.B., and Phillips, W.A. (2004). Mutation of the PIK3CA gene in ovarian and breast cancer. Cancer research *64*, 7678-7681.

Cancer Genome Atlas, N. (2012). Comprehensive molecular portraits of human breast tumours. Nature *490*, 61-70.

Cantley, L.C. (2002). The phosphoinositide 3-kinase pathway. Science *296*, 1655-1657.

Cardiff, R.D. (2001). Validity of mouse mammary tumour models for human breast cancer: comparative pathology. Microscopy research and technique *52*, 224-230.

References

Cardiff, R.D., Anver, M.R., Gusterson, B.A., Hennighausen, L., Jensen, R.A., Merino, M.J., Rehm, S., Russo, J., Tavassoli, F.A., Wakefield, L.M., *et al.* (2000). The mammary pathology of genetically engineered mice: the consensus report and recommendations from the Annapolis meeting. Oncogene *19*, 968-988.

Carvalho, B.S., and Irizarry, R.A. (2010). A framework for oligonucleotide microarray preprocessing. Bioinformatics *26*, 2363-2367.

Cassidy, J.W., Caldas, C., and Bruna, A. (2015). Maintaining Tumor Heterogeneity in Patient-Derived Tumor Xenografts. Cancer research *75*, 2963-2968.

Cha, T.L., Zhou, B.P., Xia, W., Wu, Y., Yang, C.C., Chen, C.T., Ping, B., Otte, A.P., and Hung, M.C. (2005). Akt-mediated phosphorylation of EZH2 suppresses methylation of lysine 27 in histone H3. Science *310*, 306-310.

Chaffer, C.L., Brueckmann, I., Scheel, C., Kaestli, A.J., Wiggins, P.A., Rodrigues, L.O., Brooks, M., Reinhardt, F., Su, Y., Polyak, K., *et al.* (2011). Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. Proceedings of the National Academy of Sciences of the United States of America *108*, 7950-7955.

Chaffer, C.L., Marjanovic, N.D., Lee, T., Bell, G., Kleer, C.G., Reinhardt, F., D'Alessio, A.C., Young, R.A., and Weinberg, R.A. (2013). Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity. Cell *154*, 61-74.

Choi, N., Zhang, B., Zhang, L., Ittmann, M., and Xin, L. (2012). Adult murine prostate basal and luminal cells are self-sustained lineages that can both serve as targets for prostate cancer initiation. Cancer cell *21*, 253-265.

Cicalese, A., Bonizzi, G., Pasi, C.E., Faretta, M., Ronzoni, S., Giulini, B., Brisken, C., Minucci, S., Di Fiore, P.P., and Pelicci, P.G. (2009). The tumor suppressor p53 regulates polarity of self-renewing divisions in mammary stem cells. Cell *138*, 1083-1095.

Clark, S.E., Warwick, J., Carpenter, R., Bowen, R.L., Duffy, S.W., and Jones, J.L. (2011). Molecular subtyping of DCIS: heterogeneity of breast cancer reflected in pre-invasive disease. British journal of cancer *104*, 120-127.

Cleary, A.S., Leonard, T.L., Gestl, S.A., and Gunther, E.J. (2014). Tumour cell heterogeneity maintained by cooperating subclones in Wnt-driven mammary cancers. Nature *508*, 113-117.

Cohen, J.C., Shank, P.R., Morris, V.L., Cardiff, R., and Varmus, H.E. (1979). Integration of the DNA of mouse mammary tumor virus in virus-infected normal and neoplastic tissue of the mouse. Cell *16*, 333-345.

Coordinators, N.R. (2014). Database resources of the National Center for Biotechnology Information. Nucleic acids research *42*, D7-17.

References

Courtney, K.D., Corcoran, R.B., and Engelman, J.A. (2010). The PI3K pathway as drug target in human cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology *28*, 1075-1083.

Curtis, C., Shah, S.P., Chin, S.F., Turashvili, G., Rueda, O.M., Dunning, M.J., Speed, D., Lynch, A.G., Samarajiwa, S., Yuan, Y.*, et al.* (2012). The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature *486*, 346-352.

D'Amato, N.C., Ostrander, J.H., Bowie, M.L., Sistrunk, C., Borowsky, A., Cardiff, R.D., Bell, K., Young, L.J., Simin, K., Bachelder, R.E.*, et al.* (2012). Evidence for phenotypic plasticity in aggressive triple-negative breast cancer: human biology is recapitulated by a novel model system. PloS one *7*, e45684.

De Palma, M., and Hanahan, D. (2012). The biology of personalized cancer medicine: facing individual complexities underlying hallmark capabilities. Molecular oncology *6*, 111-127.

de Visser, K.E., Ciampricotti, M., Michalak, E.M., Tan, D.W., Speksnijder, E.N., Hau, C.S., Clevers, H., Barker, N., and Jonkers, J. (2012). Developmental stage-specific contribution of LGR5(+) cells to basal and luminal epithelial lineages in the postnatal mammary gland. The Journal of pathology *228*, 300-309.

Deome, K.B., Faulkin, L.J., Jr., Bern, H.A., and Blair, P.B. (1959). Development of mammary tumors from hyperplastic alveolar nodules transplanted into gland-free mammary fat pads of female C3H mice. Cancer research *19*, 515-520.

DeRose, Y.S., Wang, G., Lin, Y.C., Bernard, P.S., Buys, S.S., Ebbert, M.T., Factor, R., Matsen, C., Milash, B.A., Nelson, E.*, et al.* (2011). Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes. Nature medicine *17*, 1514-1520.

Desmedt, C., Fumagalli, D., Pietri, E., Zoppoli, G., Brown, D., Nik-Zainal, S., Gundem, G., Rothe, F., Majjaj, S., Garuti, A.*, et al.* (2015). Uncovering the genomic heterogeneity of multifocal breast cancer. The Journal of pathology *236*, 457-466.

Ding, L., Ellis, M.J., Li, S., Larson, D.E., Chen, K., Wallis, J.W., Harris, C.C., McLellan, M.D., Fulton, R.S., Fulton, L.L.*, et al.* (2010). Genome remodelling in a basal-like breast cancer metastasis and xenograft. Nature *464*, 999-1005.

Dourdin, N., Schade, B., Lesurf, R., Hallett, M., Munn, R.J., Cardiff, R.D., and Muller, W.J. (2008). Phosphatase and tensin homologue deleted on chromosome 10 deficiency accelerates tumor induction in a mouse model of ErbB-2 mammary tumorigenesis. Cancer research *68*, 2122-2131.

Dumont, N., Liu, B., Defilippis, R.A., Chang, H., Rabban, J.T., Karnezis, A.N., Tjoe, J.A., Marx, J., Parvin, B., and Tlsty, T.D. (2013). Breast fibroblasts modulate early dissemination, tumorigenesis, and metastasis through alteration of extracellular matrix characteristics. Neoplasia *15*, 249-262.

References

Dunlap, J., Le, C., Shukla, A., Patterson, J., Presnell, A., Heinrich, M.C., Corless, C.L., and Troxell, M.L. (2010). Phosphatidylinositol-3-kinase and AKT1 mutations occur early in breast carcinoma. Breast cancer research and treatment *120*, 409-418.

Dupont Jensen, J., Laenkholm, A.V., Knoop, A., Ewertz, M., Bandaru, R., Liu, W., Hackl, W., Barrett, J.C., and Gardner, H. (2011). PIK3CA mutations may be discordant between primary and corresponding metastatic disease in breast cancer. Clinical cancer research : an official journal of the American Association for Cancer Research *17*, 667-677.

Eichhorn, P.J., Gili, M., Scaltriti, M., Serra, V., Guzman, M., Nijkamp, W., Beijersbergen, R.L., Valero, V., Seoane, J., Bernards, R.*, et al.* (2008). Phosphatidylinositol 3-kinase hyperactivation results in lapatinib resistance that is reversed by the mTOR/phosphatidylinositol 3-kinase inhibitor NVP-BEZ235. Cancer research *68*, 9221-9230.

Eirew, P., Steif, A., Khattra, J., Ha, G., Yap, D., Farahani, H., Gelmon, K., Chia, S., Mar, C., Wan, A.*, et al.* (2015). Dynamics of genomic clones in breast cancer patient xenografts at single-cell resolution. Nature *518*, 422-426.

Ellis, M.J., Ding, L., Shen, D., Luo, J., Suman, V.J., Wallis, J.W., Van Tine, B.A., Hoog, J., Goiffon, R.J., Goldstein, T.C.*, et al.* (2012). Whole-genome analysis informs breast cancer response to aromatase inhibition. Nature *486*, 353-360.

Engelman, J.A. (2009). Targeting PI3K signalling in cancer: opportunities, challenges and limitations. Nature reviews. Cancer *9*, 550-562.

Engelman, J.A., Chen, L., Tan, X., Crosby, K., Guimaraes, A.R., Upadhyay, R., Maira, M., McNamara, K., Perera, S.A., Song, Y.*, et al.* (2008). Effective use of PI3K and MEK inhibitors to treat mutant Kras G12D and PIK3CA H1047R murine lung cancers. Nature medicine *14*, 1351-1356.

Engelman, J.A., Luo, J., and Cantley, L.C. (2006). The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism. Nature reviews. Genetics *7*, 606-619.

Faulkin, L.J., Jr., and Deome, K.B. (1960). Regulation of growth and spacing of gland elements in the mammary fat pad of the C3H mouse. Journal of the National Cancer Institute *24*, 953-969.

Gaujoux, R., and Seoighe, C. (2010). A flexible R package for nonnegative matrix factorization. BMC bioinformatics *11*, 367.

Gerlinger, M., Rowan, A.J., Horswell, S., Larkin, J., Endesfelder, D., Gronroos, E., Martinez, P., Matthews, N., Stewart, A., Tarpey, P.*, et al.* (2012). Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. The New England journal of medicine *366*, 883-892.

Ghajar, C.M., Peinado, H., Mori, H., Matei, I.R., Evason, K.J., Brazier, H., Almeida, D., Koller, A., Hajjar, K.A., Stainier, D.Y.*, et al.* (2013). The perivascular niche regulates breast tumour dormancy. Nature cell biology *15*, 807-817.

References

Goubran, H.A., Kotb, R.R., Stakiw, J., Emara, M.E., and Burnouf, T. (2014). Regulation of tumor growth and metastasis: the role of tumor microenvironment. Cancer growth and metastasis *7*, 9-18.

Gunther, E.J., Belka, G.K., Wertheim, G.B., Wang, J., Hartman, J.L., Boxer, R.B., and Chodosh, L.A. (2002). A novel doxycycline-inducible system for the transgenic analysis of mammary gland biology. FASEB journal : official publication of the Federation of American Societies for Experimental Biology *16*, 283-292.

Guo, W., Keckesova, Z., Donaher, J.L., Shibue, T., Tischler, V., Reinhardt, F., Itzkovitz, S., Noske, A., Zurrer-Hardi, U., Bell, G.*, et al.* (2012). Slug and Sox9 cooperatively determine the mammary stem cell state. Cell *148*, 1015-1028.

Gupta, P.B., Fillmore, C.M., Jiang, G., Shapira, S.D., Tao, K., Kuperwasser, C., and Lander, E.S. (2011). Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. Cell *146*, 633-644.

Gusterson, B. (2009). Do 'basal-like' breast cancers really exist? Nature reviews. Cancer *9*, 128-134.

Guy, C.T., Cardiff, R.D., and Muller, W.J. (1992). Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. Molecular and cellular biology *12*, 954-961.

Hao, Y., Wang, C., Cao, B., Hirsch, B.M., Song, J., Markowitz, S.D., Ewing, R.M., Sedwick, D., Liu, L., Zheng, W.*, et al.* (2013). Gain of interaction with IRS1 by p110alpha-helical domain mutants is crucial for their oncogenic functions. Cancer cell *23*, 583-593.

Hennighausen, L., and Robinson, G.W. (2001). Signaling pathways in mammary gland development. Developmental cell *1*, 467-475.

Herschkowitz, J.I., Simin, K., Weigman, V.J., Mikaelian, I., Usary, J., Hu, Z., Rasmussen, K.E., Jones, L.P., Assefnia, S., Chandrasekharan, S.*, et al.* (2007). Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome biology *8*, R76.

Hidalgo, M., Amant, F., Biankin, A.V., Budinska, E., Byrne, A.T., Caldas, C., Clarke, R.B., de Jong, S., Jonkers, J., Maelandsmo, G.M.*, et al.* (2014). Patient-derived xenograft models: an emerging platform for translational cancer research. Cancer discovery *4*, 998-1013.

Hollander, M.C., Blumenthal, G.M., and Dennis, P.A. (2011). PTEN loss in the continuum of common cancers, rare syndromes and mouse models. Nature reviews. Cancer *11*, 289-301.

Hollern, D.P., and Andrechek, E.R. (2014). A genomic analysis of mouse models of breast cancer reveals molecular features of mouse models and relationships to human breast cancer. Breast cancer research : BCR *16*, R59.

References

Hon, W.C., Berndt, A., and Williams, R.L. (2012). Regulation of lipid binding underlies the activation mechanism of class IA PI3-kinases. Oncogene *31*, 3655-3666.

Hu, Y., and Smyth, G.K. (2009). ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. Journal of immunological methods *347*, 70-78.

Ince, T.A., Richardson, A.L., Bell, G.W., Saitoh, M., Godar, S., Karnoub, A.E., Iglehart, J.D., and Weinberg, R.A. (2007). Transformation of different human breast epithelial cell types leads to distinct tumor phenotypes. Cancer cell *12*, 160-170.

Isakoff, S.J., Engelman, J.A., Irie, H.Y., Luo, J., Brachmann, S.M., Pearline, R.V., Cantley, L.C., and Brugge, J.S. (2005). Breast cancer-associated PIK3CA mutations are oncogenic in mammary epithelial cells. Cancer research *65*, 10992-11000.

Jaks, V., Barker, N., Kasper, M., van Es, J.H., Snippert, H.J., Clevers, H., and Toftgard, R. (2008). Lgr5 marks cycling, yet long-lived, hair follicle stem cells. Nature genetics *40*, 1291-1299.

Jean, S., and Kiger, A.A. (2014). Classes of phosphoinositide 3-kinases at a glance. Journal of cell science *127*, 923-928.

Jiang, Z., Deng, T., Jones, R., Li, H., Herschkowitz, J.I., Liu, J.C., Weigman, V.J., Tsao, M.S., Lane, T.F., Perou, C.M.*, et al.* (2010). Rb deletion in mouse mammary progenitors induces luminal-B or basal-like/EMT tumor subtypes depending on p53 status. The Journal of clinical investigation *120*, 3296-3309.

Jonkers, J., Meuwissen, R., van der Gulden, H., Peterse, H., van der Valk, M., and Berns, A. (2001). Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional mouse model for breast cancer. Nature genetics *29*, 418-425.

Joyce, J.A., and Pollard, J.W. (2009). Microenvironmental regulation of metastasis. Nature reviews. Cancer *9*, 239-252.

Kadota, M., Sato, M., Duncan, B., Ooshima, A., Yang, H.H., Diaz-Meyer, N., Gere, S., Kageyama, S., Fukuoka, J., Nagata, T.*, et al.* (2009). Identification of novel gene amplifications in breast cancer and coexistence of gene amplification with an activating mutation of PIK3CA. Cancer research *69*, 7357-7365.

Kalinsky, K., Jacks, L.M., Heguy, A., Patil, S., Drobnjak, M., Bhanot, U.K., Hedvat, C.V., Traina, T.A., Solit, D., Gerald, W.*, et al.* (2009). PIK3CA mutation associates with improved outcome in breast cancer. Clinical cancer research : an official journal of the American Association for Cancer Research *15*, 5049-5059.

Kataoka, Y., Mukohara, T., Shimada, H., Saijo, N., Hirai, M., and Minami, H. (2010). Association between gain-of-function mutations in PIK3CA and resistance to HER2-targeted agents in HER2-amplified breast cancer cell lines. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO *21*, 255-262.

References

Keller, P.J., Arendt, L.M., Skibinski, A., Logvinenko, T., Klebba, I., Dong, S., Smith, A.E., Prat, A., Perou, C.M., Gilmore, H.*, et al.* (2012). Defining the cellular precursors to human breast cancer. Proceedings of the National Academy of Sciences of the United States of America *109*, 2772-2777.

Kendrick, H., Regan, J.L., Magnay, F.A., Grigoriadis, A., Mitsopoulos, C., Zvelebil, M., and Smalley, M.J. (2008). Transcriptome analysis of mammary epithelial subpopulations identifies novel determinants of lineage commitment and cell fate. BMC genomics *9*, 591.

Kinzel, B., Pikiolek, M., Orsini, V., Sprunger, J., Isken, A., Zietzling, S., Desplanches, M., Dubost, V., Breustedt, D., Valdez, R.*, et al.* (2014). Functional roles of Lgr4 and Lgr5 in embryonic gut, kidney and skin development in mice. Developmental biology *390*, 181-190.

Klevebring, D., Rosin, G., Ma, R., Lindberg, J., Czene, K., Kere, J., Fredriksson, I., Bergh, J., and Hartman, J. (2014). Sequencing of breast cancer stem cell populations indicates a dynamic conversion between differentiation states in vivo. Breast cancer research : BCR *16*, R72.

Komaki, K., Sano, N., and Tangoku, A. (2006). Problems in histological grading of malignancy and its clinical significance in patients with operable breast cancer. Breast cancer *13*, 249-253.

Kordon, E.C., and Smith, G.H. (1998). An entire functional mammary gland may comprise the progeny from a single cell. Development *125*, 1921-1930.

Koren, S., and Bentires-Alj, M. (2013). Mouse models of PIK3CA mutations: one mutation initiates heterogeneous mammary tumors. The FEBS journal *280*, 2758-2765.

Koren, S., Reavie, L., Couto, J.P., De Silva, D., Stadler, M.B., Roloff, T., Britschgi, A., Eichlisberger, T., Kohler, H., Aina, O.*, et al.* (2015). PIK3CA(H1047R) induces multipotency and multi-lineage mammary tumours. Nature *525*, 114-118.

Kreso, A., and Dick, J.E. (2014). Evolution of the cancer stem cell model. Cell stem cell *14*, 275-291.

Lai, Y.L., Mau, B.L., Cheng, W.H., Chen, H.M., Chiu, H.H., and Tzen, C.Y. (2008). PIK3CA exon 20 mutation is independently associated with a poor prognosis in breast cancer patients. Annals of surgical oncology *15*, 1064-1069.

Leroy, C., Amante, R.J., and Bentires-Alj, M. (2014). Anticipating mechanisms of resistance to PI3K inhibition in breast cancer: a challenge in the era of precision medicine. Biochemical Society transactions *42*, 733-741.

Levine, D.A., Bogomolniy, F., Yee, C.J., Lash, A., Barakat, R.R., Borgen, P.I., and Boyd, J. (2005). Frequent mutation of the PIK3CA gene in ovarian and breast cancers. Clinical cancer research : an official journal of the American Association for Cancer Research *11*, 2875-2878.

References

Li, G., Robinson, G.W., Lesche, R., Martinez-Diaz, H., Jiang, Z., Rozengurt, N., Wagner, K.U., Wu, D.C., Lane, T.F., Liu, X.*, et al.* (2002). Conditional loss of PTEN leads to precocious development and neoplasia in the mammary gland. Development *129*, 4159-4170.

Li, H., Zhu, R., Wang, L., Zhu, T., Li, Q., Chen, Q., Wang, H., and Zhu, H. (2010). PIK3CA mutations mostly begin to develop in ductal carcinoma of the breast. Experimental and molecular pathology *88*, 150-155.

Li, S.Y., Rong, M., Grieu, F., and Iacopetta, B. (2006). PIK3CA mutations in breast cancer are associated with poor outcome. Breast cancer research and treatment *96*, 91-95.

Lim, E., Vaillant, F., Wu, D., Forrest, N.C., Pal, B., Hart, A.H., Asselin-Labat, M.L., Gyorki, D.E., Ward, T., Partanen, A.*, et al.* (2009). Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. Nature medicine *15*, 907-913.

Lim, E., Wu, D., Pal, B., Bouras, T., Asselin-Labat, M.L., Vaillant, F., Yagita, H., Lindeman, G.J., Smyth, G.K., and Visvader, J.E. (2010). Transcriptome analyses of mouse and human mammary cell subpopulations reveal multiple conserved genes and pathways. Breast cancer research : BCR *12*, R21.

Lin, S.C., Lee, K.F., Nikitin, A.Y., Hilsenbeck, S.G., Cardiff, R.D., Li, A., Kang, K.W., Frank, S.A., Lee, W.H., and Lee, E.Y. (2004). Somatic mutation of p53 leads to estrogen receptor alpha-positive and -negative mouse mammary tumors with high frequency of metastasis. Cancer research *64*, 3525-3532.

Liu, P., Cheng, H., Santiago, S., Raeder, M., Zhang, F., Isabella, A., Yang, J., Semaan, D.J., Chen, C., Fox, E.A.*, et al.* (2011). Oncogenic PIK3CA-driven mammary tumors frequently recur via PI3K pathway-dependent and PI3K pathway-independent mechanisms. Nature medicine *17*, 1116-1120.

Liu, X., Holstege, H., van der Gulden, H., Treur-Mulder, M., Zevenhoven, J., Velds, A., Kerkhoven, R.M., van Vliet, M.H., Wessels, L.F., Peterse, J.L.*, et al.* (2007). Somatic loss of BRCA1 and p53 in mice induces mammary tumors with features of human BRCA1-mutated basal-like breast cancer. Proceedings of the National Academy of Sciences of the United States of America *104*, 12111-12116.

Macias, H., and Hinck, L. (2012). Mammary gland development. Wiley interdisciplinary reviews. Developmental biology *1*, 533-557.

Mandelker, D., Gabelli, S.B., Schmidt-Kittler, O., Zhu, J., Cheong, I., Huang, C.H., Kinzler, K.W., Vogelstein, B., and Amzel, L.M. (2009). A frequent kinase domain mutation that changes the interaction between PI3Kalpha and the membrane. Proceedings of the National Academy of Sciences of the United States of America *106*, 16996-17001.

Mangone, F.R., Bobrovnitchaia, I.G., Salaorni, S., Manuli, E., and Nagai, M.A. (2012). PIK3CA exon 20 mutations are associated with poor prognosis in breast cancer patients. Clinics *67*, 1285-1290.

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

References

Marusyk, A., and Polyak, K. (2010). Tumor heterogeneity: causes and consequences. Biochimica et biophysica acta *1805*, 105-117.

Marusyk, A., Tabassum, D.P., Altrock, P.M., Almendro, V., Michor, F., and Polyak, K. (2014). Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity. Nature *514*, 54-58.

Maruyama, N., Miyoshi, Y., Taguchi, T., Tamaki, Y., Monden, M., and Noguchi, S. (2007). Clinicopathologic analysis of breast cancers with PIK3CA mutations in Japanese women. Clinical cancer research : an official journal of the American Association for Cancer Research *13*, 408-414.

Meier-Abt, F., Milani, E., Roloff, T., Brinkhaus, H., Duss, S., Meyer, D.S., Klebba, I., Balwierz, P.J., van Nimwegen, E., and Bentires-Alj, M. (2013). Parity induces differentiation and reduces Wnt/Notch signaling ratio and proliferation potential of basal stem/progenitor cells isolated from mouse mammary epithelium. Breast cancer research : BCR *15*, R36.

Melchor, L., Molyneux, G., Mackay, A., Magnay, F.A., Atienza, M., Kendrick, H., Nava-Rodrigues, D., Lopez-Garcia, M.A., Milanezi, F., Greenow, K.*, et al.* (2014). Identification of cellular and genetic drivers of breast cancer heterogeneity in genetically engineered mouse tumour models. The Journal of pathology *233*, 124-137.

Meyer, D.S., and Bentires-Alj, M. (2010). Can phosphatidylinositol 3-kinase/mammalian target of rapamycin inhibition ERase them all? Breast cancer research : BCR *12*, 315.

Meyer, D.S., Brinkhaus, H., Muller, U., Muller, M., Cardiff, R.D., and Bentires-Alj, M. (2011). Luminal expression of PIK3CA mutant H1047R in the mammary gland induces heterogeneous tumors. Cancer research *71*, 4344-4351.

Meyer, D.S., Koren, S., Leroy, C., Brinkhaus, H., Muller, U., Klebba, I., Muller, M., Cardiff, R.D., and Bentires-Alj, M. (2013). Expression of PIK3CA mutant E545K in the mammary gland induces heterogeneous tumors but is less potent than mutant H1047R. Oncogenesis *2*, e74.

Miled, N., Yan, Y., Hon, W.C., Perisic, O., Zvelebil, M., Inbar, Y., Schneidman-Duhovny, D., Wolfson, H.J., Backer, J.M., and Williams, R.L. (2007). Mechanism of two classes of cancer mutations in the phosphoinositide 3-kinase catalytic subunit. Science *317*, 239-242.

Miller, T.W. (2012). Initiating breast cancer by PIK3CA mutation. Breast cancer research : BCR *14*, 301.

Miller, T.W., Forbes, J.T., Shah, C., Wyatt, S.K., Manning, H.C., Olivares, M.G., Sanchez, V., Dugger, T.C., de Matos Granja, N., Narasanna, A.*, et al.* (2009). Inhibition of mammalian target of rapamycin is required for optimal antitumor effect of HER2 inhibitors against HER2-overexpressing cancer cells. Clinical cancer research : an official journal of the American Association for Cancer Research *15*, 7266-7276.

References

Miller, T.W., Rexer, B.N., Garrett, J.T., and Arteaga, C.L. (2011). Mutations in the phosphatidylinositol 3-kinase pathway: role in tumor progression and therapeutic implications in breast cancer. Breast cancer research : BCR *13*, 224.

Miron, A., Varadi, M., Carrasco, D., Li, H., Luongo, L., Kim, H.J., Park, S.Y., Cho, E.Y., Lewis, G., Kehoe, S*., et al.* (2010). PIK3CA mutations in in situ and invasive breast carcinomas. Cancer research *70*, 5674-5678.

Molyneux, G., Geyer, F.C., Magnay, F.A., McCarthy, A., Kendrick, H., Natrajan, R., Mackay, A., Grigoriadis, A., Tutt, A., Ashworth, A*., et al.* (2010). BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. Cell stem cell *7*, 403-417.

Muller, W.J., Sinn, E., Pattengale, P.K., Wallace, R., and Leder, P. (1988). Single-step induction of mammary adenocarcinoma in transgenic mice bearing the activated c-neu oncogene. Cell *54*, 105-115.

Nardella, C., Lunardi, A., Patnaik, A., Cantley, L.C., and Pandolfi, P.P. (2011). The APL paradigm and the "co-clinical trial" project. Cancer discovery *1*, 108-116.

Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., McIndoo, J., Cook, K., Stepansky, A., Levy, D., Esposito, D*., et al.* (2011). Tumour evolution inferred by single-cell sequencing. Nature *472*, 90-94.

Neve, R.M., Chin, K., Fridlyand, J., Yeh, J., Baehner, F.L., Fevr, T., Clark, L., Bayani, N., Coppe, J.P., Tong, F*., et al.* (2006). A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer cell *10*, 515-527.

Nik-Zainal, S., Alexandrov, L.B., Wedge, D.C., Van Loo, P., Greenman, C.D., Raine, K., Jones, D., Hinton, J., Marshall, J., Stebbings, L.A*., et al.* (2012a). Mutational processes molding the genomes of 21 breast cancers. Cell *149*, 979-993.

Nik-Zainal, S., Van Loo, P., Wedge, D.C., Alexandrov, L.B., Greenman, C.D., Lau, K.W., Raine, K., Jones, D., Marshall, J., Ramakrishna, M*., et al.* (2012b). The life history of 21 breast cancers. Cell *149*, 994-1007.

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

Pal, B., Bouras, T., Shi, W., Vaillant, F., Sheridan, J.M., Fu, N., Breslin, K., Jiang, K., Ritchie, M.E., Young, M*., et al.* (2013). Global changes in the mammary epigenome are induced by hormonal cues and coordinated by Ezh2. Cell reports *3*, 411-426.

Perez-Tenorio, G., Alkhori, L., Olsson, B., Waltersson, M.A., Nordenskjold, B., Rutqvist, L.E., Skoog, L., and Stal, O. (2007). PIK3CA mutations and PTEN loss correlate with similar prognostic factors and are not mutually exclusive in breast cancer. Clinical cancer research : an official journal of the American Association for Cancer Research *13*, 3577-3584.

Perou, C.M., Sorlie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., Rees, C.A., Pollack, J.R., Ross, D.T., Johnsen, H., Akslen, L.A.*, et al.* (2000). Molecular portraits of human breast tumours. Nature *406*, 747-752.

Petersen, O.W., and Polyak, K. (2010). Stem cells in the human breast. Cold Spring Harbor perspectives in biology *2*, a003160.

Pfefferle, A.D., Herschkowitz, J.I., Usary, J., Harrell, J.C., Spike, B.T., Adams, J.R., Torres-Arzayus, M.I., Brown, M., Egan, S.E., Wahl, G.M.*, et al.* (2013). Transcriptomic classification of genetically engineered mouse models of breast cancer identifies human subtype counterparts. Genome biology *14*, R125.

Phillips, S., Prat, A., Sedic, M., Proia, T., Wronski, A., Mazumdar, S., Skibinski, A., Shirley, S.H., Perou, C.M., Gill, G.*, et al.* (2014). Cell-state transitions regulated by SLUG are critical for tissue regeneration and tumor initiation. Stem cell reports *2*, 633-647.

Plaks, V., Brenot, A., Lawson, D.A., Linnemann, J.R., Van Kappel, E.C., Wong, K.C., de Sauvage, F., Klein, O.D., and Werb, Z. (2013). Lgr5-expressing cells are sufficient and necessary for postnatal mammary gland organogenesis. Cell reports *3*, 70-78.

Prat, A., Parker, J.S., Karginova, O., Fan, C., Livasy, C., Herschkowitz, J.I., He, X., and Perou, C.M. (2010). Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast cancer research : BCR *12*, R68.

Prater, M.D., Petit, V., Alasdair Russell, I., Giraddi, R.R., Shehata, M., Menon, S., Schulte, R., Kalajzic, I., Rath, N., Olson, M.F.*, et al.* (2014). Mammary stem cells have myoepithelial cell properties. Nature cell biology *16*, 942-950, 941-947.

Quail, D.F., and Joyce, J.A. (2013). Microenvironmental regulation of tumor progression and metastasis. Nature medicine *19*, 1423-1437.

Quereda, V., Martinalbo, J., Dubus, P., Carnero, A., and Malumbres, M. (2007). Genetic cooperation between p21Cip1 and INK4 inhibitors in cellular senescence and tumor suppression. Oncogene *26*, 7665-7674.

Ramos, P., and Bentires-Alj, M. (2015). Mechanism-based cancer therapy: resistance to therapy, therapy for resistance. Oncogene *34*, 3617-3626.

Renner, O., Blanco-Aparicio, C., Grassow, M., Canamero, M., Leal, J.F., and Carnero, A. (2008). Activation of phosphatidylinositol 3-kinase by membrane localization of p110alpha predisposes mammary glands to neoplastic transformation. Cancer research *68*, 9643-9653.

Reyal, F., Guyader, C., Decraene, C., Lucchesi, C., Auger, N., Assayag, F., De Plater, L., Gentien, D., Poupon, M.F., Cottu, P.*, et al.* (2012). Molecular profiling of patient-derived breast cancer xenografts. Breast cancer research : BCR *14*, R11.

References

Rijnkels, M., Kabotyanski, E., Montazer-Torbati, M.B., Hue Beauvais, C., Vassetzky, Y., Rosen, J.M., and Devinoy, E. (2010). The epigenetic landscape of mammary gland development and functional differentiation. Journal of mammary gland biology and neoplasia *15*, 85-100.

Rios, A.C., Fu, N.Y., Lindeman, G.J., and Visvader, J.E. (2014). In situ identification of bipotent stem cells in the mammary gland. Nature *506*, 322-327.

Saal, L.H., Holm, K., Maurer, M., Memeo, L., Su, T., Wang, X., Yu, J.S., Malmstrom, P.O., Mansukhani, M., Enoksson, J.*, et al.* (2005). PIK3CA mutations correlate with hormone receptors, node metastasis, and ERBB2, and are mutually exclusive with PTEN loss in human breast carcinoma. Cancer research *65*, 2554-2559.

Samuels, Y., Wang, Z., Bardelli, A., Silliman, N., Ptak, J., Szabo, S., Yan, H., Gazdar, A., Powell, S.M., Riggins, G.J.*, et al.* (2004). High frequency of mutations of the PIK3CA gene in human cancers. Science *304*, 554.

Schepers, A.G., Snippert, H.J., Stange, D.E., van den Born, M., van Es, J.H., van de Wetering, M., and Clevers, H. (2012). Lineage tracing reveals Lgr5+ stem cell activity in mouse intestinal adenomas. Science *337*, 730-735.

Schwitalla, S., Fingerle, A.A., Cammareri, P., Nebelsiek, T., Goktuna, S.I., Ziegler, P.K., Canli, O., Heijmans, J., Huels, D.J., Moreaux, G.*, et al.* (2013). Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem-cell-like properties. Cell *152*, 25-38.

Shackleton, M., Vaillant, F., Simpson, K.J., Stingl, J., Smyth, G.K., Asselin-Labat, M.L., Wu, L., Lindeman, G.J., and Visvader, J.E. (2006). Generation of a functional mammary gland from a single stem cell. Nature *439*, 84-88.

Shah, S.P., Morin, R.D., Khattra, J., Prentice, L., Pugh, T., Burleigh, A., Delaney, A., Gelmon, K., Guliany, R., Senz, J.*, et al.* (2009). Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. Nature *461*, 809-813.

Shah, S.P., Roth, A., Goya, R., Oloumi, A., Ha, G., Zhao, Y., Turashvili, G., Ding, J., Tse, K., Haffari, G.*, et al.* (2012). The clonal and mutational evolution spectrum of primary triple-negative breast cancers. Nature *486*, 395-399.

Shen, H., and Laird, P.W. (2013). Interplay between the cancer genome and epigenome. Cell *153*, 38-55.

Sinn, E., Muller, W., Pattengale, P., Tepler, I., Wallace, R., and Leder, P. (1987). Coexpression of MMTV/v-Ha-ras and MMTV/c-myc genes in transgenic mice: synergistic action of oncogenes in vivo. Cell *49*, 465-475.

Skibinski, A., and Kuperwasser, C. (2015). The origin of breast tumor heterogeneity. Oncogene *34*, 5309-5316.

References

Sleeman, K.E., Kendrick, H., Ashworth, A., Isacke, C.M., and Smalley, M.J. (2006). CD24 staining of mouse mammary gland cells defines luminal epithelial, myoepithelial/basal and non-epithelial cells. Breast cancer research : BCR *8*, R7.

Sleeman, K.E., Kendrick, H., Robertson, D., Isacke, C.M., Ashworth, A., and Smalley, M.J. (2007). Dissociation of estrogen receptor expression and in vivo stem cell activity in the mammary gland. The Journal of cell biology *176*, 19-26.

Sommer, C., Strähle, C., Köthe, U., and Hamprecht, F.A. (2011). Ilastik: Interactive Learning and Segmentation Toolkit. Proceedings, 230-233.

Soriano, P. (1999). Generalized lacZ expression with the ROSA26 Cre reporter strain. Nature genetics *21*, 70-71.

Sorlie, T., Perou, C.M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S.*, et al.* (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proceedings of the National Academy of Sciences of the United States of America *98*, 10869-10874.

Sorlie, T., Tibshirani, R., Parker, J., Hastie, T., Marron, J.S., Nobel, A., Deng, S., Johnsen, H., Pesich, R., Geisler, S.*, et al.* (2003). Repeated observation of breast tumor subtypes in independent gene expression data sets. Proceedings of the National Academy of Sciences of the United States of America *100*, 8418-8423.

Sotillo, R., Dubus, P., Martin, J., de la Cueva, E., Ortega, S., Malumbres, M., and Barbacid, M. (2001). Wide spectrum of tumors in knock-in mice carrying a Cdk4 protein insensitive to INK4 inhibitors. The EMBO journal *20*, 6637-6647.

Spike, B.T., Engle, D.D., Lin, J.C., Cheung, S.K., La, J., and Wahl, G.M. (2012). A mammary stem cell population identified and characterized in late embryogenesis reveals similarities to human breast cancer. Cell stem cell *10*, 183-197.

Srinivas, S., Watanabe, T., Lin, C.S., William, C.M., Tanabe, Y., Jessell, T.M., and Costantini, F. (2001). Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. BMC developmental biology *1*, 4.

Stambolic, V., Tsao, M.S., Macpherson, D., Suzuki, A., Chapman, W.B., and Mak, T.W. (2000). High incidence of breast and endometrial neoplasia resembling human Cowden syndrome in pten+/- mice. Cancer research *60*, 3605-3611.

Stefansson, O.A., and Esteller, M. (2013). Epigenetic modifications in breast cancer and their role in personalized medicine. The American journal of pathology *183*, 1052-1063.

Stephens, P.J., Tarpey, P.S., Davies, H., Van Loo, P., Greenman, C., Wedge, D.C., Nik-Zainal, S., Martin, S., Varela, I., Bignell, G.R.*, et al.* (2012). The landscape of cancer genes and mutational processes in breast cancer. Nature *486*, 400-404.

Sternlicht, M.D. (2006). Key stages in mammary gland development: the cues that regulate ductal branching morphogenesis. Breast cancer research : BCR *8*, 201.

Stewart, T.A., Pattengale, P.K., and Leder, P. (1984). Spontaneous mammary adenocarcinomas in transgenic mice that carry and express MTV/myc fusion genes. Cell *38*, 627-637.

Stingl, J., Eirew, P., Ricketson, I., Shackleton, M., Vaillant, F., Choi, D., Li, H.I., and Eaves, C.J. (2006). Purification and unique properties of mammary epithelial stem cells. Nature *439*, 993-997.

Tabassum, D.P., and Polyak, K. (2015). Tumorigenesis: it takes a village. Nature reviews. Cancer *15*, 473-483.

Tao, L., van Bragt, M.P., Laudadio, E., and Li, Z. (2014). Lineage tracing of mammary epithelial cells using cell-type-specific cre-expressing adenoviruses. Stem cell reports *2*, 770-779.

Tao, L., van Bragt, M.P., and Li, Z. (2015). A Long-Lived Luminal Subpopulation Enriched with Alveolar Progenitors Serves as Cellular Origin of Heterogeneous Mammary Tumors. Stem cell reports *5*, 60-74.

Tchorz, J.S., Kinter, J., Muller, M., Tornillo, L., Heim, M.H., and Bettler, B. (2009). Notch2 signaling promotes biliary epithelial cell fate specification and tubulogenesis during bile duct development in mice. Hepatology *50*, 871-879.

Tikoo, A., Roh, V., Montgomery, K.G., Ivetac, I., Waring, P., Pelzer, R., Hare, L., Shackleton, M., Humbert, P., and Phillips, W.A. (2012). Physiological levels of Pik3ca(H1047R) mutation in the mouse mammary gland results in ductal hyperplasia and formation of ERalpha-positive tumors. PloS one *7*, e36924.

Torre, L.A., Bray, F., Siegel, R.L., Ferlay, J., Lortet-Tieulent, J., and Jemal, A. (2015). Global cancer statistics, 2012. CA: a cancer journal for clinicians *65*, 87-108.

Troester, M.A., Hoadley, K.A., Sorlie, T., Herbert, B.S., Borresen-Dale, A.L., Lonning, P.E., Shay, J.W., Kaufmann, W.K., and Perou, C.M. (2004). Cell-type-specific responses to chemotherapeutics in breast cancer. Cancer research *64*, 4218-4226.

Uchida, H., Yamazaki, K., Fukuma, M., Yamada, T., Hayashida, T., Hasegawa, H., Kitajima, M., Kitagawa, Y., and Sakamoto, M. (2010). Overexpression of leucine-rich repeat-containing G protein-coupled receptor 5 in colorectal cancer. Cancer science *101*, 1731-1737.

van Amerongen, R., Bowman, A.N., and Nusse, R. (2012). Developmental stage and time dictate the fate of Wnt/beta-catenin-responsive stem cells in the mammary gland. Cell stem cell *11*, 387-400.

Van Keymeulen, A., Lee, M.Y., Ousset, M., Brohee, S., Rorive, S., Giraddi, R.R., Wuidart, A., Bouvencourt, G., Dubois, C., Salmon, I*., et al.* (2015). Reactivation of multipotency by oncogenic PIK3CA induces breast tumour heterogeneity. Nature *525*, 119-123.

References

Van Keymeulen, A., Rocha, A.S., Ousset, M., Beck, B., Bouvencourt, G., Rock, J., Sharma, N., Dekoninck, S., and Blanpain, C. (2011). Distinct stem cells contribute to mammary gland development and maintenance. Nature *479*, 189-193.

Viale, G. (2012). The current state of breast cancer classification. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO *23 Suppl 10*, x207-210.

Visvader, J.E., and Stingl, J. (2014). Mammary stem cells and the differentiation hierarchy: current status and perspectives. Genes & development *28*, 1143-1158.

Voskoglou-Nomikos, T., Pater, J.L., and Seymour, L. (2003). Clinical predictive value of the in vitro cell line, human xenograft, and mouse allograft preclinical cancer models. Clinical cancer research : an official journal of the American Association for Cancer Research *9*, 4227-4239.

Wagner, K.U., McAllister, K., Ward, T., Davis, B., Wiseman, R., and Hennighausen, L. (2001). Spatial and temporal expression of the Cre gene under the control of the MMTV-LTR in different lines of transgenic mice. Transgenic research *10*, 545-553.

Wagner, K.U., Wall, R.J., St-Onge, L., Gruss, P., Wynshaw-Boris, A., Garrett, L., Li, M., Furth, P.A., and Hennighausen, L. (1997). Cre-mediated gene deletion in the mammary gland. Nucleic acids research *25*, 4323-4330.

Wang, D., Cai, C., Dong, X., Yu, Q.C., Zhang, X.O., Yang, L., and Zeng, Y.A. (2015). Identification of multipotent mammary stem cells by protein C receptor expression. Nature *517*, 81-84.

Wang, Z.A., Mitrofanova, A., Bergren, S.K., Abate-Shen, C., Cardiff, R.D., Califano, A., and Shen, M.M. (2013). Lineage analysis of basal epithelial cells reveals their unexpected plasticity and supports a cell-of-origin model for prostate cancer heterogeneity. Nature cell biology *15*, 274-283.

Whittle, J.R., Lewis, M.T., Lindeman, G.J., and Visvader, J.E. (2015). Patient-derived xenograft models of breast cancer and their predictive power. Breast cancer research : BCR *17*, 17.

Wintermantel, T.M., Mayer, A.K., Schutz, G., and Greiner, E.F. (2002). Targeting mammary epithelial cells using a bacterial artificial chromosome. Genesis *33*, 125-130.

Wu, G., Xing, M., Mambo, E., Huang, X., Liu, J., Guo, Z., Chatterjee, A., Goldenberg, D., Gollin, S.M., Sukumar, S.*, et al.* (2005). Somatic mutation and gain of copy number of PIK3CA in human breast cancer. Breast cancer research : BCR *7*, R609-616.

Yates, L.R., Gerstung, M., Knappskog, S., Desmedt, C., Gundem, G., Van Loo, P., Aas, T., Alexandrov, L.B., Larsimont, D., Davies, H.*, et al.* (2015). Subclonal diversification of primary breast cancer revealed by multiregion sequencing. Nature medicine *21*, 751-759.

Yuan, T.L., and Cantley, L.C. (2008). PI3K pathway alterations in cancer: variations on a theme. Oncogene *27*, 5497-5510.

Yuan, W., Stawiski, E., Janakiraman, V., Chan, E., Durinck, S., Edgar, K.A., Kljavin, N.M., Rivers, C.S., Gnad, F., Roose-Girma, M.*, et al.* (2013). Conditional activation of Pik3ca(H1047R) in a knock-in mouse model promotes mammary tumorigenesis and emergence of mutations. Oncogene *32*, 318-326.

Zardavas, D., Baselga, J., and Piccart, M. (2013). Emerging targeted agents in metastatic breast cancer. Nature reviews. Clinical oncology *10*, 191-210.

Zhang, J., Grindley, J.C., Yin, T., Jayasinghe, S., He, X.C., Ross, J.T., Haug, J.S., Rupp, D., Porter-Westpfahl, K.S., Wiedemann, L.M.*, et al.* (2006). PTEN maintains haematopoietic stem cells and acts in lineage choice and leukaemia prevention. Nature *441*, 518-522.

Zhang, X., Claerhout, S., Prat, A., Dobrolecki, L.E., Petrovic, I., Lai, Q., Landis, M.D., Wiechmann, L., Schiff, R., Giuliano, M.*, et al.* (2013). A renewable tissue resource of phenotypically stable, biologically and ethnically diverse, patient-derived human breast cancer xenograft models. Cancer research *73*, 4885-4897.

Zhao, J.J., Liu, Z., Wang, L., Shin, E., Loda, M.F., and Roberts, T.M. (2005). The oncogenic properties of mutant p110alpha and p110beta phosphatidylinositol 3-kinases in human mammary epithelial cells. Proceedings of the National Academy of Sciences of the United States of America *102*, 18443-18448.

Zhao, L., and Vogt, P.K. (2008a). Class I PI3K in oncogenic cellular transformation. Oncogene *27*, 5486-5496.

Zhao, L., and Vogt, P.K. (2008b). Helical domain and kinase domain mutations in p110alpha of phosphatidylinositol 3-kinase induce gain of function by different mechanisms. Proceedings of the National Academy of Sciences of the United States of America *105*, 2652-2657.

Zuo, T., Liu, T.M., Lan, X., Weng, Y.I., Shen, R., Gu, F., Huang, Y.W., Liyanarachchi, S., Deatherage, D.E., Hsu, P.Y.*, et al.* (2011). Epigenetic silencing mediated through activated PI3K/AKT signaling in breast cancer. Cancer research *71*, 1752-1762.

9 | Acknowledgements

First, I would like to thank Momo for giving me the opportunity to work on such an exciting project and for all the support especially in stressful times during my PhD. I further would like to thank Nancy and Matt for being in my thesis committee and for giving me advice and support.

Big thanks go to all the 31 current and past Bentires lab members I worked with over the last 4 years: Adrian Britschgi, Joana Silva, Pedro Ramos, Milan Obradovic, Romain Amante, Duvini De Silva, Tobias Eichlisberger, Marie-May Coissieux, Cedric Leroy, Ana Correia, Linsey Reavie, Ryoko Okamoto, Markus Ackerknecht, Federica Zilli, Milica Vulin, Sandro Bruno, Atul Sethi, Nina Sausgruber, Laura Bonapace, Astrid Pentz, Heike Brinkhaus, Dominique Meyer, Emanuela Milani, Stephan Duss, Nicola Aceto, Ina Klebba, Sophie Sarret, Urs Müller, Jeff Wyckoff, Fabienne Meier-Abt and Karen Cornille. You guys helped me with numerous advices and critical discussions. I would not be where I am today without all your input and help. I appreciate all the hours you spent with me in the lab, at the Rhein, bars and clubs not only discussing science but also making my PhD enjoyable and full of good memories. We had millions of good jokes and laughs and I wish that we stay in touch.

I would like to acknowledge the FMI facilities for all the expertise they provide. Particularly, I would like to thank Hubertus Kohler, Sandrine Bichet, Augustyn Bogucki, Michael Stadler, Tim Roloff, Stephane Thiry, Laurent Gelman, Raphael Thierry, Steven Bourke and everybody from the animal facility. I really appreciated the great help in times of need.

I thank Robert D. Cardiff and Olulanu Aina for the expertise on mouse pathology.

I am thankful for the people who provided us with mouse models: Blanpain Lab (IRIBHM Brussels), B. Kinzel and J. Tchorz (NIBR) and Roska Lab (FMI).

Last but not least I would like to express my deepest gratitude to my friends in and outside the FMI and my family. Special thanks go to my best buddies Joana and Adrian for all their support and good spirit. Huge thanks go to my love Michi and my Mama for all the mental support that constantly helped me during my PhD and especially during the preparation of the manuscript.

<div align="center">

THE END

THANK YOU ALL SO MUCH

</div>

10 | Appendices

10.1 Abbreviations

3D	Three-dimensional
APC	Adenomatous polyposis coli
CAF	Cancer-associated fibroblast
CAGs	Modified chicken β-actin promoter
CD24	Cluster of differentiation 24/heat stable antigen CD24
CD29	Cluster of differentiation 29/Integrin beta-1
CD45	Cluster of differentiation 45/Protein tyrosine phosphatase, receptor type C
CD49f	Cluster of differentiation 49f/Integrin alpha-6
CD61	Cluster of differentiation 29/Integrin beta-3
CDK4	Cyclin-dependent kinase 4
Cre-ERT2	Tamoxifen inducible activation of Cre-recombinase
CSC	Cancer stem cell
Ctrl	Control
DAPI	4′,6-Diamidin-2-phenylindol
EGFP	Enhanced Green fluorescent protein
EMT	Epithelial-mesenchymal transition
EpCAM	Epithelial cell adhesion molecule
ER	Estrogen receptor
ERK	Extracellular-signal-regulated kinase
ES cell	Embryonic stem cell
EZH2	Enhancer of zeste homolog 2
FACS	Fluorescence-activated cell sorting

GPCR	G-protein-coupled receptor
GTP	Guanosine-5'-triphosphate
h	Hour
HA	Hemagglutinin
HER2	Human epidermal growth factor receptor 2
HMEC	Human mammary epithelial cell
ICD	Invasive ductal carcinoma
IRES	Internal ribosome entry site
IRS	Insulin receptor substrate
K14	Keratin 14
K5	Keratin 5
K8	Keratin 8
Lgr5	Leucine-rich repeat-containing G-protein coupled receptor 5
LSD1	Lysin-specific demethylase 1
MaSC	Mammary stem cell
MMTV	Mouse mammary tumor virus
NS	Not significant
PDK1	Phosphoinositide-dependent kinase 1
PDX	Patient-derived xenograft
PGK	Phosphoglycerate kinase
PI3K	Phosphoinositol 3-kinase
PI-4-P	Phosphatidylinositol-4-phosphate
PIP2	Phosphatidylinositol 4,5-bisphosphate
PIP3	Phosphatidylinositol (3,4,5)-trisphosphate
PR	Progesterone receptor

PTEN	Phosphatase and Tensin homolog
PyMT	Polyomavirus middle T
RMCE	Recombinase-mediated cassette exchange
RTK	Receptor tyrosine kinase
rtTA	Tetracycline transactivator protein
s.d.	Standard deviation
s.e.m.	Standard error of the mean
Sca1	Stem cell antigen 1
SMA	Smooth muscle actin
Stat3	Signal transducer and activator of transcription 3
TCGA	The cancer genome atlas
WAP	Whey acidic protein
Wnt1	Wingless-type member 1
ZEB1	Zinc finger E-box-binding homeobox 1

10.2 List of figures

10.3 List of tables

Introduction

10.4 Published manuscripts

Mouse models of *PIK3CA* mutations: one mutation initiates heterogeneous mammary tumors.

<u>Koren S</u>, Bentires-Alj M. **FEBS J. 2013** Jun;280(12):2758-65. doi: 10.1111/febs.12175. Epub 2013 Mar 1. Review.

Expression of *PIK3CA* mutant E545K in the mammary gland induces heterogeneous tumors but is less potent than mutant H1047R.

Meyer DS*, **<u>Koren S*</u>**, Leroy C, Brinkhaus H, Müller U, Klebba I, Müller M, Cardiff RD, **Bentires-Alj** M. **Oncogenesis. 2013** Sep 30;2:e74. doi: 10.1038/oncsis.2013.38. **(*equal contribution)**

***PIK3CA*^{H1047R} induces multipotency and multi-lineage mammary tumours.**

<u>Koren S</u>, Reavie L, Couto JP, De Silva D, Stadler MB, Roloff T, Britschgi A, Eichlisberger T, Kohler H, Aina O, Cardiff RD, Bentires-Alj M. **Nature. 2015** Sep 3;525(7567):114-8. doi: 10.1038/nature14669. Epub 2015 Aug 12.

Editors` choice in Science Signaling 2015 September 08

Highlighted in Cancer Discovery 2015 October 05

Mouse models of *PIK3CA* mutations: one mutation initiates heterogeneous mammary tumors

Shany Koren and Mohamed Bentires-Alj

Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland

The phosphoinositide 3-kinase (PI3K) signaling pathway is crucial for cell growth, proliferation, metabolism, and survival, and is frequently deregulated in human cancer, including ~ 70% of breast tumors. *PIK3CA*, the gene encoding the catalytic subunit p110α of PI3K, is mutated in ~ 30% of breast cancers. However, the exact mechanism of *PIK3CA*-evoked breast tumorigenesis has not yet been defined. Genetically engineered mouse models are valuable for examining the initiation, development and progression of cancer. Transgenic mice harboring hotspot mutations in p110α have helped to elucidate breast cancer pathogenesis and increase our knowledge about molecular and cellular alterations *in vivo*. They are also useful for the development of therapeutic strategies. Here, we describe current mouse models of mutant *PIK3CA* in the mammary gland, and discuss differences in tumor latency and pathogenesis.

Introduction

Phosphoinositide 3-kinases (PI3Ks) belong to a family of lipid kinases involved in metabolism, growth, proliferation, and survival signaling. Class Ia PI3Ks phosphorylate the 3-hydroxyl group of phosphatidylinositol 4,5-bisphosphate, resulting in the production of the second messenger phosphatidylinositol 3,4,5-trisphosphate. This recruits and activates several signaling proteins, including AKT and PDK1, leading to the activation of their downstream effectors. PI3K action is reversed by the PTEN phosphatase [1,2]. PI3Ks are heterodimers of regulatory (p85α, p50α, p55α, p85β, and p55γ) and catalytic (p110α, p110β, p110γ, and p110δ) subunits [2,3].

Genomic alterations of components of the PI3K pathway are found in ~ 70% of breast cancers [4]. The gene *PIK3CA* encodes the catalytic subunit p110α, and its amplification and/or mutation is associated with several kinds of human solid tumors [5–9]. Activating somatic mutations in *PIK3CA* are present in ~ 30% of human breast cancers at all stages [5,8,10–12]. In 47% of these cases, mutations occur in the kinase domain; the most frequent one is H1047R in exon 20. In 33% of these cases, mutations occur in the helical domain; the most frequent ones are E545K and E542K in exon 9 [8,13]. These mutations lead to a constitutively active enzyme with oncogenic capacity in cell culture [14–16]. Alterations in *PIK3CA* are found at similar frequencies in pure ductal carcinoma *in situ*, ductal carcinoma *in situ* adjacent to invasive ductal carcinoma, and in invasive ductal carcinoma, indicating that *PIK3CA* mutations occur early in carcinoma development [17]. In addition, mutant p110α has been found in distinct human breast cancer subtypes, such as estrogen receptor (ER)α-positive, progesterone

receptor (PR)-positive, human epidermal growth factor 2 (HER2)/Neu-positive and triple-negative breast cancers [12,18], but the correlation between *PIK3CA* mutations and pathological parameters remains controversial [5,12,19–25]. Also, assessment of the clinical outcome associated with these hotspot mutations showed contradictory results: some studies reported poor prognosis in breast cancer patients harboring *PIK3CA* exon 20 [26,27] or exon 9 mutations [10], whereas others reported favorable prognosis with improved overall survival in patients with exon 20 mutations [10,22]. Notably, *PIK3CA* mutations were shown to reduce the efficacy of HER2-targeted and ER-targeted therapies [28–30].

Adding to the list of mammary tumor mouse models [31], several groups have generated transgenic mice expressing *PIK3CA* H1047R in the mammary gland [32–36]. In contrast to mouse models such as Neu, Myc or polyoma middle T-antigen, which evoke tumors with a very specific phenotype, *PIK3CA* H1047R mutant mice have heterogeneous tumors. These new models should not only help to delineate the cellular and molecular mechanisms of action of mutant p110α *in vivo*, but also improve our understanding of tumor initiation, development and progression in breast cancer, provide a resource for the development of cancer therapies, and help to elucidate mechanisms of resistance to current PI3K pathway inhibitors.

Tumor formation in *PIK3CA* H1047R mutant mice

Transgenic mice expressing the amino acid substitution H1047R in the mammary gland were generated in order to model mutant *PIK3CA* breast cancer (Table 1; Fig. 1). For conditional mammary-specific expression of human *PIK3CA* or murine *Pik3ca* H1047R, two different promoters were used to drive Cre recombinase expression. First, Cre driven by the mouse mammary tumor virus (MMTV) long terminal repeat (MMTV-Cre) promoter results in mosaic expression of mutant *PIK3CA*/*Pik3ca* in differentiated mammary luminal cells and progenitor cells, and in other organs, depending on the MMTV-Cre line [35,37–41]. Second, Cre driven by the whey acidic protein promoter (WAPi-Cre) [42] (Fig. 1A) results in expression of mutant *PIK3CA* in alveolar progenitor cells and differentiated secretory luminal cells [34]. Tetracycline-inducible promoter systems (combined with MMTV-rtTA [43]; Fig. 1B) were also used to drive overexpression of H1047R, leading to a seven-fold to eight-fold change in expression of mutant *PIK3CA* as compared with *Pik3ca* [33]. Other groups have used a knock-in system to express endogenous levels of *Pik3ca* H1047R under the control of the native promoter (combined with MMTV-Cre [35,36,39,41]) (Fig. 1C,D).

Meyer *et al.* found mammary tumor-independent high lethality (~ 75%) in MMTV-Cre *PIK3CA* H1047R mice. Although the cause of death could not be identified, promoter leakiness leading to expression of *PIK3CA* H1047R in other tissues was suggested [34]. Using two different MMTV-Cre lines [37–39] to induce expression of *PIK3CA* H1047 (Fig. 1E), Adams *et al.* found that some MMTV-Cre[lineA] *Pik3ca* H1047R and MMTV-Cre[NLST] *Pik3ca* H1047R mice reached endpoint (lethargy, impaired breathing, tumor) independently of mammary tumors [32]. These observations raise concerns about whether MMTV-Cre is the optimal promoter system with which to study *PIK3CA*-induced mammary cancer in mice.

Each of these systems leads to the development of heterogeneous mammary tumors. The most prominent phenotypes, adenosquamous carcinoma and adenomyoepithelioma, express ERα, as well as basal [e.g. cytokeratin (CK)5 and CK14] and luminal (e.g. CK8 and CK18) markers. The *Pik3ca* H1047R knock-in models of Yuan *et al.* and Tikoo *et al.* led mostly to hormone receptor-positive fibroadenomas (76.9% and 45%, respectively) or sarcomas (42.5%). Other histopathological features, such as adenocarcinoma, carcinosarcoma, and osteosarcoma, were also observed [35,36] (Table 1). Heterogeneity, a feature of human breast cancer, was also reported in mouse models of *Pten* inactivation [44,45]. In another study, however, loss of PTEN resulted only in adenomyoepithelioma [46] (for a review of PTEN mouse models, see [47]).

Nulliparous WAPi-Cre *PIK3CA* H1047R animals developed tumors at an average of 219 days (K.S. and M.B.A., unpublished data), and nulliparous MMTV-Cre *Pik3ca* knock-in H1047R animals developed tumors at an average of 484 and 492 days [35,36]. Notably, tumor latencies in parous WAPi-Cre *PIK3CA* H1047R animals were, on average, 140 days [34], and those in MMTV-Cre *Pik3ca* knock-in H1047R animals were 392 and 465 days [35,36], showing that pregnancy accelerates tumorigenesis in these models. An increase in the number of H1047R-expressing cells after pregnancy and a delay in involution (days 2 and 8), owing to a reduced number of apoptotic cells, were observed in WAPi-Cre H1047R mice [34].

Tumor formation was also investigated by using MMTV-driven expression of nonmutated p110α fused to a Src myristolyation sequence (Fig. 1F), which results in the recruitment of p110α to the membrane and constitutive activation of PI3K signaling (MMTV-MYR-p110α). Transgenic mice developed heterogeneous ER-positive

Table 1. Summary of current mouse models of *PIK3CA* alterations. α-SMA, α-smooth muscle actin.

Mouse model	Mean age at tumor onset	Pathology	Ref.
WAPi-Cre *PIK3CA* H1047R (transgenic model; Fig. 1A)	Parous: 140.3 ± 6.9 days (= 36.8 ± 4.9 days after delivery) Nulliparous: 219 ± 12 days	Adenosquamous carcinoma (54.6%), adenomyoepithelioma (22.7%, PR+), adenocarcinoma with squamous metaplasia (13.6%), adenocarcinoma (9.1%) (all ER+, CK14+, CK18+, and CK14/CK18+)	[34]
MMTV-Cre *PIK3CA* H1047R (transgenic model; Fig. 1A)	214 ± 22.6 days	Adenomyoepithelioma (100%) (ER+, PR+, CK14+, CK18+, and α-SMA+)	[34]
MMTV-rtTA TetO-*PIK3CA* H1047R (inducible transgenic model; Fig. 1B)	7 months	Adenocarcinoma and adenosquamous carcinoma	[33]
MMTV-Cre *Pik3ca* H1047R (knock-in model; Fig. 1C)	Parous: 465 days Nulliparous: 492 days	Fibroadenoma (76.9%), adenocarcinoma (15.4%) (both CK5+, CK18+, ER+ and PR+); spindle cell neoplasia (7.7%) (ER−, PR−, CK5−, CK18+, vimentin+)	[36]
MMTV-Cre *Pik3ca* H1047R (knock-in model; Fig. 1D)	Parous: 393 days Nulliparous: 484 days	Benign fibroadenoma (45%), carcinosarcoma (both ER+, CK5/CK6+, CK8/CK18+, CK5/CK8+, CK8/E-cadherin+) or sarcoma (42.5%); adenosquamous carcinoma (10%) (CK5/CK6+, CK8−, E-cadherin−); osteosarcoma (2.5%)	[35]
MMTV-Cre^NLST *Pik3ca* H1047R (transgenic model; Fig. 1E)	5 months	Adenosquamous carcinoma (51%) (ER+, CK8+, CK14+, CK8/CK14+, CK8/CK14−, N-cadherin+, vimentin+, Atf3+, CK10+, β-catenin+), adenomyoepithelioma (45%) (ER+, CK8+, CK14+, N-cadherin+, Atf3+, desmin+, β-catenin+), spindle cell tumors (1%), poorly differentiated adenocarcinoma (3%)	[32]
MMTV Cre^NLST *Pik3ca* H1047R; p53 ^fl/+ (transgenic model)	< 5 months	Spindle cell/EMT tumors (33%) (ER+, CK8+, CK14+, N-cadherin+, desmin+), adenosquamous carcinoma (52%) (ER+, CK8+, CK14+, CK8/CK14+, CK8/CK14−, N-cadherin+, desmin−, CK10+), radial scar type lesions (10%) and poorly differentiated adenocarcinoma (5%) (ER+, CK8+, CK14+, CK8/CK14+, CK8/CK14−)	[32]
MMTV-MYR-p110α (Fig. 1F) and MMTV-MYR-p110α; p53+/− (transgenic model)	Not reported	Nulliparous: adenosquamous carcinoma. Parous: 66% adenosquamous carcinoma, 34% carcinoma (all ER+, cathepsin D+)	[48]
MMTV-MYR-p110α; CDK4(R24C) (transgenic model)	Not reported	Nulliparous: adenosquamous carcinoma, papillary adenocarcinoma, carcinoma, sarcoma Parous: adenosquamous carcinoma, complex adenocarcinoma, carcinoma, sarcoma (all ER+ except sarcoma)	[48]

mammary tumors, but at a frequency lower than in mice expressing mutant *PIK3CA* [48].

Synergism between *PIK3CA* H1047R and *p53* alterations

Whole-exome capture and sequencing of mammary tumors from MMTV-Cre knock-in *Pik3ca* H1047R mice of various histotypes has revealed greater increases in somatic mutations in spindle cell tumors (∼ 44–88) and adenocarcinoma (∼ 4–61) than in fibroadenoma (∼ 2–13) [36]. Moreover, comparative genomic hybridization-array profiling showed a greater accumulation of chromosomal copy number alterations in spindle cell tumors than in adenocarcinoma and fibroadenoma [36]. Functional validation and examination of the clinical relevance of these secondary genomic alterations are now warranted.
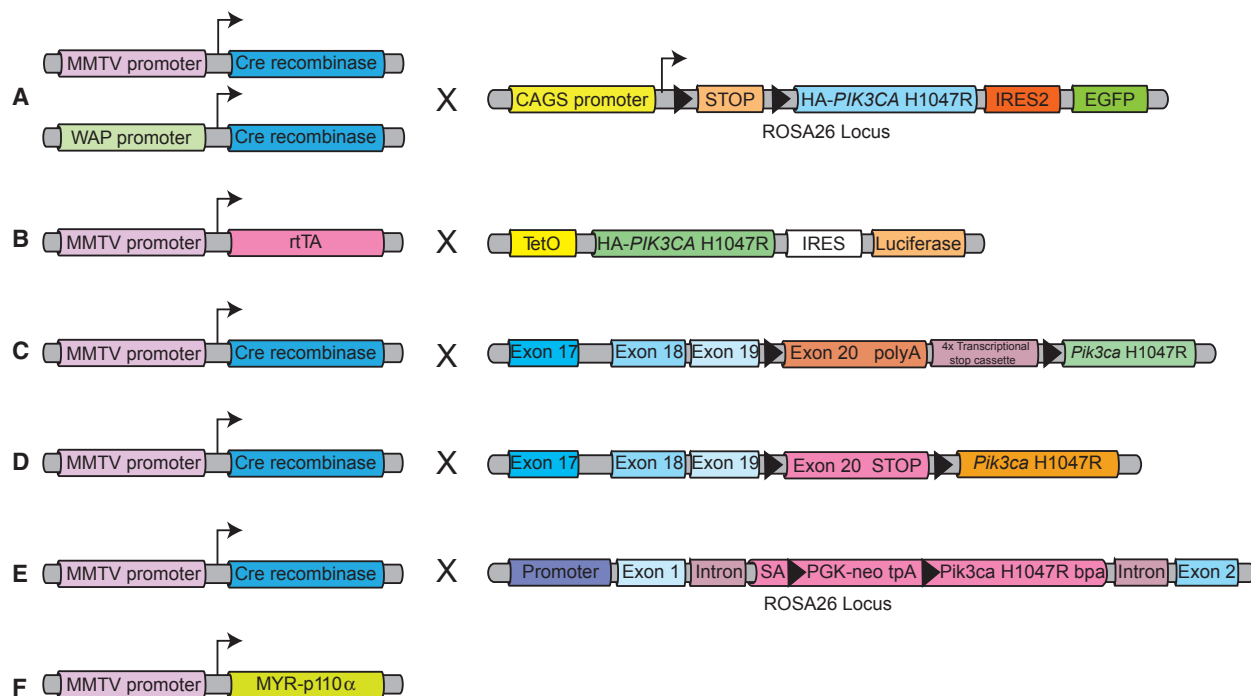
Fig. 1. Schematic overview of constructs that were used to generate mutant *PIK3CA/Pik3ca* models. Triangles represent loxp sites. (A) WAPi-Cre *PIK3CA* H1047R and MMTV-Cre *PIK3CA* H1047R models [34]. (B) MMTV-rtTA TetO-*PIK3CA* H1047R model [33]. (C) MMTV-Cre *Pik3ca* H1047R model [36]. (D) MMTV-Cre *Pik3ca* H1047R model [35]. (E) MMTV-Cre^{NLST} *Pik3ca* H1047R model [32]. (F) MMTV-MYR-p110α model [48]. EGFP, enhanced green fluorescent protein; HA, hemagglutinin; IRES, internal ribosome entry site; PGK, phosphoglycerate kinase; SA, splice acceptor sequence.

Some human breast tumors harbor alterations in *PIK3CA* in combination with mutant *p53* [18,49,50]. *p53* mutations (R245H, A135V, and I192N) were among the secondary mutations identified by Yuan *et al.* [36] in adenocarcinoma and spindle cell neoplasia. It is likely that these mutations prevented the well-established p53-dependent tumor suppression. Notably, *Pik3ca* mutant mouse models were used to investigate the interaction of *Pik3ca* H1047R and *p53* [32,36]. Heterozygosity in *p53* was shown to accelerate tumor onset in MMTV-Cre *Pik3ca* H1047R mice [32]. The tumor histotype in double mutants consisted mostly of ER-positive, CK14-positive and CK8-positive spindle cell tumors that express epithelial–mesenchymal transition markers, or adenosquamous carcinoma [32] (Table 1).

p53 has been found to be inactivated in MMTV-MYR-p110α-evoked tumors, suggesting that p53 loss is important for tumorigenesis in this model. Consistently, no difference was found in tumor latency or tumor phenotype between MMTV-MYR-p110α mice in a heterozygous *p53* background and MMTV-MYR-p110α mice [48]. Notably, MMTV-MYR-p110α mice in an inactive retinoblastoma (pRB) background

(CDK4 R24C knock-in line [51,52]) showed enhanced mammary tumorigenesis. These data suggest that *PIK3CA* H1047R-evoked tumor suppression mechanisms can be circumvented by inactivation of either p53 or pRB [48].

PIK3CA H1047R mutations and metastasis

Mouse models with an altered PI3K pathway can increase our understanding of breast cancer progression and metastatic spread. Metastases were reported in *Pten* heterozygous mice: one study found a metastatic tumor in the regional lymph node of one mouse, and three other mice had lung metastases. Their morphological appearance was similar to that of the primary tumor [44]. In contrast, metastases are rarely found in *PIK3CA/Pik3ca* H1047R mouse models. There is only one report of isolated lung metastasis in MMTV-Cre^{NLST} *Pik3ca* H1047R mice [32]. However, *PIK3CA* mutations occur at high frequencies in metastatic human breast cancer [21]. Surprisingly, oncogenic *PIK3CA*-driven breast tumors have a longer time to recurrence after surgery [21], and some clinical

studies reported a good prognosis [10,22]. These observations may mean that mutant *PIK3CA* results in a selective advantage for breast cancer cells at the primary site, but not during metastatic progression and colonization of distant sites. The analysis of *PIK3CA* status in a large number of metastatic lesions, circulating tumor cells and matched primary tumors should clarify this '*PIK3CA* paradox' [53]. An alternative explanation is that patients with mutations in *PIK3CA* respond well to the current standard of care, resulting in this apparent paradox.

Constitutively active PI3K signaling, in association with further genomic alterations, induces mammary cancer in mice, suggesting a causative role for *PIK3CA* mutations in breast tumorigenesis. Additional gain-of-function or loss-of-function genomic alterations may also contribute to metastasis in breast tumor progression. The identification of these synergistic oncogenic pathways is of paramount importance for the elucidation of the 'wiring diagram' of tumor cells with *PIK3CA* mutations.

Therapeutic strategies and resistance in mutant *PIK3CA* mouse models

At least 26 PI3K pathway-targeting compounds are currently undergoing over 150 cancer-related clinical trials [54]. *PIK3CA* mouse models also serve as a valuable tool for testing anticancer drugs [55,56]. A test of the efficacy of the PI3K inhibitor GDC-0941 by Yuan *et al.* showed a decline in tumor growth in spindle cell tumors with *Pik3ca* H1047R and *p53* mutations [36].

These mouse models have already proved useful for investigating resistance mechanisms. For example, *PIK3CA* H1047R-driven tumors were shown to recur after *PIK3CA* H1047R inactivation in a PI3K-dependent or PI3K-independent manner [33]. Tumor survival in *c-MET*-elevated tumors was shown to depend on an active endogenous PI3K pathway, whereas *c-MYC* elevation contributed to oncogene independence and GDC-0941 resistance [33].

The PI3K-independent recurrence of *PIK3CA* H1047R-initiated mammary tumors shows how important it is to investigate associated pathways involved in tumor formation that may result in escape from treatment. Their delineation should pave the way for the development of mechanism-based combination therapies.

Future perspectives

Current mouse models of mutant *PIK3CA* are contributing to a better understanding of breast cancer pathogenesis associated with alterations in the PI3K pathway. So far, characterization of *PIK3CA* H1047R-evoked tumors has clarified some pathophysiological aspects, but has also raised further relevant questions.

Tumor heterogeneity and the observation that *PIK3CA* H1047R mutants develop CK5/CK14-positive and CK8/CK18-positive mammary carcinomas [32,34–36] suggest either a luminal and basal tumor cell of origin or the dedifferentiation of cell lineage-committed tumor cells to multipotent progenitors that then give rise to CK5/CK14-positive and CK8/CK18-positive cells. Consistent with the latter possibility, Tikoo *et al.* found that expression of *Pik3ca* H1047R results in expansion of the luminal progenitor population. Furthermore, the putative mammary stem cell-enriched basal population and the luminal progenitors of mutants show enhanced colony-forming ability and a larger colony size [35], but the molecular mechanisms underlying these effects have not yet been defined. The current mouse models do not definitively answer the question of what causes heterogeneity in *PIK3CA* H1047R-evoked tumors and which cell type gives rise to which subtype of mammary cancer. However, lineage-tracing experiments should provide further information about the cell of origin and cellular hierarchy in tumors.

There have been few studies of synergistic pathways contributing to tumor progression *in vivo*, especially metastatic spread, and to resistance to therapy. Such investigations should lead to an increase in our understanding of tumor relapse and therapy failure, and should help to define targets for preventing *PIK3CA*-driven tumor initiation and progression.

Acknowledgements

References

1 Cantley LC (2002) The phosphoinositide 3-kinase pathway. *Science* **296**, 1655–1657.
2 Yuan TL & Cantley LC (2008) PI3K pathway alterations in cancer: variations on a theme. *Oncogene* **27**, 5497–5510.
3 Zhao L & Vogt PK (2008) Class I PI3K in oncogenic cellular transformation. *Oncogene* **27**, 5486–5496.

4 Miller TW, Rexer BN, Garrett JT & Arteaga CL (2011) Mutations in the phosphatidylinositol 3-kinase pathway: role in tumor progression and therapeutic implications in breast cancer. *Breast Cancer Res* **13**, 224. doi: 10. 1186/bcr3039.

5 Bachman KE, Argani P, Samuels Y, Silliman N, Ptak J, Szabo S, Konishi H, Karakas B, Blair BG, Lin C *et al.* (2004) The PIK3CA gene is mutated with high frequency in human breast cancers. *Cancer Biol Ther* **3**, 772–775.

6 Kadota M, Sato M, Duncan B, Ooshima A, Yang HH, Diaz-Meyer N, Gere S, Kageyama S, Fukuoka J, Nagata T *et al.* (2009) Identification of novel gene amplifications in breast cancer and coexistence of gene amplification with an activating mutation of PIK3CA. *Cancer Res* **69**, 7357–7365.

7 Levine DA, Bogomolniy F, Yee CJ, Lash A, Barakat RR, Borgen PI & Boyd J (2005) Frequent mutation of the PIK3CA gene in ovarian and breast cancers. *Clin Cancer Res* **11**, 2875–2878.

8 Samuels Y, Wang Z, Bardelli A, Silliman N, Ptak J, Szabo S, Yan H, Gazdar A, Powell SM, Riggins GJ *et al.* (2004) High frequency of mutations of the PIK3CA gene in human cancers. *Science* **304**, 554. doi: 10.1126/science.1096502.

9 Wu G, Xing M, Mambo E, Huang X, Liu J, Guo Z, Chatterjee A, Goldenberg D, Gollin SM, Sukumar S *et al.* (2005) Somatic mutation and gain of copy number of PIK3CA in human breast cancer. *Breast Cancer Res* **7**, R609–R616.

10 Barbareschi M, Buttitta F, Felicioni L, Cotrupi S, Barassi F, Del Grammastro M, Ferro A, Dalla Palma P, Galligioni E & Marchetti A (2007) Different prognostic roles of mutations in the helical and kinase domains of the PIK3CA gene in breast carcinomas. *Clin Cancer Res* **13**, 6064–6069.

11 Miller TW (2012) Initiating breast cancer by PIK3CA mutation. *Breast Cancer Res* **14**, 301. doi:10.1186/bcr3103.

12 Saal LH, Holm K, Maurer M, Memeo L, Su T, Wang X, Yu JS, Malmstrom PO, Mansukhani M, Enoksson J *et al.* (2005) PIK3CA mutations correlate with hormone receptors, node metastasis, and ERBB2, and are mutually exclusive with PTEN loss in human breast carcinoma. *Cancer Res* **65**, 2554–2559.

13 Bader AG, Kang S, Zhao L & Vogt PK (2005) Oncogenic PI3K deregulates transcription and translation. *Nat Rev Cancer* **5**, 921–929.

14 Bader AG, Kang S & Vogt PK (2006) Cancer-specific mutations in PIK3CA are oncogenic in vivo. *Proc Natl Acad Sci USA* **103**, 1475–1479.

15 Isakoff SJ, Engelman JA, Irie HY, Luo J, Brachmann SM, Pearline RV, Cantley LC & Brugge JS (2005) Breast cancer-associated PIK3CA mutations are oncogenic in mammary epithelial cells. *Cancer Res* **65**, 10992–11000.

16 Zhao JJ, Liu Z, Wang L, Shin E, Loda MF & Roberts TM (2005) The oncogenic properties of mutant p110alpha and p110beta phosphatidylinositol 3-kinases in human mammary epithelial cells. *Proc Natl Acad Sci USA* **102**, 18443–18448.

17 Miron A, Varadi M, Carrasco D, Li H, Luongo L, Kim HJ, Park SY, Cho EY, Lewis G, Kehoe S *et al.* (2010) PIK3CA mutations in in situ and invasive breast carcinomas. *Cancer Res* **70**, 5674–5678.

18 Boyault S, Drouet Y, Navarro C, Bachelot T, Lasset C, Treilleux I, Tabone E, Puisieux A & Wang Q (2012) Mutational characterization of individual breast tumors: TP53 and PI3K pathway genes are frequently and distinctively mutated in different subtypes. *Breast Cancer Res Treat* **132**, 29–39.

19 Campbell IG, Russell SE, Choong DY, Montgomery KG, Ciavarella ML, Hooi CS, Cristiano BE, Pearson RB & Phillips WA (2004) Mutation of the PIK3CA gene in ovarian and breast cancer. *Cancer Res* **64**, 7678–7681.

20 Dunlap J, Le C, Shukla A, Patterson J, Presnell A, Heinrich MC, Corless CL & Troxell ML (2010) Phosphatidylinositol-3-kinase and AKT1 mutations occur early in breast carcinoma. *Breast Cancer Res Treat* **120**, 409–418.

21 Dupont Jensen J, Laenkholm AV, Knoop A, Ewertz M, Bandaru R, Liu W, Hackl W, Barrett JC & Gardner H (2011) PIK3CA mutations may be discordant between primary and corresponding metastatic disease in breast cancer. *Clin Cancer Res* **17**, 667–677.

22 Kalinsky K, Jacks LM, Heguy A, Patil S, Drobnjak M, Bhanot UK, Hedvat CV, Traina TA, Solit D, Gerald W *et al.* (2009) PIK3CA mutation associates with improved outcome in breast cancer. *Clin Cancer Res* **15**, 5049–5059.

23 Li H, Zhu R, Wang L, Zhu T, Li Q, Chen Q, Wang H & Zhu H (2010) PIK3CA mutations mostly begin to develop in ductal carcinoma of the breast. *Exp Mol Pathol* **88**, 150–155.

24 Li SY, Rong M, Grieu F & Iacopetta B (2006) PIK3CA mutations in breast cancer are associated with poor outcome. *Breast Cancer Res Treat* **96**, 91–95.

25 Perez-Tenorio G, Alkhori L, Olsson B, Waltersson MA, Nordenskjold B, Rutqvist LE, Skoog L & Stal O (2007) PIK3CA mutations and PTEN loss correlate with similar prognostic factors and are not mutually exclusive in breast cancer. *Clin Cancer Res* **13**, 3577–3584.

26 Lai YL, Mau BL, Cheng WH, Chen HM, Chiu HH & Tzen CY (2008) PIK3CA exon 20 mutation is independently associated with a poor prognosis in breast cancer patients. *Ann Surg Oncol* **15**, 1064–1069.

27 Mangone FR, Bobrovnitchaia IG, Salaorni S, Manuli E & Nagai MA (2012) PIK3CA exon 20 mutations are associated with poor prognosis in breast cancer patients. *Clinics* **67**, 1285–1290.

28 Eichhorn PJ, Gili M, Scaltriti M, Serra V, Guzman M, Nijkamp W, Beijersbergen RL, Valero V, Seoane J, Bernards R *et al.* (2008) Phosphatidylinositol 3-kinase hyperactivation results in lapatinib resistance that is reversed by the mTOR/phosphatidylinositol 3-kinase inhibitor NVP-BEZ235. *Cancer Res* **68**, 9221–9230.

29 Kataoka Y, Mukohara T, Shimada H, Saijo N, Hirai M & Minami H (2010) Association between gain-of-function mutations in PIK3CA and resistance to HER2-targeted agents in HER2-amplified breast cancer cell lines. *Ann Oncol* **21**, 255–262.

30 Miller TW, Forbes JT, Shah C, Wyatt SK, Manning HC, Olivares MG, Sanchez V, Dugger TC, de Matos Granja N, Narasanna A *et al.* (2009) Inhibition of mammalian target of rapamycin is required for optimal antitumor effect of HER2 inhibitors against HER2-overexpressing cancer cells. *Clin Cancer Res* **15**, 7266–7276.

31 Borowsky AD (2011) Choosing a mouse model: experimental biology in context – the utility and limitations of mouse models of breast cancer. *Cold Spring Harbor Perspect Biol* **3**, a009670. doi: 10.1101/cshperspect.a009670.

32 Adams JR, Xu K, Liu JC, Agamez NM, Loch AJ, Wong RG, Wang W, Wright KL, Lane TF, Zacksenhaus E *et al.* (2011) Cooperation between Pik3ca and p53 mutations in mouse mammary tumor formation. *Cancer Res* **71**, 2706–2717.

33 Liu P, Cheng H, Santiago S, Raeder M, Zhang F, Isabella A, Yang J, Semaan DJ, Chen C, Fox EA *et al.* (2011) Oncogenic PIK3CA-driven mammary tumors frequently recur via PI3K pathway-dependent and PI3K pathway-independent mechanisms. *Nat Med* **17**, 1116–1120.

34 Meyer DS, Brinkhaus H, Muller U, Muller M, Cardiff RD & Bentires-Alj M (2011) Luminal expression of PIK3CA mutant H1047R in the mammary gland induces heterogeneous tumors. *Cancer Res* **71**, 4344–4351.

35 Tikoo A, Roh V, Montgomery KG, Ivetac I, Waring P, Pelzer R, Hare L, Shackleton M, Humbert P & Phillips WA (2012) Physiological levels of Pik3ca(H1047R) mutation in the mouse mammary gland results in ductal hyperplasia and formation of ERalpha-positive tumors. *PLoS One* **7**, e36924.

36 Yuan W, Stawiski E, Janakiraman V, Chan E, Durinck S, Edgar KA, Kljavin NM, Rivers CS, Gnad F, Roose-Girma M *et al.* (2012) Conditional activation of Pik3ca (H1047R) in a knock-in mouse model promotes mammary tumorigenesis and emergence of mutations. *Oncogene* **32**, 318–326.

37 Soriano P (1999) Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat Genet* **21**, 70–71.

38 Srinivas S, Watanabe T, Lin CS, William CM, Tanabe Y, Jessell TM & Costantini F (2001) Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev Biol* **1**, 4. doi: 10.1186/1471-213X-1-4.

39 Wagner KU, Wall RJ, St-Onge L, Gruss P, Wynshaw-Boris A, Garrett L, Li M, Furth PA & Hennighausen L (1997) Cre-mediated gene deletion in the mammary gland. *Nucleic Acids Res* **25**, 4323–4330.

40 Andrechek ER, White D & Muller WJ (2005) Targeted disruption of ErbB2/Neu in the mammary epithelium results in impaired ductal outgrowth. *Oncogene* **24**, 932–937.

41 Wagner KU, McAllister K, Ward T, Davis B, Wiseman R & Hennighausen L (2001) Spatial and temporal expression of the Cre gene under the control of the MMTV-LTR in different lines of transgenic mice. *Transgenic Res* **10**, 545–553.

42 Wintermantel TM, Mayer AK, Schutz G & Greiner EF (2002) Targeting mammary epithelial cells using a bacterial artificial chromosome. *Genesis* **33**, 125–130.

43 Gunther EJ, Belka GK, Wertheim GB, Wang J, Hartman JL, Boxer RB & Chodosh LA (2002) A novel doxycycline-inducible system for the transgenic analysis of mammary gland biology. *FASEB J* **16**, 283–292.

44 Stambolic V, Tsao MS, Macpherson D, Suzuki A, Chapman WB & Mak TW (2000) High incidence of breast and endometrial neoplasia resembling human Cowden syndrome in pten+/– mice. *Cancer Res* **60**, 3605–3611.

45 Li G, Robinson GW, Lesche R, Martinez-Diaz H, Jiang Z, Rozengurt N, Wagner KU, Wu DC, Lane TF, Liu X *et al.* (2002) Conditional loss of PTEN leads to precocious development and neoplasia in the mammary gland. *Development* **129**, 4159–4170.

46 Dourdin N, Schade B, Lesurf R, Hallett M, Munn RJ, Cardiff RD & Muller WJ (2008) Phosphatase and tensin homologue deleted on chromosome 10 deficiency accelerates tumor induction in a mouse model of ErbB-2 mammary tumorigenesis. *Cancer Res* **68**, 2122–2131.

47 Hollander MC, Blumenthal GM & Dennis PA (2011) PTEN loss in the continuum of common cancers, rare syndromes and mouse models. *Nat Rev Cancer* **11**, 289–301.

48 Renner O, Blanco-Aparicio C, Grassow M, Canamero M, Leal JF & Carnero A (2008) Activation of phosphatidylinositol 3-kinase by membrane localization of p110alpha predisposes mammary glands to neoplastic transformation. *Cancer Res* **68**, 9643–9653.

49 Buttitta F, Felicioni L, Barassi F, Martella C, Paolizzi D, Fresu G, Salvatore S, Cuccurullo F, Mezzetti A, Campani D *et al.* (2006) PIK3CA mutation and histologic type in breast carcinoma: high frequency of mutations in lobular carcinoma. *J Pathol* **208**, 350–355.

50 Maruyama N, Miyoshi Y, Taguchi T, Tamaki Y, Monden M & Noguchi S (2007) Clinicopathologic

analysis of breast cancers with PIK3CA mutations in Japanese women. *Clin Cancer Res* **13**, 408–414.

51 Quereda V, Martinalbo J, Dubus P, Carnero A & Malumbres M (2007) Genetic cooperation between p21Cip1 and INK4 inhibitors in cellular senescence and tumor suppression. *Oncogene* **26**, 7665–7674.

52 Sotillo R, Dubus P, Martin J, de la Cueva E, Ortega S, Malumbres M & Barbacid M (2001) Wide spectrum of tumors in knock-in mice carrying a Cdk4 protein insensitive to INK4 inhibitors. *EMBO J* **20**, 6637–6647.

53 Meyer DS & Bentires-Alj M (2010) Can phosphatidylinositol 3-kinase/mammalian target of rapamycin inhibition ERase them all? *Breast Cancer Res* **12**, 315. doi: 10.1186/bcr2718.

54 Sheppard K, Kinross KM, Solomon B, Pearson RB & Phillips WA (2012) Targeting PI3 kinase/AKT/mTOR signaling in cancer. *Crit Rev Oncog* **17**, 69–95.

55 Courtney KD, Corcoran RB & Engelman JA (2010) The PI3K pathway as drug target in human cancer. *J Clin Oncol* **28**, 1075–1083.

56 Engelman JA, Chen L, Tan X, Crosby K, Guimaraes AR, Upadhyay R, Maira M, McNamara K, Perera SA, Song Y *et al.* (2008) Effective use of PI3K and MEK inhibitors to treat mutant Kras G12D and PIK3CA H1047R murine lung cancers. *Nat Med* **14**, 1351–1356.

npg

SHORT COMMUNICATION

Expression of *PIK3CA* mutant E545K in the mammary gland induces heterogeneous tumors but is less potent than mutant H1047R

DS Meyer[1,4,5], S Koren[1,4], C Leroy[1,2], H Brinkhaus[1], U Müller[1], I Klebba[1,6], M Müller[2], RD Cardiff[3] and M Bentires-Alj[1]

The phosphoinositide 3-kinase (PI3K) signaling cascade is a key mediator of cellular growth, survival and metabolism and is frequently subverted in human cancer. The gene encoding for the alpha catalytic subunit of PI3K (*PIK3CA*) is mutated and/or amplified in ~30% of breast cancers. Mutations in either the kinase domain (H1047R) or the helical domain (E545K) are most common and result in a constitutively active enzyme with oncogenic capacity. *PIK3CA*[H1047R] was previously demonstrated to induce tumors in transgenic mouse models; however, it was not known whether overexpression of *PIK3CA*[E545K] is sufficient to induce mammary tumors and whether tumor initiation by these two types of mutants differs. Here, we demonstrate that expression of *PIK3CA*[E545K] in the mouse mammary gland induces heterogenous mammary carcinomas but with a longer latency than *PIK3CA*[H1047R]-expressing mice. Our results suggest that the helical domain mutant *PIK3CA*[E545K] is a less potent inducer of mammary tumors due to less efficient activation of downstream Akt signaling.

INTRODUCTION

The phosphoinositide 3-kinase (PI3K) pathway is a key regulator of cell growth, proliferation, metabolism and survival and is often found to be hyperactivated in human cancer.[1,2] The most common aberrations of the PI3K pathway include mutation and/or amplification of *PIK3CA*,[3–7] the gene encoding the alpha catalytic subunit of the kinase (p110α), loss of expression of the PTEN phosphatase that reverses PI3K action, activation downstream of oncogenic receptor tyrosine kinases and mutation/amplification of Akt.[1] Hyperactivation of the PI3K pathway increases tumorigenicity by reducing cell death and increasing cell proliferation, migration, invasion, metabolism and angiogenesis.[1,2] It also enhances resistance to chemotherapy.[8]

The majority of mutations in *PIK3CA* occur at two 'hotspots' within the kinase (H1047R) and helical domains (E542K and E545K) of p110α.[6,9] These mutations lead to a constitutively active enzyme, transform cells *in vitro*, and enhance tumorigenicity in xenograft models.[10–13] Notably, different mechanisms underlie the gain-of-function activities of helical- and kinase domain mutants. While *PIK3CA*[E545K] is independent of binding to the adaptor molecule p85 but requires interaction with Ras-GTP, the *PIK3CA*[H1047R] mutant is highly dependent on p85 for its oncogenic capacity but independent of Ras-GTP.[14]

PIK3CA gain-of-function mutations are found in ~30% of human breast cancers[3,6,15–17] and most likely occur at an early stage of breast carcinoma development, as suggested by the similar mutation frequencies in *PIK3CA* found in pure ductal carcinoma *in situ*, ductal carcinoma *in situ* adjacent to invasive ductal carcinoma, and invasive ductal carcinoma.[18] Evaluation of the clinical outcome of genomic alterations in *PIK3CA* has produced contradictory results.[15,19,20] However, these studies showed that alterations in different exons of *PIK3CA* have varying impacts on tumor development and progression and, therefore, differ in prognostic value. For example, both mutations are associated with lower grade and hormone receptor-positive tumors, but *PIK3CA*[H1047R] mutants are strongly associated with lymph-node negativity and *PIK3CA*[E545K] mutants with older age at diagnosis, indicating the different oncogenic potentials of the H1047R and E545K mutations.[19] This is further supported by the different frequencies of E545K (~6%) and H1047R (~15%) mutations in breast cancer.[17,21] *In vivo* transplantation assays have demonstrated *PIK3CA*[H1047R] to be more potent in inducing tumors[10] but another study found no trend,[12] and the exact impact of these mutations on breast cancer has remained controversial.

We and others have reported that expression of *PIK3CA*[H1047R] in the mammary gland induces heterogeneous tumors.[22–25] To determine which *PIK3CA* mutant shows higher oncogenic activity *in vivo*, we generated a novel conditional mouse model expressing *PIK3CA*[E545K]. We have demonstrated that *PIK3CA*[E545K] induces heterogeneous mammary tumors that express basal and luminal markers but is a less potent oncogene *in vivo* than *PIK3CA*[H1047R].

---

[1]Mechanisms of Cancer, Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland; [2]Developmental and Molecular Pathways, Novartis Institutes for Biomedical Research, Basel, Switzerland and [3]Department of Pathology, Center for Comparative Medicine, University of California Davis, Davis, CA, USA. Correspondence: Dr M Bentires-Alj, Mechanisms of Cancer, Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058, Basel, Switzerland.
E-mail: bentires@fmi.ch
[4]These authors contributed equally to this work.
[5]Current address: Helen Diller Family Comprehensive Cancer Center, School of Medicine, University of California San Francisco (UCSF), San Francisco, CA, USA.
[6]Current address: Department of Biosystems Science and Engineering (D-BSSE), ETH Zürich, Basel, Switzerland.

RESULTS AND DISCUSSION

Expression of PIK3CA[E545K] but not wild-type PIK3CA induces mammary tumors

We and others have shown that PIK3CA[H1047R] induces mouse mammary carcinomas.[22–25] To test whether overexpression of wild-type human PIK3CA (PIK3CA[wt]) or PIK3CA[E545K] also induces mammary tumors, we generated novel transgenic mice that conditionally express PIK3CA[wt] or PIK3CA[E545K] (Figure 1a). To achieve equivalent transgene expression, we integrated PIK3CA[wt] or PIK3CA[E545K] into the ROSA26 locus using recombinase-mediated cassette exchange.[26] Correct integration of the target cassettes was confirmed in the resulting PIK3CA[wt] and PIK3CA[E545K] lines (Figure 1b, left). PIK3CA[wt] and PIK3CA[E545K] animals were then crossed to WAPiCre mice in which expression of recombinase Cre is controlled by the whey acidic protein (WAP) promoter, which is mainly active in secretory mammary epithelial cells, and expression of the transgenes confirmed (Figure 1b, right).[27–30] This enabled us to directly compare the kinetics of tumor onset in PIK3CA[wt] and PIK3CA[E545K] mice, and the previously reported WAPiCre PIK3CA[H1047R] mice.[22]

The resulting bi-transgenic WAPiCre PIK3CA[wt] and WAPiCre PIK3CA[E545K] female mice were impregnated to achieve maximal Cre-mediated recombination and the pups removed 1 day after delivery. All WAPiCre PIK3CA[E545K] mice developed mammary tumors on average 80 ($\pm$10) days after delivery, whereas parous WAPiCre PIK3CA[wt] mice did not form tumors within 520 days (Figure 1c). This indicates that overexpression of wild-type PIK3CA itself is insufficient to induce mammary tumors. Of note, the latency to tumor onset in WAPiCre PIK3CA[E545K] animals was significantly longer than that observed previously for WAPiCre PIK3CA[H1047R] mice (36 ($\pm$4.9) days).[22] We also crossed PIK3CA[E545K] and PIK3CA[H1047R] lines to CAGs-CreERT2 mice that express a tamoxifen-inducible Cre/estrogen receptor (ER) fusion protein under the control of a modified β-actin promoter; this results in the expression of Cre-ER in virtually all cells. Unexpectedly, bi-transgenic CAGs-Cre PIK3CA[E545K] and PIK3CA[H1047R] mice died by the age of 4 months even when no tamoxifen was administered. Although we were unable to identify the exact cause of death, we concluded that leakiness of the CAGs-CreERT2 system caused premature and deleterious PIK3CA[E545K] or PIK3CA[H1047R] expression in various tissues of these mice (DSM and MB-A, unpublished observations).

To compare the tumor-initiating potential of the two different PIK3CA mutants, we then transplanted pieces of mammary gland tissue from CAGs-CreERT2 PIK3CA[E545K] or PIK3CA[H1047R] donor mice previously treated with tamoxifen into cleared fat pads of Balb/c mice. The mammary glands reconstituted by either CAGs-CreERT2 PIK3CA[E545K] or PIK3CA[H1047R]-derived epithelium were hyperplastic (data not shown) and eventually formed tumors after 229 ($\pm$17, PIK3CA[H1047R]) and 336 days ($\pm$20, PIK3CA[E545K]), respectively (Figure 1d). As observed in the WAPiCre mouse cohorts, PIK3CA[E545K] was significantly less potent than PIK3CA[H1047R] in the induction of mammary carcinomas, which is a possible explanation for the lower frequency of E542K/E545K mutations in human breast cancer.[17,21]
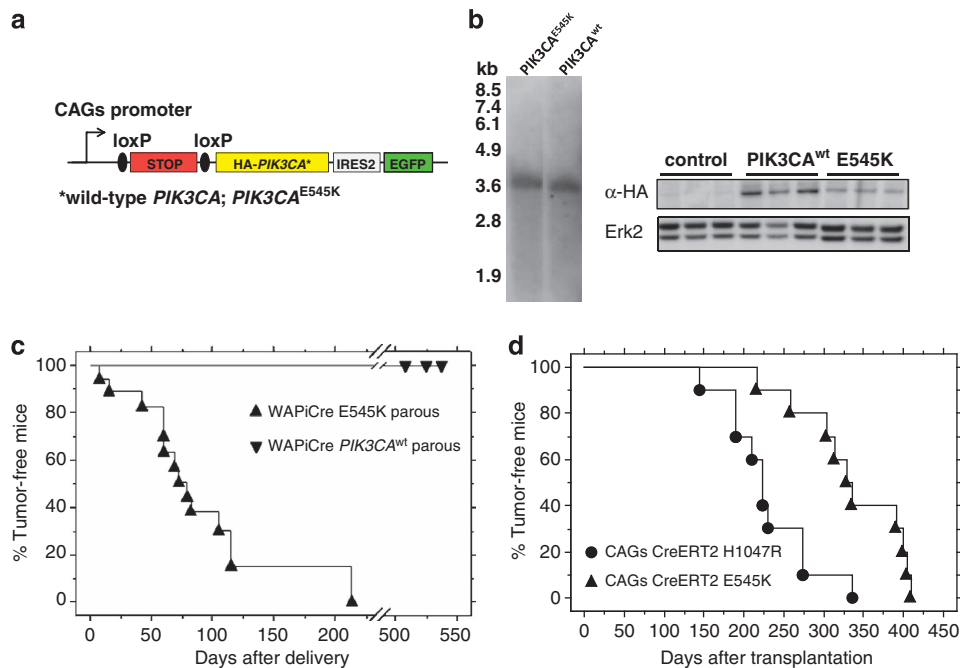
Figure 1. Overexpression of PIK3CA mutant PIK3CA[E545K] but not PIK3CA[wt] induces mammary tumors. (**a**) Schematic of the constructs used for generating transgenic mice conditionally expressing human wild-type and mutant 5′-terminally HA-tagged PIK3CA. Vectors were constructed in which the PIK3CA cDNA is flanked by a floxed STOP cassette upstream and an IRES2-EGFP reporter element downstream. The transgene is driven by a modified chicken β-actin (CAGs) promoter. The vector was introduced into a modified Rosa26 locus of Balb/c mouse embryonic stem cells by recombinase-mediated cassette exchange. (**b**) Southern blot of genomic DNA from PIK3CA[E545K] and PIK3CA[wt] transgenic mice (left) and immunoblots of lysates from mammary glands isolated 12 h after onset of involution from WAPiCre control, WAPiCre PIK3CA[wt] and WAPiCre PIK3CA[E545K] mice (each n = 3) probed for HA. Erk2 levels were used as a control for equal loading (right). (**c**) Kaplan–Meier plot showing tumor onset in parous WAPiCre PIK3CA[wt] (n = 8) and WAPiCre E545K (n = 16) mice. The mice were impregnated and the pups weaned 1 day after delivery. WAPiCre PIK3CA[wt] mice did not develop palpable tumors within 520 days whereas mice expressing PIK3CA[E545K] developed tumors on average 80 ($\pm$10) days after delivery. (**d**) Seven-week-old CAGs-CreERT2 PIK3CA[E545K] and PIK3CA[H1047R] donor mice were treated with tamoxifen on 3 consecutive days for transgene induction and fragments of glands were transplanted into cleared fat pads of three-week-old Balb/c recipient mice. Kaplan–Meier curves show tumor onset in recipient Balb/c mice transplanted with CAGs-CreERT2 PIK3CA[E545K]- (n = 10) or CAGS-CreERT2 PIK3CA[H1047R]-derived mammary glands (n = 10). Balb/c mice developed palpable tumors on average 336 ($\pm$20) days (PIK3CA[E545K]) or 229 ($\pm$17) days (PIK3CA[H1047R]) after transplantation; P = 0.0033.

WAPiCre PIK3CA[E545K]-evoked mammary tumors are heterogeneous
Examination of 30 WAPiCre PIK3CA[E545K]-derived tumors identified 6 distinct histotypes. By far the most prevalent tumor phenotype was adenosquamous carcinoma (60%) (Figures 2a and b), which was also the most common histotype formed by WAPiCre PIK3CA[H1047R] mice (54.6%).[22] Adenocarcinomas (23.3%) and carcinomas (6.7%) were also observed albeit at lower frequencies (Figures 2a and b). An adenocarcinoma with squamous metaplasia (3.3%), an adenomyoepithelioma (3.3%) and spindle cell tumor (3.3%) were observed in one tumor only (Figures 2a and b). The low frequency of adenomyoepithelioma in WAPiCre PIK3CA[E545K] mice is in stark contrast to the WAPiCre PIK3CA[H1047R] animals, in which adenomyoepitheliomas accounted for ∼23% of the tumors.[22] A further discrepancy between mice expressing PIK3CA[E545K] or PIK3CA[H1047R] was the complete absence of diffuse and invasive adenocarcinomatosis in WAPiCre PIK3CA[E545K]-derived glands, a histological feature that was displayed by all tumor-surrounding tissue in WAPiCre PIK3CA[H1047R] mice.[22]

The PIK3CA[E545K]-induced tumors were stained for luminal cytokeratin 18 (K18), basal/myoepithelial cytokeratin 14 (K14), and myoepithelial α-smooth muscle actin (α-SMA) markers. The most frequent histotypes, adenosquamous carcinoma and adenocarcinoma, were positive for both luminal K18 and basal K14 (Figure 2c). In tumors of the adenosquamous carcinoma type, the relative tumor areas positive for K18 and K14 were ∼35% and ∼39%, respectively (Figure 2d) and largely negative for α-SMA (<1%) (Figure 2d). WAPiCre PIK3CA[E545K]-evoked adenosquamous carcinomas also stained positive for ER (∼8% of the tumor cells) and displayed a high proportion of Ki-67-positive cells (∼35%) (Figures 2c and d). The relative tumor areas and cells positive for K14, K18, α-SMA, ER and Ki-67 were very similar to those observed in PIK3CA[H1047R]-driven adenosquamous carcinomas.[22] However, the number of apoptotic cells staining positively for cleaved

Figure 2. WAPiCre PIK3CA[E545K]-evoked tumors are heterogeneous and express basal and luminal cytokeratins. (**a**) Diagram showing relative abundance of adenosquamous carcinoma (60%, red), adenocarcinoma (23.3%, green), carcinoma (6.7%, purple), adenocarcinoma with squamous metaplasia (3.3%, dark blue), adenomyoepithelioma (3.3%, light blue) and spindle cell tumor (3.3%, orange) among tumors (n = 30) from parous WAPiCre PIK3CA[E545K] mice. (**b**) H&E-stained tumor sections of the indicated histotypes from parous WAPiCre PIK3CA[E545K] mice. The top-left image shows a representative adenosquamous carcinoma with glands and squamous features. The top-right image shows an adenocarcinoma; the arrows indicate the gland lumen. The glands are lined by malignant epithelium. The center-left image shows an adenocarcinoma with squamous metaplasia; the asterisk shows an area with glands and the arrow indicates areas of metaplasia. The center-right image shows a spindle cell tumor with possible osseous metaplasia, intense pink stroma and large cells in the interstices. The bottom image shows an adenomyoepithelioma. Scale bar = 100 μm. (**c**) Immunostaining for K14, K18, α-SMA, ER, Ki67 and cleaved caspase-3. Scale bars = 50 μm. (**d**) Quantification of immunostaining for K18, K14, α-SMA, ER, Ki67 and cleaved caspase-3 isolated from parous WAPiCre PIK3CA[E545K]-evoked adenosquamous carcinomas (n = 8). The data are presented as percentages of positive tumor area and tumor cells. Histological features of WAPiCre PIK3CA[E545K]-evoked adenosquamous carcinomas are compared with those of WAPiCre PIK3CA[H1047R]-evoked adenosquamous carcinomas previously reported.[22]

caspase-3 was higher in *PIK3CA* *PIK3CA*[E545K]-evoked adeno-squamous carcinomas ($\sim$7%) (Figures 2c and d) than those derived from WAPiCre H1047R mice ($\sim$1%),[22] indicating that *PIK3CA*[H1047R] is a more potent suppressor of apoptosis than *PIK3CA*[E545K] in mammary tumors.

In summary, both *PIK3CA*[E545K] and *PIK3CA*[H1047R] produced K14/K18-positive tumors of various histotypes, with the adenosquamous carcinoma type being the most common in both transgenic models. However, differences between the mouse models included low abundance of adenomyoepitheliomas and the absence of adenocarcinomatosis in WAPiCre *PIK3CA*[E545K] mice.

The variations in tumor histotypes and the discrepancy in tumor latency in WAPiCre *PIK3CA*[E545K] and *PIK3CA*[H1047R] mice suggest that different mechanisms underlie tumor initiation by these mutants. To gain a mechanistic insight that might explain these differences, we investigated whether pregnancy accelerates tumor onset in WAPiCre *PIK3CA*[E545K] as it does in WAPiCre *PIK3CA*[H1047R] mice.[22] Pregnancy accelerated tumor onset in WAPiCre *PIK3CA*[E545K] mice, reducing latency from 228 ± 15 days in nulliparous to 165 ± 10 days in parous mice (Figure 3a). Interestingly, pregnancy appeared to accentuate the difference in tumor

latency between WAPiCre *PIK3CA*[E545K] and *PIK3CA*[H1047R] mice, shown by 32 days difference in nulliparous vs 48 days difference in parous mice (Figure 3a).[22] We showed previously that a pregnancy-induced delay in mammary gland involution accounts, at least in part, for accelerated tumor kinetics in parous vs nulliparous *PIK3CA*[H1047R] mice.[22] Thus, we hypothesized here that the longer tumor latency of parous WAPiCre *PIK3CA*[E545K] compared with parous WAPiCre *PIK3CA*[H1047R] animals is the result of a less-pronounced involution delay. Comparison of WAPiCre *PIK3CA*[E545K] and WAPiCre *PIK3CA*[H1047R] glands 15 days after weaning revealed a dramatic delay in involution compared with control animals (Figure 3b). The relative gland area occupied by epithelial cells was the same in WAPiCre *PIK3CA*[E545K] and *PIK3CA*[H1047R] mice and significantly larger than in WAPiCre control glands (Figure 3c). Similarly, there was no difference in the number of apoptotic or proliferating cells in glands expressing either of the *PIK3CA* mutations (Figure 3d). Interestingly, glands from WAPiCre *PIK3CA*[wt] mice, which did not form tumors, displayed normal involution and numbers of apoptotic and proliferating cells similar to the controls (Figure 3), indicating that the delay in involution is caused by mutant *PIK3CA* rather than by
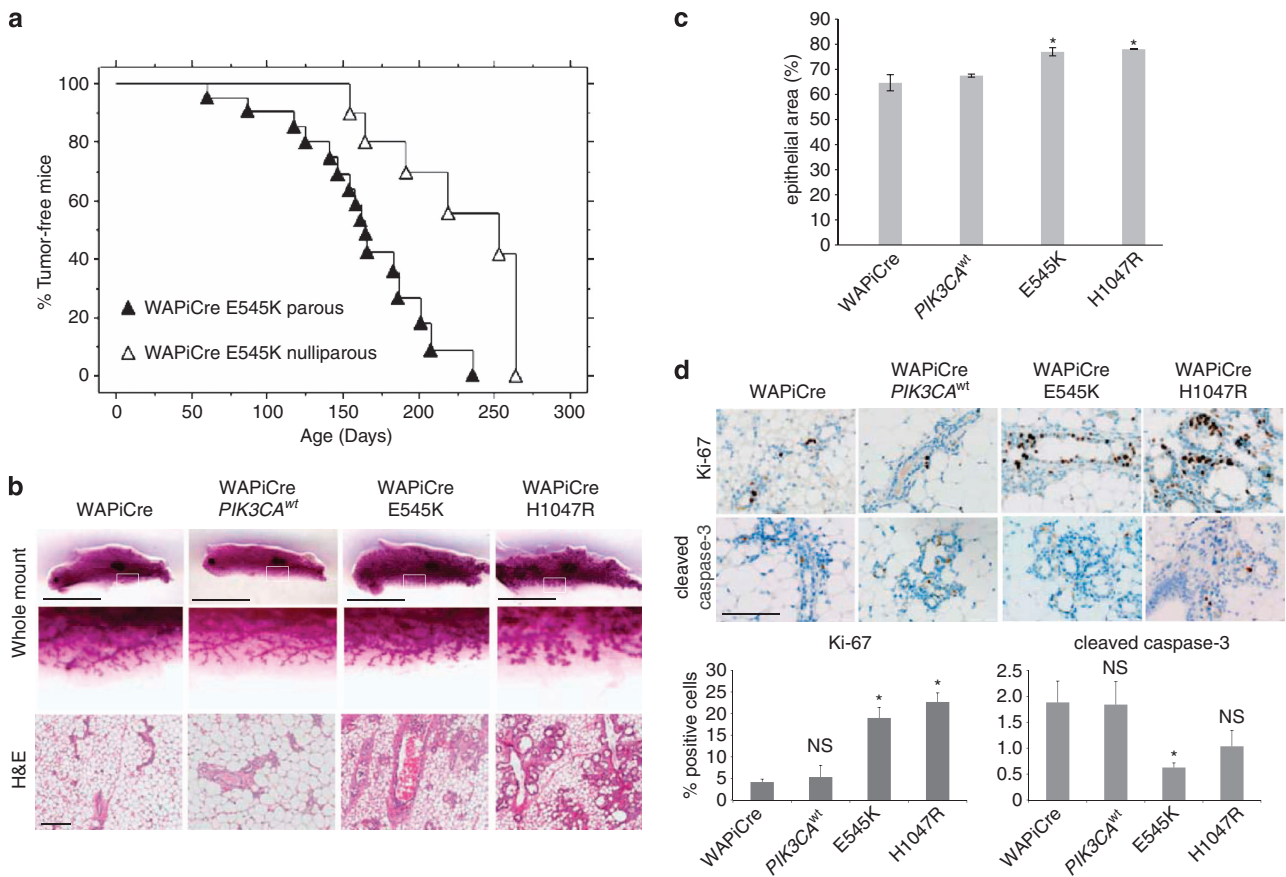
Figure 3. Pregnancy accelerates *PIK3CA*-evoked tumorigenesis and *PIK3CA* mutants delay mammary gland involution. (**a**) Kaplan–Meier curves showing tumor onset in parous WAPiCre *PIK3CA*[E545K] ($n = 16$) and nulliparous WAPiCre *PIK3CA*[E545K] ($n = 6$) animals. Parous WAPiCre *PIK3CA*[E545K] mice developed palpable tumors on average after 165 (± 10) days whereas nulliparous mice developed tumors on average after 228 (± 15) days (*PIK3CA*[E545K]); $P = 0.0023$. (**b**) Representative images of whole mount (*top panels*), magnification of whole mount (*center panels*) and H&E (*lower panels*) staining of involuting glands from WAPiCre control, WAPiCre *PIK3CA*[wt], WAPiCre *PIK3CA*[E545K] and WAPiCre *PIK3CA*[H1047R] mice as indicated. The glands were isolated 15 days after removal of the pups. Scale bar = 1 cm (whole mounts). Scale bar = 100 μm (H&E sections). (**c**) Bar graph showing relative epithelium to total gland area of involution at day 15 in whole mounts prepared from WAPiCre control ($n = 3$), WAPiCre *PIK3CA*[wt] ($n = 4$), WAPiCre *PIK3CA*[E545K] ($n = 4$) and WAPiCre *PIK3CA*[H1047R] ($n = 2$). Means $\pm$ s.d. are shown; $P = 0.001$ (WAPiCre vs WAPiCre *PIK3CA*[E545K]); $P = 0.01$ (WAPiCre vs WAPiCre *PIK3CA*[H1047R]); $P = 0.46$ (WAPiCre *PIK3CA*[E545K] vs WAPiCre *PIK3CA*[H1047R]). (**d**) Immunostaining for Ki67 and cleaved caspase-3 of day 15 involuting glands from WAPiCre control, WAPiCre *PIK3CA*[E545K] and WAPiCre[H1047R] mice (*upper panels*). Scale bar = 50 μm. Quantification of Ki67- and cleaved caspase-3-positive cells (*lower panels*). Means $\pm$ s.e.m. are shown. *For Ki67-positive cells: $P = 2.18 \times 10^{-5}$ (WAPiCre vs WAPiCre *PIK3CA*[E545K]), $P = 2.90 \times 10^{-7}$ (WAPiCre vs WAPiCre *PIK3CA*[H1047R]). For cleaved caspase-3-positive cells: $P = 6.55 \times 10^{-3}$ (WAPiCre vs WAPiCre *PIK3CA*[E545K]). NS = not significant.
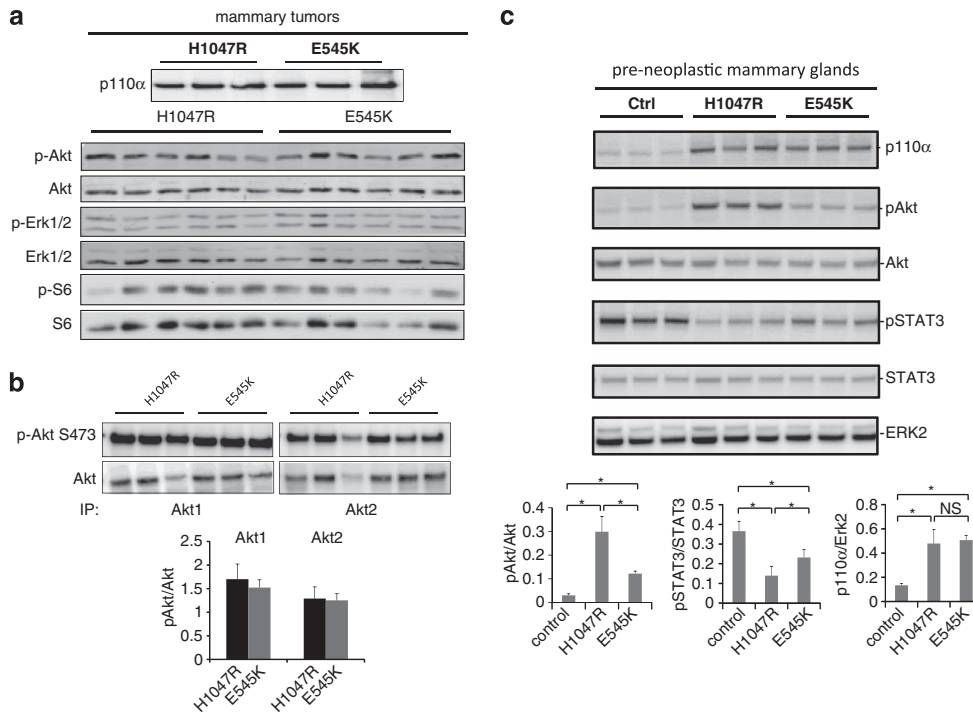
Figure 4. WAPiCre *PIK3CA*[E545K] involuting glands show reduced pAkt and increased pSTAT3 compared with WAPiCre *PIK3CA*[H1047R] at 12 h of involution. (**a**) Immunoblots of lysates from WAPiCre *PIK3CA*[E545K] and WAPiCre *PIK3CA*[H1047R] mammary tumors probed for the indicated proteins. (**b**) Lysates of mammary tumors from WAPiCre *PIK3CA*[E545K] and WAPiCre *PIK3CA*[H1047R] mice were first immunoprecipitated with antibodies against either Akt1 or Akt2 and then probed for total Akt or S473 pAkt (*upper panels*). Bar graph showing relative levels of S473 pAkt normalized to total Akt in Akt1 or Akt2 immunoprecipitates (*lower panel*). (**c**) Lysates of mammary glands isolated 12 h after onset of involution from WAPiCre control, WAPiCre *PIK3CA*[H1047R] and WAPiCre *PIK3CA*[E545K] mice (each $n = 3$) probed for p110α, pAkt, Akt, pSTAT3, STAT3 and ERK2 as a loading control (*upper panels*). Bar graphs showing relative amounts of pAkt (normalized to total Akt), pSTAT3 (normalized to total STAT3) and p110α (normalized to ERK2) in lysates of WAPiCre control, WAPiCre *PIK3CA*[H1047R] and WAPiCre *PIK3CA*[E545K] mammary glands (*lower panels*). *WAPiCre vs WAPiCre *PIK3CA*[H1047R]: For pAkt $P = 3.8 \times 10^{-5}$; for pSTAT3 $P = 0.002$; for p110α $P = 0.007$. WAPiCre vs WAPiCre *PIK3CA*[E545K]: For pAkt $P = 1.8 \times 10^{-4}$; for pSTAT3 $P = 0.02$; for p110α $P = 8.8 \times 10^{-5}$. WAPiCre *PIK3CA*[H1047R] vs WAPiCre *PIK3CA*[E545K]: for pAkt $P = 2.5 \times 10^{-4}$; for pSTAT3 $P = 0.047$; for p110α $P = 0.7$. NS = not significant.

overexpression of the transgene. In summary, *PIK3CA*[E545K] and *PIK3CA*[H1047R] transgene expression caused a dramatic but comparable delay in involution and, therefore, involution does not explain the different tumor kinetics observed in parous vs nulliparous mice expressing these mutations.

Comparison of lysates from WAPiCre *PIK3CA*[E545K]- and WAPiCre *PIK3CA*[H1047R]-derived mammary glands and tumors showed equal expression of p110α in tumors from both transgenic models (Figure 4a). Despite the enhanced oncogenic potential of the PIK3CA[H1047R] mutant, no differences in activation of the PI3K/Akt or the Erk pathways were observed in the tumors (Figure 4a). Similarly, a more detailed analysis of Akt1 and Akt2 isoform-specific phosphorylation revealed no difference between *PIK3CA*[E545K]- and WAPiCre *PIK3CA*[H1047R]-induced signaling (Figure 4b). Conceivably, by the time mammary tumors were established, numerous secondary mutations had resulted in a tumor heterogeneity that compromises the detection of potentially subtle differences in oncogenic signaling induced by either *PIK3CA* mutant. To circumvent this, we investigated molecular signaling events in mutant *PIK3CA*-expressing epithelial cells at an early pre-neoplastic stage. Protein lysates from mammary glands isolated 12 h after the onset of involution revealed increased activation of Akt and decreased phosphorylation of the signal transducer and activator of transcription (STAT) 3 in mutant relative to control glands. Notably, both hyperactivation of Akt and hypoactivation of STAT3 were more pronounced in *PIK3CA*[H1047R] than in *PIK3CA*[E545K] glands (for pAkt $P = 2.5 \times 10^{-4}$; for pSTAT3 $P = 0.047$) (Figure 4c).

In summary, we found that overexpression of *PIK3CA*[E545K] in a transgenic mouse model potently induces heterogeneous mammary tumors whereas overexpression of wild-type *PIK3CA* does not. Notably, although *PIK3CA*[E545K] evokes tumors with 100% penetrance it is a weaker inducer of mammary tumors than *PIK3CA*[H1047R] in two independent mouse models in which mutant *PIK3CA* is either driven by the WAP or by the CAGs promoter. This may explain the lower frequency of helical vs kinase domain mutations in human breast cancer.[13] We found differences in Akt and STAT3 activation in pre-neoplastic mammary glands from *PIK3CA*[E545K] and *PIK3CA*[H1047R] transgenic mice that may explain the longer tumor latency observed in WAPiCre *PIK3CA*[E545K] compared with WAPiCre *PIK3CA*[H1047R] mice.

The novel transgenic mouse models reported here provide excellent tools to further dissect the activities of different *PIK3CA* mutants in tumor initiation *in vivo* and to investigate drug responses to the ever-increasing number of PI3K pathway inhibitors.

MATERIALS AND METHODS

Transgenic mice
We constructed a vector with a transcriptional STOP sequence flanked by *loxP* sites upstream of the 5'-terminally HA-tagged human *PIK3CA* cDNA (Addgene, Cambridge, MA, USA) and an *IRES2-EGFP* reporter element (pIRES2-EGFP vector; Clontech, Mountain View, CA, USA). The resulting construct was introduced into the modified Rosa26 locus of Balb/c mouse embryonic stem cells by recombinase-mediated cassette exchange as described earlier.[22] Chimeric mice were backcrossed to Balb/c mice and transgenic mice identified by genotyping.[22]

Immunoblotting

Protein lysates were extracted from inguinal mammary glands or tumors using LB buffer (50 mM Tris–HCl pH8, 150 mM NaCl, 1% NP-40) supplemented with 0.5 mM sodium orthovanadate. Anti-p110α, anti-pAKT (Ser473), anti-Akt, anti-pERK1/2 (Thr202/Tyr204), anti-ERK1/2, anti-pS6 (Ser235/236), anti-S6, anti-Akt1, anti-Akt2, anti-pSTAT3 (Tyr705) and anti-STAT3 antibodies were purchased from Cell Signaling Technology, Danvers, MA, USA.

Immunohistochemistry

The following antibodies were used: K14 (Thermo Scientific, , Waltham, MA, USA, RB-9020, 1:100), K18 (Fitzgerald, Acton, MA, USA, #GP11, 1:200), ER (Santa Cruz, , Dallas, TX, USA, SC-542, 1:1000), α-SMA (Thermo Scientific, RB-9010, 1:500), cleaved caspase-3 (Cell Signaling, #9661, 1:100) Ki-67 (Thermo Scientific, RB-9106, 1:50).

Southern blotting

Genomic DNA from mouse tails was digested with 8 U of AvrII enzyme (New England BioLabs (NEB), Ipswich, MA, USA) and separated on a 1% agarose gel. A DIG-labeled DNA probe targeting the neomycin resistance cassette was amplified using the PCR DIG Probe Synthesis Kit (Roche, Basel, Switzerland) and the primers 5′-ATGGGATCGGCCATTGAACAAGAT-3′ and 5′-CGGCCATTTTCCACCATGATAT-3′.

REFERENCES

1 Yuan TL, Cantley LC. PI3K pathway alterations in cancer: variations on a theme. *Oncogene* 2008; **27**: 5497–5510.
2 Manning BD, Cantley LC. AKT/PKB signaling: navigating downstream. *Cell* 2007; **129**: 1261–1274.
3 Bachman KE, Argani P, Samuels Y, Silliman N, Ptak J, Szabo S *et al*. The PIK3CA gene is mutated with high frequency in human breast cancers. *Cancer Biol Ther* 2004; **3**: 772–775.
4 Kadota M, Sato M, Duncan B, Ooshima A, Yang HH, Diaz-Meyer N *et al*. Identification of novel gene amplifications in breast cancer and coexistence of gene amplification with an activating mutation of PIK3CA. *Cancer Res* 2009; **69**: 7357–7365.
5 Levine DA, Bogomolniy F, Yee CJ, Lash A, Barakat RR, Borgen PI *et al*. Frequent mutation of the PIK3CA gene in ovarian and breast cancers. Clinical cancer research: an official journal of the American Association for Cancer Research 2005; **11**: 2875–2878.
6 Samuels Y, Wang Z, Bardelli A, Silliman N, Ptak J, Szabo S *et al*. High frequency of mutations of the PIK3CA gene in human cancers. Science 2004; **304**: 554.
7 Wu G, Xing M, Mambo E, Huang X, Liu J, Guo Z *et al*. Somatic mutation and gain of copy number of PIK3CA in human breast cancer. *Breast Cancer Res* 2005; **7**: R609–R616.
8 Hafsi S, Pezzino FM, Candido S, Ligresti G, Spandidos DA, Soua Z *et al*. Gene alterations in the PI3K/PTEN/AKT pathway as a mechanism of drug-resistance (review). *Int J Oncol* 2012; **40**: 639–644.
9 Bader AG, Kang S, Zhao L, Vogt PK. Oncogenic PI3K deregulates transcription and translation. *Nat Rev Cancer* 2005; **5**: 921–929.
10 Bader AG, Kang S, Vogt PK. Cancer-specific mutations in PIK3CA are oncogenic in vivo. *Proc Natl Acad Sci USA* 2006; **103**: 1475–1479.
11 Isakoff SJ, Engelman JA, Irie HY, Luo J, Brachmann SM, Pearline RV *et al*. Breast cancer-associated PIK3CA mutations are oncogenic in mammary epithelial cells. *Cancer Res* 2005; **65**: 10992–11000.
12 Zhao JJ, Liu Z, Wang L, Shin E, Loda MF, Roberts TM. The oncogenic properties of mutant p110alpha and p110beta phosphatidylinositol 3-kinases in human mammary epithelial cells. *Proc Natl Acad Sci USA* 2005; **102**: 18443–18448.
13 Koren S, Bentires-Alj M. Mouse models of PIK3CA mutations: one mutation initiates heterogeneous mammary tumors. *FEBS J* 2013; **280**: 2758–2765.
14 Zhao L, Vogt PK. Helical domain and kinase domain mutations in p110alpha of phosphatidylinositol 3-kinase induce gain of function by different mechanisms. *Proc Natl Acad Sci USA* 2008; **105**: 2652–2657.
15 Barbareschi M, Buttitta F, Felicioni L, Cotrupi S, Barassi F, Del Grammastro M *et al*. Different prognostic roles of mutations in the helical and kinase domains of the PIK3CA gene in breast carcinomas. *Clin Cancer Res* 2007; **13**: 6064–6069.
16 Miller TW. Initiating breast cancer by PIK3CA mutation. *Breast Cancer Res* 2012; **14**: 301.
17 Saal LH, Holm K, Maurer M, Memeo L, Su T, Wang X *et al*. PIK3CA mutations correlate with hormone receptors, node metastasis, and ERBB2, and are mutually exclusive with PTEN loss in human breast carcinoma. *Cancer Res* 2005; **65**: 2554–2559.
18 Miron A, Varadi M, Carrasco D, Li H, Luongo L, Kim HJ *et al*. PIK3CA mutations in in situ and invasive breast carcinomas. *Cancer Res* 2010; **70**: 5674–5678.
19 Kalinsky K, Jacks LM, Heguy A, Patil S, Drobnjak M, Bhanot UK *et al*. PIK3CA mutation associates with improved outcome in breast cancer. *Clin Cancer Res* 2009; **15**: 5049–5059.
20 Lai YL, Mau BL, Cheng WH, Chen HM, Chiu HH, Tzen CY. PIK3CA exon 20 mutation is independently associated with a poor prognosis in breast cancer patients. *Ann Surg Oncol* 2008; **15**: 1064–1069.
21 Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* 2012; **490**: 61–70.
22 Meyer DS, Brinkhaus H, Muller U, Muller M, Cardiff RD, Bentires-Alj M. Luminal expression of PIK3CA mutant H1047R in the mammary gland induces heterogeneous tumors. *Cancer Res* 2011; **71**: 4344–4351.
23 Adams JR, Xu K, Liu JC, Agamez NM, Loch AJ, Wong RG *et al*. Cooperation between Pik3ca and p53 mutations in mouse mammary tumor formation. *Cancer Res* 2011; **71**: 2706–2717.
24 Liu P, Cheng H, Santiago S, Raeder M, Zhang F, Isabella A *et al*. Oncogenic PIK3CA-driven mammary tumors frequently recur via PI3K pathway-dependent and PI3K pathway-independent mechanisms. *Nat Med* 2011; **17**: 1116–1120.
25 Tikoo A, Roh V, Montgomery KG, Ivetac I, Waring P, Pelzer R *et al*. Physiological levels of Pik3ca(H1047R) mutation in the mouse mammary gland results in ductal hyperplasia and formation of ERalpha-positive tumors. *PLoS One* 2012; **7**: e36924.
26 Tchorz JS, Kinter J, Muller M, Tornillo L, Heim MH, Bettler B. Notch2 signaling promotes biliary epithelial cell fate specification and tubulogenesis during bile duct development in mice. *Hepatology* 2009; **50**: 871–879.
27 Wintermantel TM, Mayer AK, Schutz G, Greiner EF. Targeting mammary epithelial cells using a bacterial artificial chromosome. *Genesis* 2002; **33**: 125–130.
28 Boulanger CA, Wagner KU, Smith GH. Parity-induced mouse mammary epithelial cells are pluripotent, self-renewing and sensitive to TGF-beta1 expression. *Oncogene* 2005; **24**: 552–560.
29 Booth BW, Boulanger CA, Smith GH. Alveolar progenitor cells develop in mouse mammary glands independent of pregnancy and lactation. *J Cell Physiol* 2007; **212**: 729–736.
30 Bruno RD, Smith GH. Functional characterization of stem cell activity in the mouse mammary gland. *Stem Cell Rev* 2011; **7**: 238–247.

LETTER

$PIK3CA^{H1047R}$ induces multipotency and multi-lineage mammary tumours

Shany Koren[1], Linsey Reavie[1], Joana Pinto Couto[1], Duvini De Silva[1], Michael B. Stadler[1,2], Tim Roloff[1], Adrian Britschgi[1], Tobias Eichlisberger[1], Hubertus Kohler[1], Olulanu Aina[3], Robert D. Cardiff[3] & Mohamed Bentires-Alj[1]

The adult mouse mammary epithelium contains self-sustained cell lineages that form the inner luminal and outer basal cell layers, with stem and progenitor cells contributing to its proliferative and regenerative potential[1–4]. A key issue in breast cancer biology is the effect of genomic lesions in specific mammary cell lineages on tumour heterogeneity and progression. The impact of transforming events on fate conversion in cancer cells of origin and thus their contribution to tumour heterogeneity remains largely elusive. Using *in situ* genetic lineage tracing and limiting dilution transplantation, we have unravelled the potential of $PIK3CA^{H1047R}$, one of the most frequent mutations occurring in human breast cancer[5], to induce multipotency during tumorigenesis in the mammary gland. Here we show that expression of $PIK3CA^{H1047R}$ in lineage-committed basal Lgr5-positive and luminal keratin-8-positive cells of the adult mouse mammary gland evokes cell dedifferentiation into a multipotent stem-like state, suggesting this to be a mechanism involved in the formation of heterogeneous, multi-lineage mammary tumours. Moreover, we show that the tumour cell of origin influences the frequency of malignant mammary tumours. Our results define a key effect of $PIK3CA^{H1047R}$ on mammary cell fate in the pre-neoplastic mammary gland and show that the cell of origin of $PIK3CA^{H1047R}$ tumours dictates their malignancy, thus revealing a mechanism underlying tumour heterogeneity and aggressiveness.

The mammary gland epithelium is composed of two major cell lineages: the luminal layer contains cells expressing keratin 8/18 (K8/18) and the basal layer with cells expressing K5/14 and/or smooth muscle actin (SMA) and/or p63 (ref. 6).

Multipotent cells that generate both the luminal and basal lineages are found in the mouse embryonic mammary gland[1,7] but their existence in the adult gland is still under debate. Studies using serial transplantation into cleared mammary fat pad supposed the existence in the adult mouse mammary gland of multipotent stem cells with myoepithelial features[8–11]. Arguably, these assays reflected the regenerative potential of the transplanted cells rather than their properties *in situ*[1–3]. Lineage-tracing studies, which permit targeted expression of a fluorescent reporter in a given cell and its progeny, showed that tissue homeostasis is maintained by unipotent luminal K8/18-positive and basal K5/14/Lgr5-positive stem cells after birth. Lineage tracing of K8/18, K5/14 and Lgr5 progeny found no evidence for the presence of multipotent stem cells in the adult mammary gland[1] but did not exclude the possibility that rare cells not targeted by these reporters, or only at a very low frequency, have multipotent potential. While tracing of the progeny of axin-2-positive cells showed the presence of multipotent stem cells during puberty and pregnancy, this and other studies revealed that the basal and luminal lineages are self-sustained in the adult virgin gland[2–4]. By contrast, recent three-dimensional wholemount imaging[12] and the identification of the Procr-positive subset[13] argue for the presence of multipotent stem cells in the adult virgin mouse mammary gland, thus reopening the debate.

The phosphatidylinositol 3-kinase (PI3K) pathway is activated in ~70% of breast cancers. Several mechanisms may account for the activation of this pathway in cancer, including amplification and/or activating mutations of the *PIK3CA* gene that encodes the p110α catalytic subunit of PI3K found in 20–40% of breast cancers[5,14]. The most recurrent mutation, H1047R, leads to constitutive PI3K signalling and heterogeneous mammary tumours[15–17]. Despite frequent alterations of the PI3K pathway in breast cancer, its impact on lineage organization during tumorigenesis and the importance of the cell of origin for heterogeneity and aggressiveness of PI3K-driven tumours has remained unclear.

To address the effects of mutant $PIK3CA^{H1047R}$ on basal- or luminal-lineage-restricted cells, we performed lineage tracing in adult Lgr5-CreER[T2]/Tomato-reporter and K8-CreER[T2]/Tomato-reporter mice with or without $PIK3CA^{H1047R}$ (Extended Data Fig. 1a)[16]. Moreover, we used Lgr5- and K8-CreER[T2]/$PIK3CA^{H1047R}$ or *PIK3CA* wild-type ($PIK3CA^{WT}$) animals[16,18] for tracing of the green fluorescent protein (GFP) reporter, and Lgr5-CreER[T2] and K8-CreER[T2] mice as controls (Extended Data Fig. 1b). As previously reported[1,19], we found Lgr5 activity only in a subset of basal cells in the nipple area of the mammary gland (Extended Data Fig. 2a–e). We further assessed the effects of $PIK3CA^{H1047R}$ expression on the distribution of mammary subpopulations by fluorescence-activated cell sorting (FACS) on isolated mammary epithelial cells labelled with CD24 and Sca1, which were shown to enrich for luminal (CD24[Hi]Sca1[−], CD24[Hi]Sca1[+]) and basal (CD24[Lo]Sca1[−]) cells[10,20] (Extended Data Fig. 3). Four days lineage tracing confirmed Tomato labelling of basal cells in Lgr5-CreER[T2]/Tomato and Lgr5-CreER[T2]/$PIK3CA^{H1047R}$/Tomato and of luminal cells in K8-CreER[T2]/Tomato and K8-CreER[T2]/$PIK3CA^{H1047R}$/Tomato animals (Extended Data Figs 2b, f and 4a, b). We further performed 4-, 8- and 13-week lineage tracing. The distribution of Tomato-labelled subsets in Lgr5-CreER[T2]/Tomato glands did not change with time where the progeny of Lgr5-positive cells was mostly of basal origin (Fig. 1a, b and Extended Data Fig. 2g, h). In K8-CreER[T2]/Tomato mice, labelling was restricted to the luminal subset, marking mostly mature CD24[Hi]Sca1[+] in 4 weeks and mostly CD24[Hi]Sca1[−] luminal progenitors in 13 weeks tracing, indicating the targeting of a long-term unipotent luminal subset (Fig. 1a, c and Extended Data Fig. 4c, d). By contrast, expression of $PIK3CA^{H1047R}$ in Lgr5-CreER[T2]/$PIK3CA^{H1047R}$/Tomato and K8-CreER[T2]/$PIK3CA^{H1047R}$/Tomato mice resulted in labelling of both the luminal and the basal compartments (Fig. 1b, c and Extended Data Figs 2g, h, 4c, d). $PIK3CA^{H1047R}$-evoked multi-lineage labelling was not observed at 4–7 days after tamoxifen induction (Extended Data Figs 2f, i and 4b). Since $PIK3CA^{H1047R/WT}$ targeting vectors contain an internal ribosome entry site (IRES)–GFP construct, we also used GFP as a readout of transgene expression. At 4 days after tamoxifen induction, K8-CreER[T2]/$PIK3CA^{WT}$ and K8-CreER[T2]/$PIK3CA^{H1047R}$ glands expressed similar levels of GFP, indicating similar Cre recombination efficiency. In both models, 4- and 8–11-week lineage tracing revealed an increase in GFP-labelled
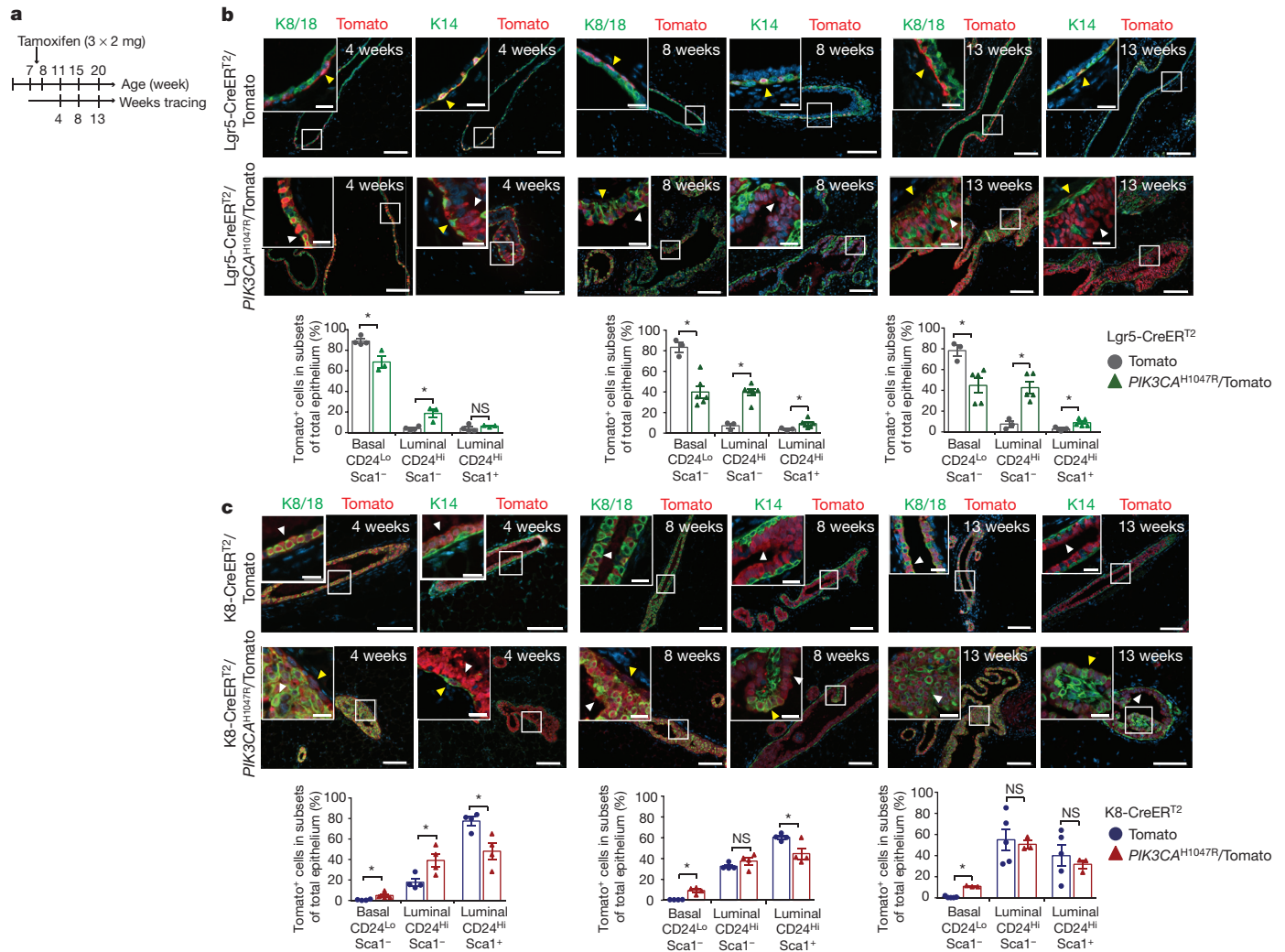
Figure 1 | Mutant *PIK3CA* induces mammary cell plasticity. a, Timeline for lineage-tracing studies. **b, c**, Representative images of 4-, 8- and 13-week tracing and FACS quantification of Tomato-positive epithelial basal (CD24^{Lo}Sca1$^-$) and luminal (CD24^{Hi}Sca1$^{-/+}$) subsets from Lgr5-CreER^{T2}/Tomato (**b**, immunofluorescence: left $n = 3$, middle $n = 11$, right $n = 6$ mice; FACS: left $n = 4$ technical replicates (each 1–3 pooled mice), middle $n = 3$ technical replicates (each 1–2 pooled mice), right $n = 3$ technical replicates (each 1 mouse)), Lgr5-CreER^{T2}/*PIK3CA*^{H1047R}/Tomato (**b**, immunofluorescence: left $n = 3$, middle $n = 10$, right $n = 4$ mice; FACS: left $n = 3$ technical replicates (each 1–2 pooled mice), middle $n = 6$ technical replicates (each 1 mouse), right $n = 5$ technical replicates (each 1–2 pooled mice)), K8-CreER^{T2}/Tomato

(**c**, immunofluorescence: left $n = 4$, middle $n = 8$, right $n = 5$ mice; FACS: left $n = 4$ technical replicates (each 1–3 pooled mice), middle $n = 4$ technical replicates (each 1–2 pooled mice), right $n = 5$ technical replicates (each 1 mouse)) and K8-CreER^{T2}/*PIK3CA*^{H1047R}/Tomato animals (**c**, immunofluorescence: left $n = 4$, middle $n = 4$, right $n = 3$ mice; FACS: left $n = 4$ technical replicates (each 1–2 pooled mice), middle $n = 4$ technical replicates, right $n = 3$ technical replicates (each 1 mouse)). White arrowheads indicate luminal and yellow arrowheads indicate basal Tomato-labelled cells. Scale bars, 100 μm, 20 μm (magnifications). Bar graphs show means ± standard error of the mean (s.e.m.); two-sided unpaired Student's t-test; *$P < 0.05$; NS, not significant.

---

basal and luminal subsets in *PIK3CA*^{WT} compared with *PIK3CA*^{WT} animals that showed lineage-restricted GFP labelling, consistent with higher PI3K pathway activation (Extended Data Figs 5 and 6). These results suggest that expression of *PIK3CA*^{H1047R} in basal- or luminal-restricted mammary cells triggers lineage plasticity and cell expansion.

Next, we assessed the effects of *PIK3CA*^{H1047R} on global gene expression in the pre-neoplastic gland. Microarray and quantitative polymerase chain reaction with reverse transcription (qRT–PCR) analyses were performed on mammary epithelial cell subpopulations from Lgr5-CreER^{T2} control versus Lgr5-CreER^{T2}/*PIK3CA*^{H1047R} and K8-CreER^{T2} control versus K8-CreER^{T2}/*PIK3CA*^{H1047R} animals. Genes expressed differentially between subpopulations of control and mutant mice were compared with subpopulation signatures that have been previously described[21,22]. We found enrichment of luminal progenitor signature genes in the basal Lgr5-CreER^{T2}/*PIK3CA*^{H1047R} subset and in the newly formed basal K8-CreER^{T2}/*PIK3CA*^{H1047R} subset. Enrichment of myoepithelial signature genes was found in the newly

formed luminal Lgr5-CreER^{T2}/*PIK3CA*^{H1047R} and in the K8-CreER^{T2}/*PIK3CA*^{H1047R} luminal subsets (Fig. 2a, b and Extended Data Fig. 7). Mammary cells co-expressing basal and luminal markers in neoplastic areas confirmed that *PIK3CA*^{H1047R} induces cell plasticity (Fig. 2c).

We addressed how lineage plasticity is evoked by *PIK3CA*^{H1047R} in functional assays. Limiting dilution transplantation revealed an increase in the mammary-repopulating capacity of basal cells from Lgr5-CreER^{T2}/*PIK3CA*^{H1047R} and K8-CreER^{T2}/*PIK3CA*^{H1047R} mice compared with the respective controls (Fig. 3a). The GFP-negative basal CD24^{Lo}Sca1$^-$ population from K8-CreER^{T2} mice comprises the bulk of basal cells (myoepithelial and mammary-repopulating cells)[10,20], explaining the lack of outgrowths at the number of cells transplanted. *PIK3CA*^{H1047R}-expressing luminal cells also had repopulating capacity (Extended Data Fig. 8a, b). All outgrowths expressed luminal and basal markers (Fig. 3b and Extended Data Fig. 8c). In colony formation assays, *PIK3CA*^{H1047R} increased the percentage of double-positive (K14/K8/18) colonies derived from the newly formed
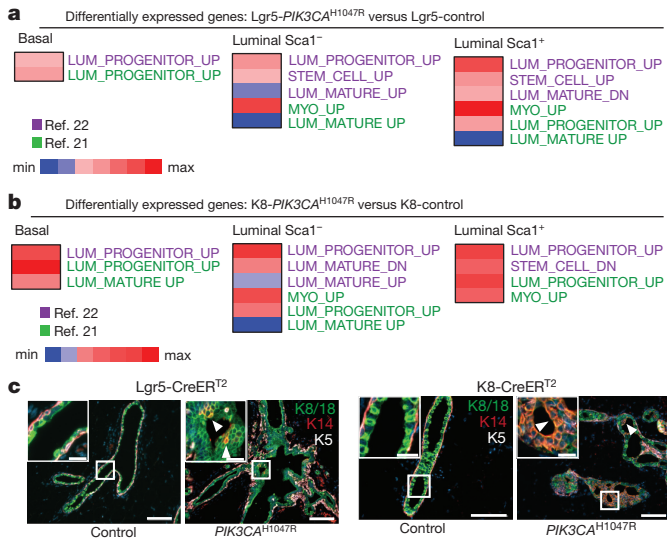
a Differentially expressed genes: Lgr5-*PIK3CA*[H1047R] versus Lgr5-control

b Differentially expressed genes: K8-*PIK3CA*[H1047R] versus K8-control

c

Figure 2 | Activation of *PIK3CA*[H1047R] leads to expression of basal- and luminal-lineage genes. **a**, **b**, Plots indicating enrichment of gene expression from FACS-sorted Lgr5-CreER[T2]/*PIK3CA*[H1047R] versus Lgr5-CreER[T2] control (**a**) and K8-CreER[T2]/*PIK3CA*[H1047R] versus K8-CreER[T2] control subsets (**b**) in signatures of mammary subpopulations from refs 21, 22. Microarray was performed 4 weeks after tamoxifen induction on basal CD24[Lo]Sca1[−] and luminal CD24[Hi]Sca1[−/+] subsets of pooled mammary glands of 2–3 oestrus-synchronized animals from three independent sortings. **c**, Representative images of immunostaining for basal (K14, red; K5, white) and luminal (K8/18, green) markers on mammary glands 4 weeks after tamoxifen treatment ($n = 3$ mice). White arrowheads indicate double-positive cells. Scale bars, 100 μm, 20 μm (magnifications).

luminal (Lgr5-CreER[T2]/*PIK3CA*[H1047R]) and basal cells (K8-CreER[T2]/*PIK3CA*[H1047R]). *PIK3CA*[H1047R]-expressing subsets overcame lineage restriction, giving rise to both lineages, albeit at low frequencies. Moreover, basal cells from Lgr5- and K8-CreER[T2]/*PIK3CA*[H1047R] animals showed increased colony formation capacity (Extended Data Fig. 8d–f). In mammosphere cultures, *PIK3CA*[H1047R] increased the sphere-forming capacity of luminal cells. While control cells formed spheres with a hollow lumen after passaging, *PIK3CA*[H1047R]-expressing cells formed filled spheres, indicating the accumulation of less differentiated cells (Extended Data Fig. 8g, h). Altogether, these data suggest that *PIK3CA*[H1047R] evokes cell dedifferentiation to a multipotent stem-like state, from which cells further differentiate into both cell lineages.

Expression of *PIK3CA*[H1047R] in Lgr5- and K8-positive cells induced mammary tumours on average after 108 and 78 days, respectively. Control and *PIK3CA*[WT] animals developed no tumours (Fig. 4a). FACS analysis of GFP-positive tumour cells revealed a similar distribution of cancer subpopulations, with an accumulation of the CD24[Hi]Sca1[−] subset. *PIK3CA*[H1047R]-evoked mammary tumours expressed basal and luminal markers (Extended Data Fig. 9a–d). Additionally, cells double positive for basal and luminal markers were found (Fig. 4b). These results suggest that *PIK3CA*[H1047R]-evoked cell plasticity results in multi-lineage tumours and that expression of basal and luminal markers is not an indicator of the origin of mammary cancers. It has been proposed that basal-like mammary tumours may originate from luminal cells[23,24]. Therefore, any inference of the cell of origin from the differentiation state of the tumour can be misleading.

Histological analysis showed heterogeneous phenotypes and differences in malignancy between both models. Lgr5-CreER[T2]/*PIK3CA*[H1047R] mice formed unique benign multi-nodular rosette-type adenomyoepitheliomas, aggressive adenosquamous carcinomas with pilosebaceous differentiation and carcinosarcomas. K8-CreER[T2]/*PIK3CA*[H1047R] mice mainly developed aggressive adenosquamous carcinoma, carcinosarcomas that infiltrated the surrounding tissue, adenocarcinomas and benign adenomyoepitheliomas. These results

a Lgr5-CreER[T2] outgrowths

	Cell no.	Control	*PIK3CA*[H1047R]
Basal CD24[Lo] Sca1[−]	50	5/17 (29%)	10/19 (52%)
	100	4/15 (26%)	9/15 (60%)
	200	5/13 (38%)	9/13 (69%)
MRU frequency (95% CI)		1 in 289 (1/169–1/496)	1 in 111 (1/74–1/166)

$P = 0.003$

K8-CreER[T2] outgrowths

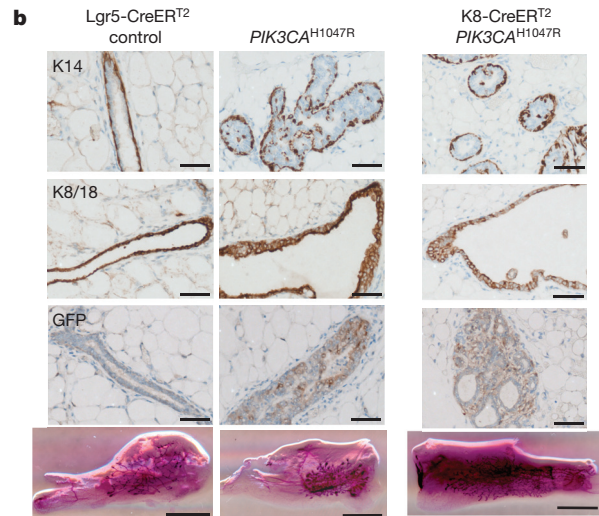|	Cell no.	Control	*PIK3CA*[H1047R]
Basal CD24[Lo] Sca1[−]	100	0/6 (0%)	4/10 (40%)
	200	0/8 (0%)	7/11 (63%)
	400	0/5 (0%)	2/4 (50%)
MRU frequency (95% CI)		N/A	1 in 253 (1/142–1/452)

$P = 4 \times 10^{-6}$

b

Figure 3 | Expression of *PIK3CA*[H1047R] evokes multipotent stem-like cells. **a**, Number of outgrowths in cleared-fat-pad transplantation of GFP-positive (Lgr5-positive) control and *PIK3CA*[H1047R]-expressing basal subsets (top). Outgrowths of GFP-negative control and GFP-positive *PIK3CA*[H1047R]-expressing basal subsets (bottom). The GFP-negative control basal subset comprises mammary repopulating cells and mostly myoepithelial cells. CI, confidence interval; MRU, mammary repopulating unit; N/A, not applicable. **b**, Representative immuno-stained sections ($n = 3$ mice) and carmine-stained whole mounts of outgrowths from cleared-fat-pad transplantation (100 cells condition). Scale bars, 50 μm (top three rows), 500 μm (bottom row). **a**, **b**, Pooled data from three independent experiments.

show that *PIK3CA*[H1047R] mostly evokes benign tumours (74%) with high intratumour heterogeneity when expressed in Lgr5-positive cells, in contrast to the mostly aggressive mammary tumours (62%) with a distinctive infiltrative densely fibrotic phenotype seen when the cell of origin is K8-positive (Fig. 4c and Extended Data Fig. 9e, f).

Microarray analysis, principle component analysis and hierarchical clustering of tumours revealed a single cluster of Lgr5-CreER[T2]/*PIK3CA*[H1047R] and three clusters of K8-CreER[T2]/*PIK3CA*[H1047R] tumours. We found no correlation between tumour phenotype and clustering, probably owing to intratumour heterogeneity. We compared tumour expression profiles to different human breast cancer subtypes[5]. The majority (7/10) of the K8-CreER[T2]/*PIK3CA*[H1047R] tumours correlated best with the malignant basal-like, HER2-enriched and luminal-B profiles, whereas 3/10 clustered with benign luminal-A and normal-like breast cancers. Lgr5-CreER[T2]/*PIK3CA*[H1047R] tumours resembled benign (2/10) but also malignant subtypes (3/10). However, 5/10 showed no similarity to a single subtype but were equidistant from all, representing the high intratumour heterogeneity in this model (Fig. 4d and Extended Data Fig. 10). The fact that

Figure 4 | The frequency of malignant tumour lesions is dictated by the cell of origin. a, Kaplan–Meier plots depicting tumour onset in Lgr5-CreER^{T2}/$PIK3CA^{H1047R}$ (average 108 days; $n = 15$) and K8-CreER^{T2}/$PIK3CA^{H1047R}$ (average 78 days; $n = 15$) mice compared with $PIK3CA^{WT}$ and control animals after tamoxifen injection (top, $n = 14$; bottom, $n = 9$ and $n = 12$, respectively). *$P < 0.0001$, log-rank test. **b**, Representative images of immunostaining and quantification of double-positive epithelial area in Lgr5-CreER^{T2}/$PIK3CA^{H1047R}$ and K8-CreER^{T2}/$PIK3CA^{H1047R}$ tumours ($n = 10$). Yellow arrowheads indicate double-positive cells. Scale bars, 100 μm (right), 20 μm (left, magnification). Graphs show means ± standard deviation (s.d.); two-sided unpaired Student's t-test; *$P = 0.0002$; NS, not significant. **c**, Pie chart showing phenotypes of tumours evoked in Lgr5-CreER^{T2}/$PIK3CA^{H1047R}$ ($n = 50$) and K8-CreER^{T2}/$PIK3CA^{H1047R}$ ($n = 29$) animals. **d**, Clustered heat map showing correlation coefficients between human breast cancer profiles and ten K8- and Lgr5-CreER^{T2}/$PIK3CA^{H1047R}$ mouse mammary tumours. Ctrl, control; Lum, luminal.
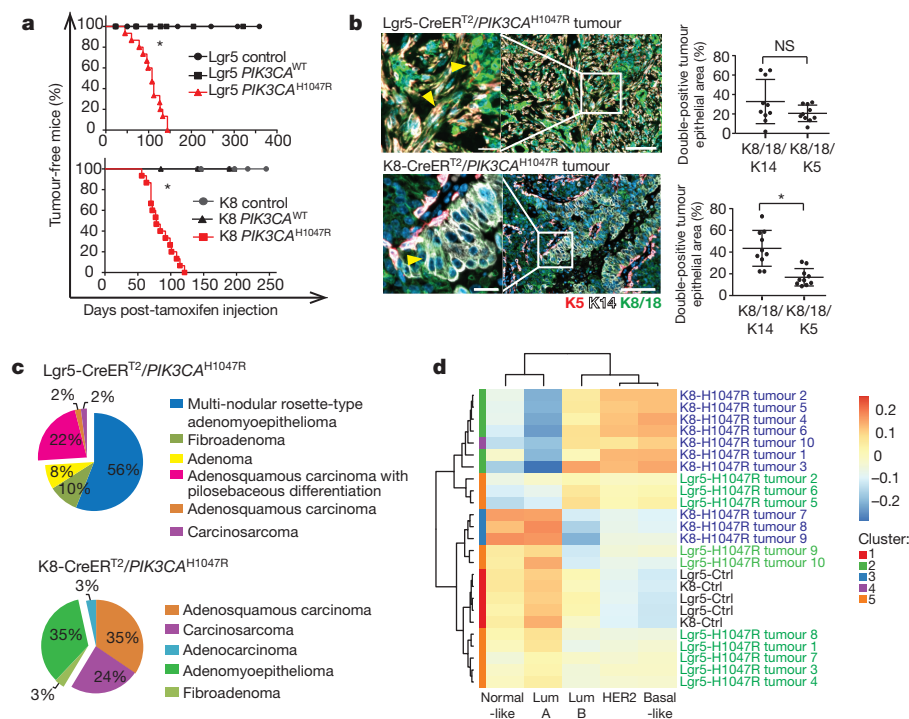
K8-CreER^{T2}/$PIK3CA^{H1047R}$ tumours but not Lgr5-CreER^{T2}/$PIK3CA^{H1047R}$ tumours clustered mostly with malignant breast cancers that have a poor prognosis is consistent with the histopathological results and suggests that, in the presence of the same initiating oncogenic mutation, the cell of origin dictates the frequency of aggressive tumours.

GFP driven by the Lgr5 promoter was found in a rare luminal subset of the control gland (Extended Data Fig. 5b, c), as shown previously[1]. It is unlikely that such rare cells expand due to $PIK3CA^{H1047R}$ and form the luminal GFP-positive population. We observed multi-lineage labelling upon $PIK3CA^{H1047R}$ expression in a basal and a luminal cell-driven model. An alternative possibility that we cannot firmly exclude is that Lgr5- and K8-positive populations contain rare bipotent subsets that are quiescent or not efficiently labelled in physiological conditions and are, therefore, not detected by lineage tracing but may expand upon $PIK3CA^{H1047R}$ expression.

We show that expression of $PIK3CA^{H1047R}$ dedifferentiates lineage-restricted epithelial cells into a multipotent stem-like state from which cells further differentiate, revealing a mechanism by which heterogeneous mixed-lineage tumours may develop. Furthermore, we show that the tumour cell of origin influences the frequency of malignant mammary tumours. The fundamental questions of which mammary cells are susceptible to which combination of oncogenes and how this impinges on tumour progression and aggressiveness warrant further investigation. Understanding these dynamic relationships is paramount for understanding tumour heterogeneity and for identifying prognostic and predictive biomarkers.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

1. Van Keymeulen, A. *et al.* Distinct stem cells contribute to mammary gland development and maintenance. *Nature* **479,** 189–193 (2011).
2. van Amerongen, R., Bowman, A. N. & Nusse, R. Developmental stage and time dictate the fate of Wnt/β-catenin-responsive stem cells in the mammary gland. *Cell Stem Cell* **11,** 387–400 (2012).
3. de Visser, K. E. *et al.* Developmental stage-specific contribution of LGR5+ cells to basal and luminal epithelial lineages in the postnatal mammary gland. *J. Pathol.* **228,** 300–309 (2012).
4. Tao, L., van Bragt, M. P., Laudadio, E. & Li, Z. Lineage tracing of mammary epithelial cells using cell-type-specific Cre-expressing adenoviruses. *Stem Cell Rep.* **2,** 770–779 (2014).
5. The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490,** 61–70 (2012).
6. Hennighausen, L. & Robinson, G. W. Signaling pathways in mammary gland development. *Dev. Cell* **1,** 467–475 (2001).
7. Spike, B. T. *et al.* A mammary stem cell population identified and characterized in late embryogenesis reveals similarities to human breast cancer. *Cell Stem Cell* **10,** 183–197 (2012).
8. Shackleton, M. *et al.* Generation of a functional mammary gland from a single stem cell. *Nature* **439,** 84–88 (2006).
9. Stingl, J. *et al.* Purification and unique properties of mammary epithelial stem cells. *Nature* **439,** 993–997 (2006).
10. Sleeman, K. E. *et al.* Dissociation of estrogen receptor expression and *in vivo* stem cell activity in the mammary gland. *J. Cell Biol.* **176,** 19–26 (2007).
11. Prater, M. D. *et al.* Mammary stem cells have myoepithelial cell properties. *Nature Cell Biol.* **16,** 942–950 (2014).
12. Rios, A. C., Fu, N. Y., Lindeman, G. J. & Visvader, J. E. *In situ* identification of bipotent stem cells in the mammary gland. *Nature* **506,** 322–327 (2014).
13. Wang, D. *et al.* Identification of multipotent mammary stem cells by protein C receptor expression. *Nature* **517,** 81–84 (2015).
14. Samuels, Y. *et al.* High frequency of mutations of the *PIK3CA* gene in human cancers. *Science* **304,** 554 (2004).
15. Koren, S. & Bentires-Alj, M. Mouse models of *PIK3CA* mutations: one mutation initiates heterogeneous mammary tumors. *FEBS J.* **280,** 2758–2765 (2013).
16. Meyer, D. S. *et al.* Luminal expression of *PIK3CA* mutant H1047R in the mammary gland induces heterogeneous tumors. *Cancer Res.* **71,** 4344–4351 (2011).
17. Liu, P. *et al.* Oncogenic PIK3CA-driven mammary tumors frequently recur via PI3K pathway-dependent and PI3K pathway-independent mechanisms. *Nature Med.* **17,** 1116–1120 (2011).
18. Meyer, D. S. *et al.* Expression of *PIK3CA* mutant E545K in the mammary gland induces heterogeneous tumors but is less potent than mutant H1047R. *Oncogenesis* **2,** e74 (2013).
19. Plaks, V. *et al.* Lgr5-expressing cells are sufficient and necessary for postnatal mammary gland organogenesis. *Cell Rep.* **3,** 70–78 (2013).
20. Sleeman, K. E., Kendrick, H., Ashworth, A., Isacke, C. M. & Smalley, M. J. CD24 staining of mouse mammary gland cells defines luminal epithelial, myoepithelial/basal and non-epithelial cells. *Breast Cancer Res.* **8,** R7 (2006).
21. Meier-Abt, F. *et al.* Parity induces differentiation and reduces Wnt/Notch signaling ratio and proliferation potential of basal stem/progenitor cells isolated from mouse mammary epithelium. *Breast Cancer Res.* **15,** R36 (2013).
22. Lim, E. *et al.* Transcriptome analyses of mouse and human mammary cell subpopulations reveal multiple conserved genes and pathways. *Breast Cancer Res.* **12,** R21 (2010).

23. Lim, E. *et al.* Aberrant luminal progenitors as the candidate target population for basal tumor development in *BRCA1* mutation carriers. *Nature Med.* **15**, 907–913 (2009).
24. Molyneux, G. *et al.* BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. *Cell Stem Cell* **7**, 403–417 (2010).

Author Contributions S.K. and M.B.-A. designed experiments, analysed the data and wrote the manuscript. L.R. contributed greatly to experimental design and data analysis. S.K. performed most of the experiments. L.R. and D.D.S. performed mammosphere cultures. L.R., D.D.S., J.P.C. and S.K. performed limiting dilution transplantations. M.B.S. and T.R. performed microarray data analysis. L.R. and J.P.C. quantified tumour immunohistochemistry. J.P.C. and A.B. isolated tumour RNA. A.B. performed immunoblotting. D.D.S. and T.E. provided technical assistance for several experiments. H.K. provided technical assistance for FACS experiments. O.A. and R.D.C. analysed histological tumour samples. All the authors discussed the data and participated in the preparation of the manuscript.

METHODS

Mice. Lgr5-CreER^{T2} mice (C57BL/6) were generated and provided by B. Kinzel and J. Tchorz[25]. Mice were backcrossed to the FVB background and used in this study. K8-CreER^{T2} mice (CD-1) were provided by C. Blanpain[1]. The generation of PIK3CA^{H1047R} and PIK3CA^{WT} (pure FVB) was described previously[16,18]. Tomato-reporter animals (C57BL/6) were provided by B. Roska. For lineage-tracing studies, mice of a mixed background (FVB/CD-1/C57BL/6 and FVB/C57BL/6) were used. For studies without a Tomato reporter, mice with a pure FVB background (Lgr5-CreER^{T2} model) or mixed FVB/CD-1 background (K8-CreER^{T2} model) were used. Mouse colonies were maintained in the animal facility of the Friedrich Miescher Institute for Biomedical Research and experiments were carried out in accordance with Swiss national guidelines on animal welfare and the regulations of the canton of Basel-Stadt, Switzerland.

Targeting Tomato and/or GFP expression. Adult 7- to 8-week-old female mice were induced by intraperitoneal injection of tamoxifen (Sigma; 2 mg per 25 g of body weight) for three consecutive days (diluted in sunflower seed oil, Sigma) to activate cell-specific expression of the Cre recombinase and thus expression of Tomato and/or the PIK3CA$^{H1047R\ or\ WT}$ transgene. Tomato and/or GFP were expressed in recombined cells derived from Lgr5-positive or K8-positive cells, respectively. For the Lgr5 model, after Cre induction, GFP expression was derived from both Lgr5 promoter activity and PIK3CA^{H1047R} expression.

Histology and immunostaining. Tumours and dissected mammary glands were spread on a glass slide and fixed in 20% formalin for 24 h at 4 °C. Samples were then processed, embedded in paraffin and sectioned (3 μm). Immunofluorescence studies were performed using a Ventana DiscoveryUltra instrument (Roche Diagnostics) following the RUO Discovery Universal method. Briefly, slides were pretreated with CC1 for 40 min and then incubated with primary antibodies for 1 h at 37 °C. After washing, secondary antibodies were incubated for 32 min at 37 °C. Slides were then washed with reaction buffer (three times), PBS (two times) and then incubated with DAPI (1 μg ml^{-1}) for 5 min. Finally, slides were rinsed with PBS (three times) and mounted in Mount Fluor (ProTaqs). The following antibodies were used: anti-keratin 8/18 (guinea-pig, 1:200, Fitzgerald, 20R-CP004), anti-RFP (rabbit, 1:500, Rockland, 600-401379), anti-keratin 14 (chicken, 1:500, Covance, Sig3476), anti-p63 (mouse, 1:500, Thermo Scientific, ma121871), anti-keratin 5 (rabbit, 1:500, Abcam, ab52635), anti-chicken Alexa Fluor 488, anti-chicken Alexa Fluor 568, anti-guinea pig Alexa Fluor 488, anti-mouse Alexa Fluor 488, anti-rabbit Alexa Fluor 647. All Alexa secondary antibodies were obtained from Molecular Probes (Invitrogen). Immunohistochemistry experiments were performed for keratin 14, keratin 8/18, GFP, and p63 using a Ventana DiscoveryXT instrument (Roche Diagnostics) following the Research IHC DAB Map XT procedure. Slides were treated with a mild CC1 and incubated with the primary antibodies for 1 h at 37 °C. After brief washes, biotinylated donkey-anti-rabbit and biotinylated anti-guinea-pig, respectively, were applied for 32 min at 37 °C. For mouse-anti p63 detection, a monoclonal rabbit-anti-mouse antibody was applied for 32 min at 37 °C, followed by incubation with a polymer anti-rabbit conjugated with horseradish peroxidase (HRP) (ImmPRESS anti-rabbit peroxidase, Vector Laboratories) for 32 min at 37 °C. Finally, sections were counterstained with haematoxylin II and bluing reagent (4 min). Staining against ERα, SMA and keratin 5 was performed by deparaffinization followed by antigen retrieval with citrate buffer and quenching with PBS plus 3% H$_2$O$_2$. Slides were then blocked with PBS plus 2.5% normal goat serum and primary antibodies incubated overnight at 4 °C in PBS plus 1% BSA plus 0.5% Tween-20. After brief washes, secondary antibodies were then incubated in PBS plus 1% BSA for 30 min at room temperature. Signals were enhanced using the Vectastain ABC system and visualized with 3,3′-diaminobenzidine (DAB; Sigma). Haematoxylin was used as counterstain. Anti-PR staining was performed without antigen retrieval. The following antibodies were used: anti-keratin 8/18 (guinea-pig, 1:500, Fitzgerald, 20R-CP004), anti-keratin 14 (rabbit, 1:500, Thermo Scientific, Rb9020), anti-p63 (mouse, 1:1,000, Thermo Scientific, ma121871), anti-keratin 5 (rabbit, 1:1,000, Abcam, ab52635), anti-SMA (rabbit, 1:500, Thermo Scientific, Rb9010), anti-ERα (rabbit, 1:1,000, Santa Cruz, sc-542), anti-PR (rabbit, 1:200, Thermo Scientific, Rm9102), anti-GFP (rabbit, 1:50, Invitrogen, A11122). Secondary antibodies were biotinylated anti-rabbit IgG (1:200, Jackson Immunoresearch), biotinylated anti-guinea pig IgG (1:200; Vector Laboratories) and biotinylated anti-mouse IgG (1:200; Abcam). Haematoxylin and eosin staining was performed using standard protocols.

Microscopy image acquisition. For immunofluorescence, images of stained sections were captured using a Zeiss Z1 wide-field fluorescent microscope, ×5/0.13, ×10/0.45 or ×20/0.8 (Plan-APOCHROMAT) objectives, an AxioCamMRc camera (1,024 × 1,024, pixel size 6.45 μm) and an Axiocam506 camera (2,752 × 2,208, pixel size 4.54 μm). Whole-mount fluorescent mammary gland and tumours sections were imaged on the Zeiss Axio Scan.Z1 slide scanner (ORCA-Flash4.0 camera, 2,048 × 2,048, pixel size 6.5 μm). Representative images

were cropped and processed using the ZenBlue software. For immunohistochemistry, stained sections were examined using a Nikon E600 Eclipse brightfield microscope (×20/0.5 and ×40/0.75 objectives) and images captured with a Nikon DXM1200 camera (2,592 × 1,944, pixel size 6.7 μm) using the IMS acquisition software. All images were scaled appropriately.

Quantification of immunohistochemistry. Five representative images of 8–15 tumours of each genotype were captured with a Nikon E600 Eclipse brightfield microscope (×20/0.5 objective) and positively stained tumour areas, luminal cells, and total epithelial cells quantified with Image J (Fiji). Evaluation of tissue sections was performed blindly by two independent investigators.

Quantification of double-positive tumour epithelial area. Ten tumours from each model were stained with anti-keratin 14, anti-keratin 5, anti-keratin 8/18, anti-rabbit Alexa Fluor 647, anti-chicken Alexa Fluor 568, and anti-guinea-pig Alexa Fluor 488. Whole-mount tumour sections were scanned using the Zeiss Axio Scan Z1 slide scanner. For computational reasons, the whole images of tumours were first tiled and the channels split using the ZenBlue software and Matlab. Images were then processed using Ilastik[26], an interactive supervised machine learning toolkit, and the subsequent prediction maps were treated with batch functions written with the Matlab programing language. For the statistical analysis shown in Fig. 4b, a total of 23 regions of tumours were subsequently tiled into 14,872 squared images of 1,024 × 1,024 pixels. In a second step, data from a set of nine selected tiles were generated by annotating a few regions representative of the predefined classes. For each of the three channels the training data were processed using the four phenotype classes: '1.background', '2.fluorescent marker (Alexa 488, 568 or 647)', '3.blood cells', and '4.Stroma'. After annotation of the different pixels of the different classes by brush stroke, features of the labelled pixels and their local neighbourhood are used to train a Random Forest classifier. In a third step, the inferred classifier was used to predict all the tiles in a batch process. The last step implemented in Matlab determined the masks for all the channels and tiles. To perform the analysis on relevant tiles, that is, tiles sufficiently covered by tissue, we performed a k-means clustering on a set of statistical features extracted from the tiles. This enabled us to suppress tiles that were merely in the background or those that contained non-epithelial structures. For each tile, the fluorescent marker mask corresponded to pixels having a probability above 50% in the class 'fluorescent marker (Alexa)'. In addition, a non-tissue mask was calculated by the mean of morphological filters on the union of background, blood cells and stroma classes. The mask for tissue represented the complement image of the non-tissue mask. We calculated on each tile the ratio between the areas of pixels in the mask of the fluorescent marker and those in the tissue mask. Finally, the distribution of double-labelled fluorescent marker within the whole tumour epithelial area was represented in Fig. 4b.

Preparation of mammary single-cell suspensions and labelling. Mammary glands were dissected and intra-mammary lymph nodes removed. To obtain mammary organoids, mammary glands were processed as described previously[20,21]. To obtain single mammary epithelial cells, organoids were washed in serum-free Leibowitz L15-medium (Gibco) and digested with Hyclone HyQTase (Thermo Scientific). Single cells were washed and filtered through a 40-μm cell strainer (BD Falcon) and counted; 10^6 cells per ml were stained with the following antibodies: PE-Cy7-CD45 (Biolegend; clone 30-F11), APC-Sca1 (Biolegend; clone E13-161.7), PerCP-Cy5.5-CD24 (Biolegend; clone M1/69), PE-CD49f (BD-Pharmingen), Alexa700-CD24 (Novus Biologicals; clone M1/69) and DAPI (2 μg ml^{-1}, Invitrogen).

Flow cytometry. FACS was carried out with a BD FACSAria III (Becton Dickinson) using a 100-μm nozzle. Cells were gated based on their forward- and sideward-scatter. Pulse-width was used to exclude doublets. DAPI-negative/CD45-negative cells were gated for Tomato or GFP. CD24, Sca1 and CD49f subsets were then gated on GFP-positive epithelium. Tomato-positive cells were gated using only CD24, CD45 and Sca1 antigens. The same numbers of living cells were recorded in each condition. FACS data were analysed using FlowJo (Tree Star). Total cell numbers were determined by enumerating the total epithelial content after single-cell isolation and calculating back based on the percentages obtained from antibody staining, sorting and FlowJo analysis. Cell numbers were subsequently normalized to one animal (3–9 independent FACS experiments of 1–5 pooled animals per time point were quantified; ± s.e.m.; $P < 0.05$, Student's t-test).

In vitro colony formation assay and quantification. Freshly sorted cells of each subpopulation (500 cells) were plated as previously described[21]. Seven days later, the colonies were fixed with acetone/methanol (1:1), washed, blocked with 2.5% normal goat serum and stained with anti-keratin 8/18 (guinea pig, 1:500, Fitzgerald, 20R-CP004), anti-keratin 14 (rabbit, 1:500, Thermo Scientific, Rb9020), DAPI (2 μg ml^{-1}, Invitrogen), anti-guinea-pig Alexa Fluor 488 and anti-rabbit Alexa Fluor 647 (1:1,000, Invitrogen). Colonies were imaged using the Zeiss Z1 wide-field fluorescent microscope (×5/0.13 DIC). Colonies were

defined as a cluster of more than five cells. The number of colonies per well was determined manually. Colonies containing more than 20% of keratin 8/18/keratin 14 double-positive cells were defined as 'double-positive'.

In vitro mammosphere culture. Mammosphere cultures were performed as described previously[27]. Freshly sorted subsets (luminal CD24[Hi]Sca1[−] and CD24[Hi]Sca1[+]) from adult FVB and uninduced *PIK3CA*[H1047R] females were plated at 20,000 cells per ml in 6-well ultra-low attachment plates (Falcon) in DMEM/F12 medium (Gibco) supplemented with 5 µg ml[−1] insulin, 0.5 µg ml[−1] hydrocortisone, 2% B27 (Invitrogen), 20 ng ml[−1] EGF and bFGF (BD Biosciences) and cholera toxin (Sigma), and cultured at 37 °C in 5% CO$_2$. After 5 days, mammospheres were collected and dissociated into single cells using HYQtase (Gibco) and counted. Luminal subsets from FVB control and *PIK3CA*[H1047R] animals were subsequently treated with TAT-Cre (Millipore) (0.5 µM) at a density of 20,000 cells per ml in 4 ml overnight at 37 °C in 5% CO$_2$. The next morning, the medium was replaced and cells were cultured for 72 h at 37 °C in 5% CO$_2$ to allow for maximal recombination and expression of *PIK3CA*[H1047R]. After 72 h, *PIK3CA*[H1047R]-GFP[+/−] cells were sorted and plated into 24-well ultra-low attachment plates in medium (described earlier) supplemented with 2% Matrigel (growth factor reduced; BD, 356230) at a density of 1,000 cells per well. Control and *PIK3CA*[H1047R] spheres from each subset were enumerated every 7 days, at which point spheres were dissociated and re-plated at a density of 1,000 cells per well.

Mammary fat pad transplantation. For limiting dilution transplantation, freshly sorted cells from Lgr5-CreER[T2] control and Lgr5-CreER[T2]/*PIK3CA*[H1047R] animals (pure FVB background) and from K8-CreER[T2] control and K8-CreER[T2]/*PIK3CA*[H1047R] F1-hybrid (FVB/CD-1) littermates were used. Sorted cells were resuspended in limiting dilution numbers in PBS plus 2% FCS with 25% Matrigel (growth factor reduced; BD, 356230) and injected in 20-µl volumes into inguinal glands of 3-week-old FVB females that had been cleared of endogenous mammary epithelium[28]. Cells from control and from *PIK3CA*[H1047R] animals were injected in the same animal on opposite sides. After 8 weeks, glands of the recipients were removed for evaluation. Glands were spread on a glass slide and fixed in Carnoy's fixative overnight. Whole-mount staining with carmine alum was performed as previously described[21] and scanned with an Epson 1600 Pro scanner. An outgrowth was defined as an epithelial structure composed of ducts arising from a central point with lobules and/or terminal end buds[8]. Frequencies of mammary-repopulating units between different cell populations were calculated and statistically compared using the Extreme Limiting Dilution Analysis (ELDA)[29] online tool (http://bioinf.wehi.edu.au/software/elda/).

Immunoblotting. Lysates from mammary glands were prepared by lysing cryo-homogenized mammary gland powder in RIPA buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) supplemented with 1× protease inhibitor cocktail (Complete Mini, Roche), 0.2 mM sodium orthovanadate, 20 mM sodium fluoride and 1 mM phenylmethylsulfonyl fluoride. Lysates (30–80 µg) were subjected to SDS–PAGE, transferred to PVDF membranes (Immobilon-P, Millipore) and blocked for 1 h at room temperature with 5% milk or BSA in PBS/0.05% Tween 20. Membranes were then incubated overnight with primary antibodies (1:200–1:3,000) and exposed to secondary HRP-coupled anti-mouse or anti-rabbit antibodies at 1:5,000–10,000 for 1 h at room temperature. Results are representative of at least three different experiments. The following antibodies were used: anti-AKT pan (Cell Signaling), anti-pAKT (Ser473, Cell Signaling), anti-p110α (Cell Signaling), anti-ERK2 (Santa Cruz) and anti-keratin pan (SantaCruz). Blot densities were quantified using ImageJ and normalized to pan-keratin for epithelial content.

RNA isolation. For pre-neoplastic gene expression profiling, mammary epithelial subsets were FACS sorted as described earlier into extraction buffer and total RNA from 250 or 2,000 sorted cells was isolated using an Arcturus PicoPure RNA Isolation Kit (Life Technologies). Subsets of pooled mammary glands of 2–3 oestrus-synchronized animals (confirmed by vaginal smear) per genotype in three independent sortings were collected for microarray analysis. RNA from mammary subsets of four animals per genotype was collected for qRT–PCR. For tumour gene expression profiling and expression of Lgr5 in the mammary gland, total RNA from 50 mg cryo-homogenized tumour tissue or nipple and distal area of mammary glands were extracted using the TRIzol method (Life Technologies) according to the manufacturer's instructions. Genomic DNA was removed by DNase I digestion (Qiagen) and RNA purified using an RNeasy Plus Mini Kit (Qiagen). RNA concentration was measured with a Nanodrop 1000 machine and RNA quality assessed using an Agilent 2100 bioanalyzer and RNA Pico Chips or RNA Nano Chips.

Microarray. Total RNA from 250 sorted cells (pre-neoplastic Lgr5), 2,000 sorted cells (pre-neoplastic K8) or 100 ng from tumour tissue was used as the input for synthesis of amplified cDNA with the NuGen Ovation Pico WTA System (NuGen). The resulting double-stranded cDNA was fragmented and labelled

using the Affymetrix GeneChip WT Terminal Labelling kit (Affymetrix). Affymetrix Gene Chip Mouse gene 1.0 ST microarrays were hybridized according to the GeneChip Whole Transcript (WT) Sense Target Labelling Assay Manual (Affymetrix) with a hybridization time of 16 h. Scanning was performed with Affymetrix GCC Scan Control Software v. 3.0.0.1214 on a GeneChip Scanner 3000 7 G with autoloader.

Normalization and analysis of mammary epithelial subset microarray data. The arrays for Gene Expression Omnibus accession numbers GSE40875 (ref. 21), GSE59870 (Fig. 2a) and GSE65411 (Fig. 2b) data sets were RMA normalized using the bioconductor package affy (R3.0.1/Bioconductor 2.13). Heat maps with unscaled normalized expression values for selected genes were plotted with the heatmap.2 function of the gplots package. Differential gene expression for the comparison of mature luminal (LM) or luminal progenitor (LP) versus myoepithelial (MYO) cells was calculated with limma. The function topTable was used to select the top 300 upregulated genes as luminal mature versus myoepithelial UP (Luminal_Mature UP, MEIER-ABT) and luminal progenitors versus myoepithelial UP (Luminal_progenitors UP, MEIER-ABT), and the top 300 upregulated genes as myoepithelial versus luminal mature UP (Myoepithelial UP, MEIER-ABT). Signatures LIM_MAMMARY_LUMINAL_MATURE_DN, LIM_MAMMARY_LUMINAL_MATURE_UP, LIM_MAMMARY_LUMINAL_PROGENITOR_DN, LIM_MAMMARY_LUMINAL_PROGENITOR_UP, LIM_MAMMARY_STEM_CELL_DN and LIM_MAMMARY_STEM_CELL_UP were downloaded from http://www.broadinstitute.org/gsea/msigdb/search.jsp in gmt format, combined with the signatures described earlier, and used in the PGSEA and smcPlot functions of the Bioconductor PGSEA package.

qRT–PCR. RNA was converted into cDNA using SuperScript III Reverse Transcriptase (Invitrogen). Quantitative real-time PCR was performed on unamplified cDNA normalized to the number of sorted cells for each subset (6,000–10,000 sorted cells). Taqman probes (Life Technologies) and Taqman Universal PCR Mastermix (Applied Biosystems) were applied. The following Taqman probe identifiers were used: Krt14 Mm00516879_m1, Krt5 Mm01305291_g1, Lgr5 Mm00438890_m1, Vim Mm01333430_m1, Krt8 Mm00835759_m1, Krt18 Mm01601702_g1, Gata3 Mm00484683_m1, Elf5 Mm00468732_m1, Csn2 Mm04207885_m1 and Axin2 Mm00443610_m1. Cycling was performed with StepOne Plus Real-time PCR Systems (Applied Biosystems). The results are representative of three qRT–PCR experiments of pooled mammary subsets from four animals of each genotype. Results of *Csn2* (CD24[Hi]Sca1[−/+]) and *Krt14* (CD24[Hi]Sca1[+]) from K8-CreER[T2]/*PIK3CA*[H1047R] cells and of *Krt5* and *Lgr5* (CD24[Hi]Sca1[−]) from Lgr5-CreER[T2]/*PIK3CA*[H1047R] cells are representative of two experiments. Statistical data analysis was performed using ΔCT values.

Normalization and analysis of tumour mouse microarray data. Mouse Affymetrix microarrays (Gene Expression Omnibus accession number GSE59872) were background corrected, quantile normalized, and log$_2$ transcript cluster expression values were calculated using the rma() function of the Bioconductor package "oligo"[30]. Transcript cluster identifiers were mapped to Entrez Gene identifiers using the Bioconductor package "mogene10sttranscriptcluster.db", and transcript clusters associated with none or multiple genes were removed. Where multiple transcript clusters were associated with a single gene that with the maximal variance across samples was selected; this resulted in a total number of 20,365 transcript clusters with unique gene assignments. Principal component analysis was performed on log$_2$ expression values from which the mean over samples had been subtracted for each gene. Hierarchical clustering of samples was performed using $(1 − r)$ as distance metric, with r being the Pearson's correlation coefficient for log$_2$ expression values between each pair of samples. The clustering dendrogram was visualized using the "dendextend" R package.

Analysis of human TCGA breast cancer data and comparison to mouse. Human breast cancer expression data and corresponding clinical data[5] were obtained from https://tcga-data.nci.nih.gov/docs/publications/brca_2012/, corresponding to the 11 November, 2011 data freeze that contains 522 tumour samples with clinical annotation. Human gene symbols were mapped to Entrez Gene identifiers and genes selected that are one-to-one homologues between human and mouse according to Homologene (build 68, downloaded from ftp://ftp.ncbi.nlm.nih.gov/pub/HomoloGene/build68/31) and that were measured on both human and mouse experimental platforms; this resulted in 13,969 genes. To reduce technical differences between human and mouse samples, all samples were scaled to a standard deviation of 1, and the within-species mean was subtracted from each gene. Heat map and clustering of the combined human and mouse expression data were performed using the function aheatmap() from the R package "NMF"[32] on the top 1,000 genes ranked by variance across human samples. Mouse samples were compared with human PAM50 tumour subtypes (Normal-like, Luminal A, Luminal B, HER2-enriched and Basal-like) by calculating Pearson's correlation coefficients for each mouse sample against the averages of all human samples within each subtype.

Survival analysis. Tumour incidence was determined by palpation. Kaplan–Meier plots were generated using the survival calculation tool from Graphpad Prism and significance was calculated using the log-rank test.

Statistical data analysis. The number of mice was calculated by performing power analysis using data from small pilot experiments. Values represent the means $\pm$ s.e.m. or $\pm$s.d. Depending on the type of experiments, data were tested using unpaired Student's t-test or log-rank test. $*P < 0.05$ was considered statistically significant. The experiments were not randomized.

25. Kinzel, B. *et al.* Functional roles of Lgr4 and Lgr5 in embryonic gut, kidney and skin development in mice. *Dev. Biol.* **390,** 181–190 (2014).
26. Sommer, C., Strähle, C., Köthe, U. & Hamprecht, F. A. Ilastik: Interactive Learning and Segmentation Toolkit. *Proc. Eighth IEEE Int. Symp. Biomed. Imag.* 230–233 (2011).
27. Cicalese, A. *et al.* The tumor suppressor p53 regulates polarity of self-renewing divisions in mammary stem cells. *Cell* **138,** 1083–1095 (2009).
28. Deome, K. B., Faulkin, L. J. Jr, Bern, H. A. & Blair, P. B. Development of mammary tumors from hyperplastic alveolar nodules transplanted into gland-free mammary fat pads of female C3H mice. *Cancer Res.* **19,** 515–520 (1959).
29. Hu, Y. & Smyth, G. K. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J. Immunol. Methods* **347,** 70–78 (2009).
30. Carvalho, B. S. & Irizarry, R. A. A framework for oligonucleotide microarray preprocessing. *Bioinformatics* **26,** 2363–2367 (2010).
31. NCBI Resource Coordinators. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **42,** D7–D17 (2014).
32. Gaujoux, R. & Seoighe, C. A flexible R package for nonnegative matrix factorization. *BMC Bioinformatics* **11,** 367 (2010).
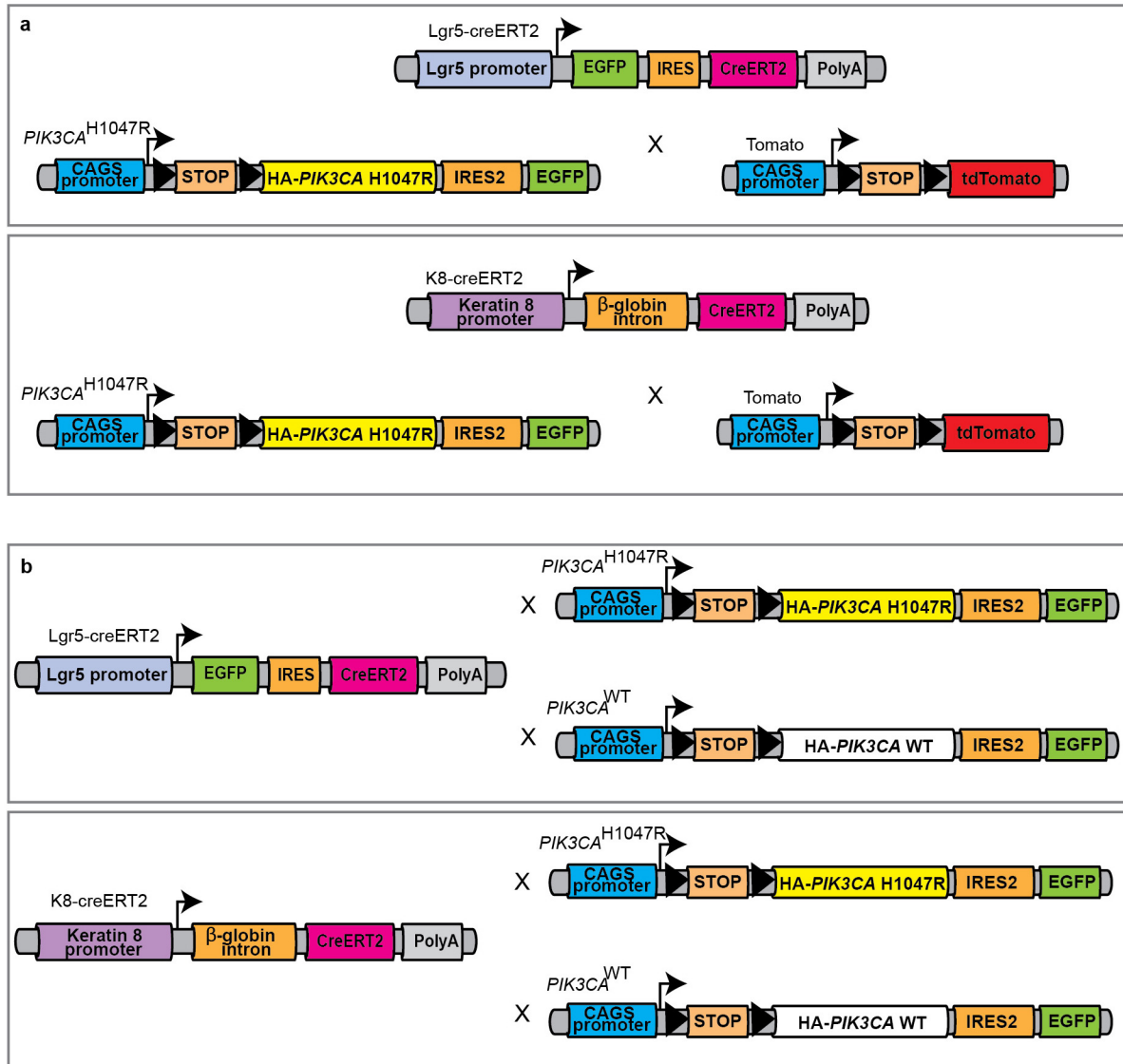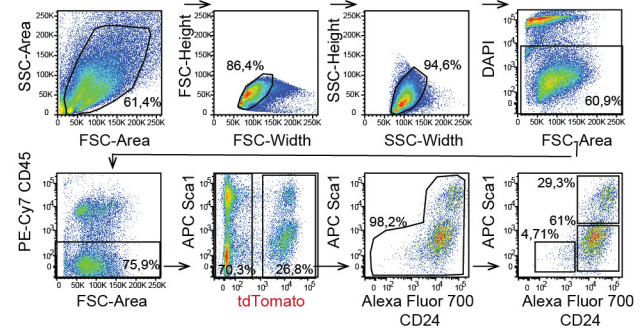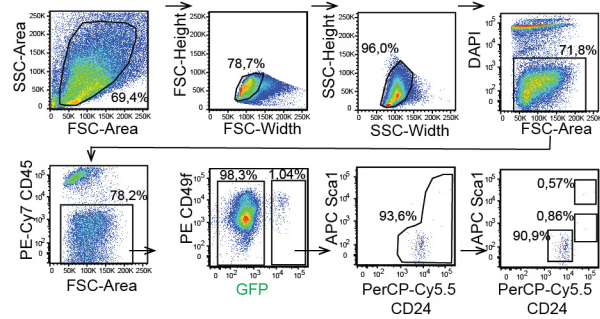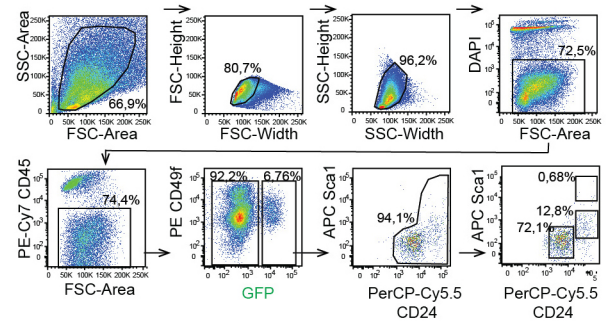
Extended Data Figure 1 | Scheme depicting mouse lines generated for lineage-tracing studies. a, Lgr5-CreER[T2] (ref. 25) or K8-CreER[T2] (ref. 1) animals were crossed to transgenic lox-STOP-lox *PIK3CA*[H1047R 16] and/or Tomato-reporter mice, generating Lgr5-CreER[T2]/Tomato, K8-CreER[T2]/Tomato, Lgr5-CreER[T2]/*PIK3CA*[H1047R]/Tomato and K8-CreER[T2]/*PIK3CA*[H1047R]/Tomato animals for lineage-tracing studies. Lgr5-CreER[T2]/Tomato and

K8-CreER[T2]/Tomato animals were used as controls. **b**, Lgr5-CreER[T2] (ref. 25) and K8-CreER[T2] (ref. 1) animals were crossed to lox-STOP-lox *PIK3CA*[H1047R] (ref. 16) or *PIK3CA*[WT] (ref. 18) animals. Lgr5-CreER[T2] and K8-CreER[T2] animals were used as controls. Tamoxifen injection induces *PIK3CA*[H1047R], *PIK3CA*[WT] and/or Tomato expression.
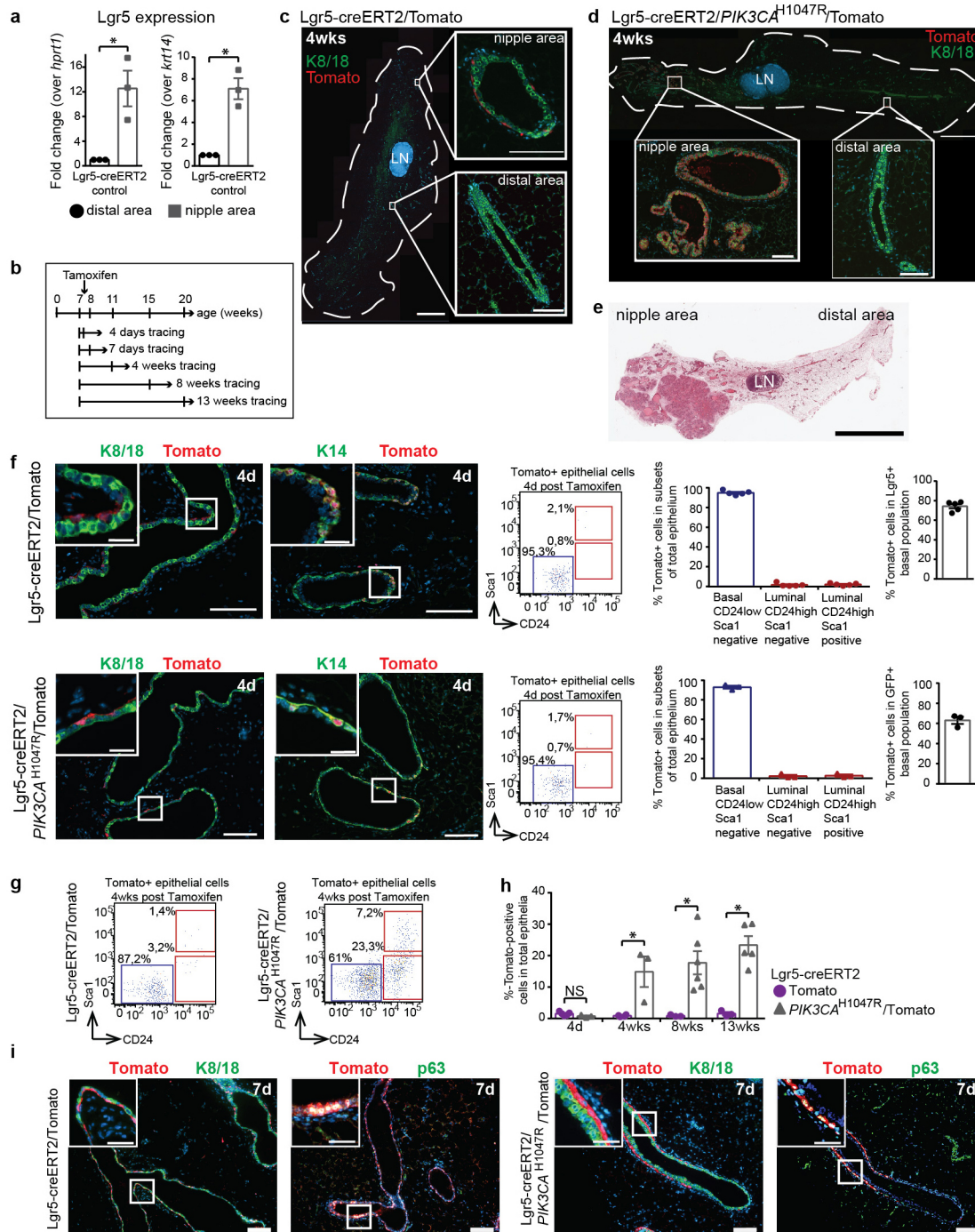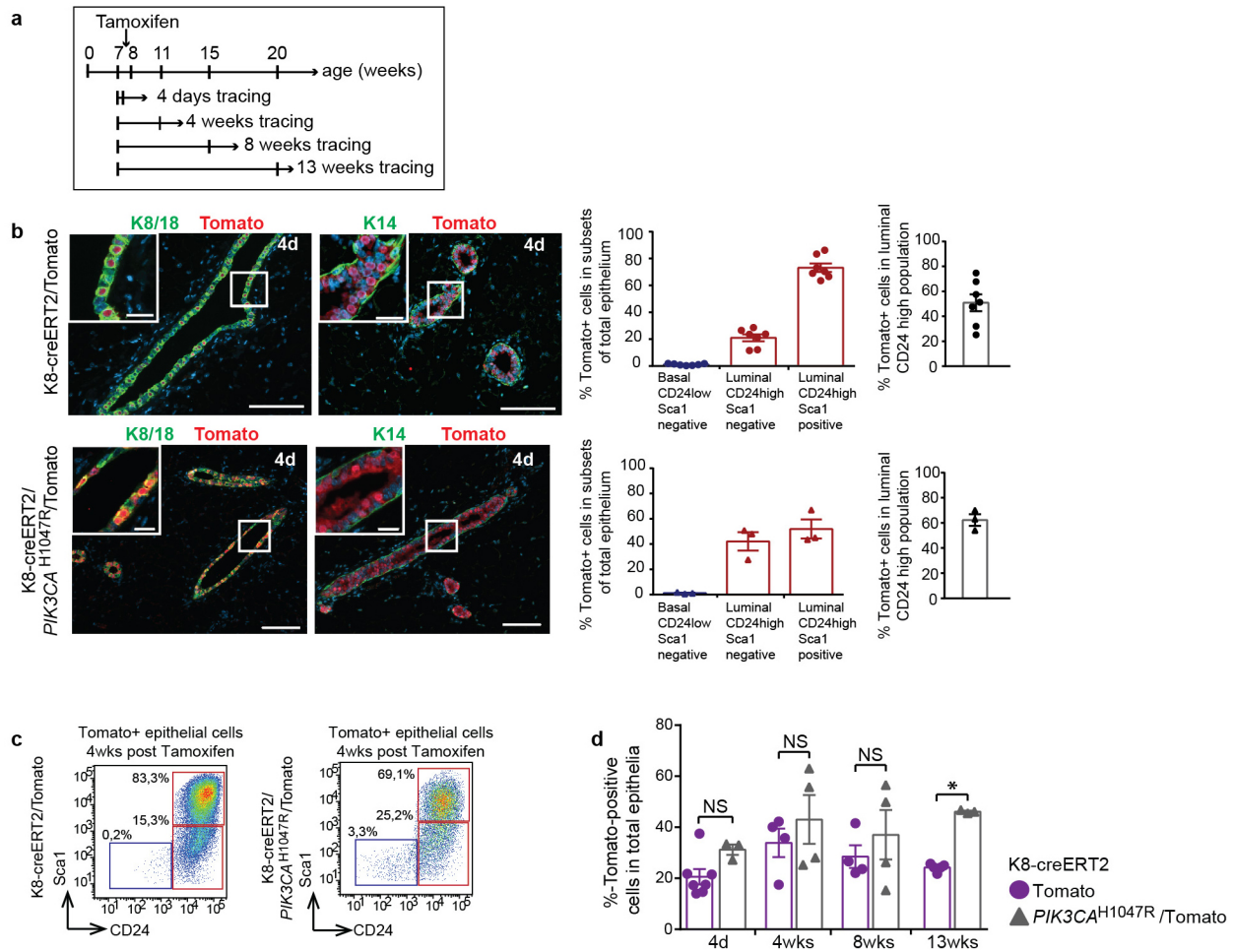
Extended Data Figure 2 | Lgr5-CreER^{T2}/Tomato and *PIK3CA*^{H1047R}/ Tomato labelling in the mammary nipple area. a, Lgr5 expression in the nipple and distal area of Lgr5-CreER^{T2} glands (*n* = 3 mice). **b,** Tracing scheme. **c, d,** Representative images of mammary glands after 4 weeks tracing (*n* = 3 mice for each genotype). Scale bars, 2 mm, 100 μm (magnifications). **e,** Representative haematoxylin and eosin staining of an Lgr5-CreER^{T2}/*PIK3CA*^{H1047R} mammary gland with a tumour. Scale bar, 500 μm. LN, lymph node. **f,** Representative images, FACS plots and quantification of 4 days tracing (24 h after the last tamoxifen injection) (top: immunofluorescence: *n* = 3 mice; FACS: *n* = 5 technical replicates (each 1–2 pooled mice); bottom:

immunofluorescence: *n* = 3 mice; FACS: *n* = 3 technical replicates (each 1 mouse)). Scale bars, 100 μm; 20 μm (magnifications). **g,** Representative FACS plots of 4-week tracing. **h,** Percentage of total Tomato-positive cells in the tracing experiments (Lgr5-CreER^{T2}/Tomato: 4 days *n* = 5, 4 weeks *n* = 4, 8 and 13 weeks *n* = 3 technical replicates (each 1–2 pooled mice); Lgr5-CreER^{T2}/ *PIK3CA*^{H1047R}/Tomato: 4 days *n* = 3, 4 weeks *n* = 3, 8 weeks *n* = 6 and 13 weeks *n* = 5 technical replicates (each 1–2 pooled mice)). **i,** Representative images of 7 days tracing (left *n* = 4 mice; right *n* = 2 mice). Scale bars, 100 μm, 50 μm (magnifications). Bar graphs show means ± s.e.m.; two-sided unpaired Student's *t*-test; *P < 0.05; NS, not significant.
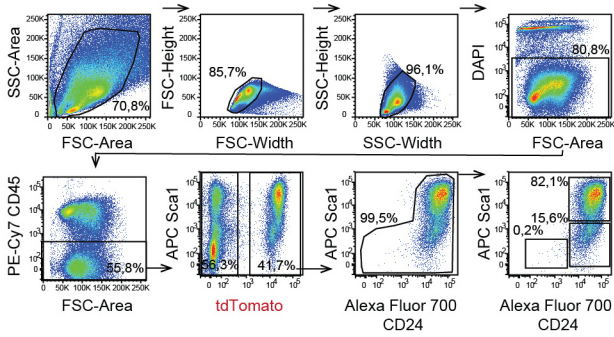
Extended Data Figure 3 | Gating scheme for FACS experiments.
a–d, Representative FACS plots of K8-CreER^{T2}/Tomato (**a**), K8-CreER^{T2}/
$PIK3CA^{H1047R}$/Tomato (**b**), Lgr5-CreER^{T2}/$PIK3CA^{WT}$ (**c**) and Lgr5-CreER^{T2}/
$PIK3CA^{H1047R}$ (**d**) animals 4 weeks after tamoxifen injection. The gating

strategy shown illustrates the elimination of doublets, dead cells (DAPI^{Hi}), and white blood cells (CD45$^+$) and the sorting of Tomato- or GFP-positive mammary epithelial subsets (basal CD24^{Lo}Sca1$^-$, luminal CD24^{Hi}Sca1$^-$; luminal CD24^{Hi}Sca1$^+$).

Extended Data Figure 4 | K8-CreERT2/Tomato and *PIK3CA*[H1047R]/Tomato labelling in the mammary gland. a, Scheme depicting timeline of tracing experiments. **b,** Representative images and FACS quantifications of K8-CreER[T2]/Tomato and K8-CreER[T2]/*PIK3CA*[H1047R]/Tomato mammary glands 4 days after tamoxifen (24 h after the last tamoxifen injection) (top: immunofluorescence: $n = 5$ mice; FACS: $n = 7$ technical replicates (each 1–2 pooled mice); bottom: immunofluorescence $n = 3$ mice; FACS: $n = 3$ technical

replicates (each 1 mouse)). Scale bars, 100 μm, 20 μm (magnifications). **c,** Representative FACS plots of 4-week Tomato tracing. **d,** Percentage of total Tomato-positive cells in mammary glands (K8-CreER[T2]/Tomato: 4 days $n = 7$, 4 and 8 weeks $n = 4$, 13 weeks $n = 5$ technical replicates (each 1–3 pooled mice); K8-CreERT2/*PIK3CA*[H1047R]/Tomato: 4 days $n = 3$, 4 and 8 weeks $n = 4$ and 13 weeks $n = 3$ technical replicates (each 1–2 pooled mice)). Bar graphs show means ± s.e.m. *$P < 0.05$; two-sided unpaired Student's *t*-test.
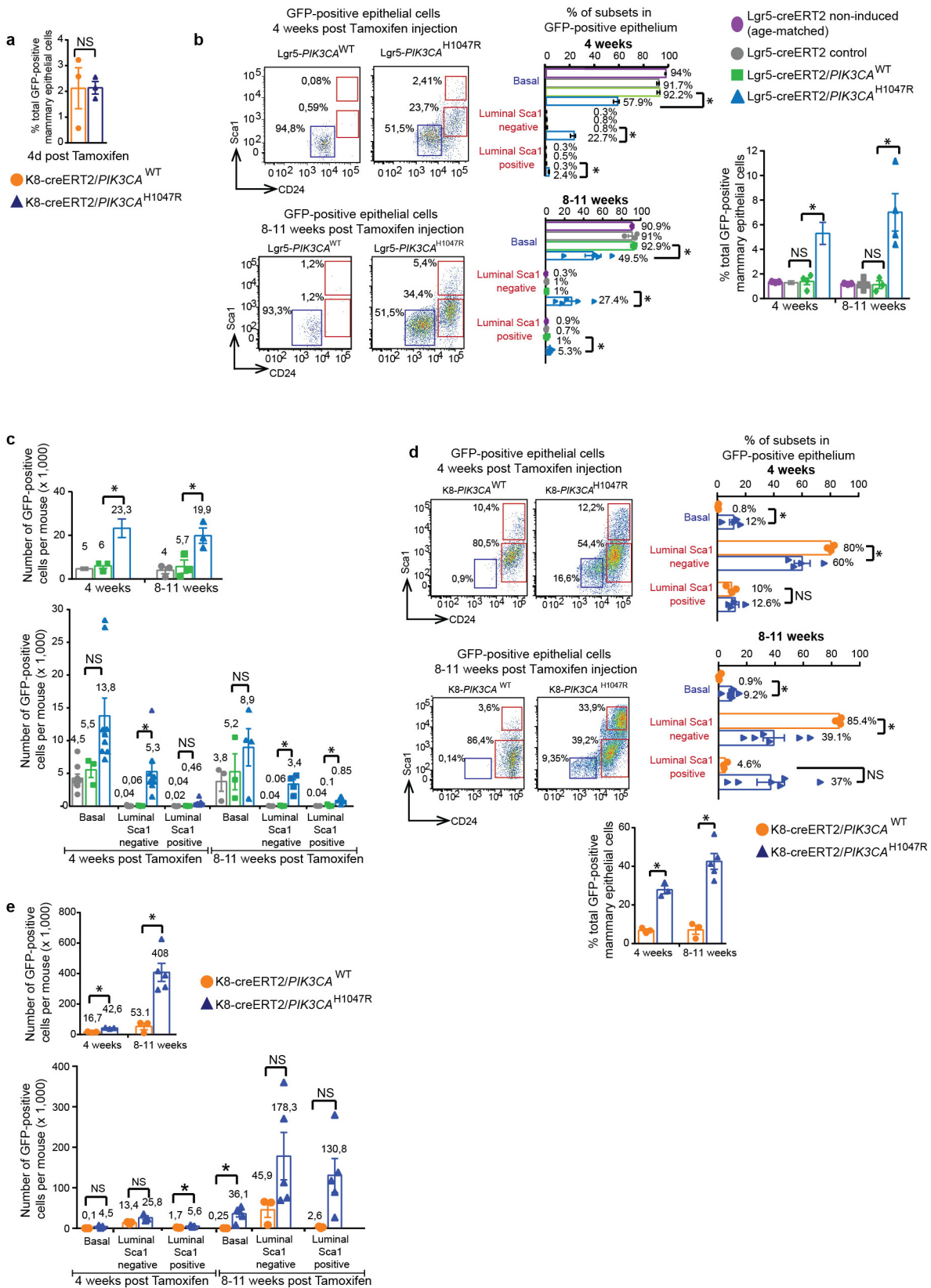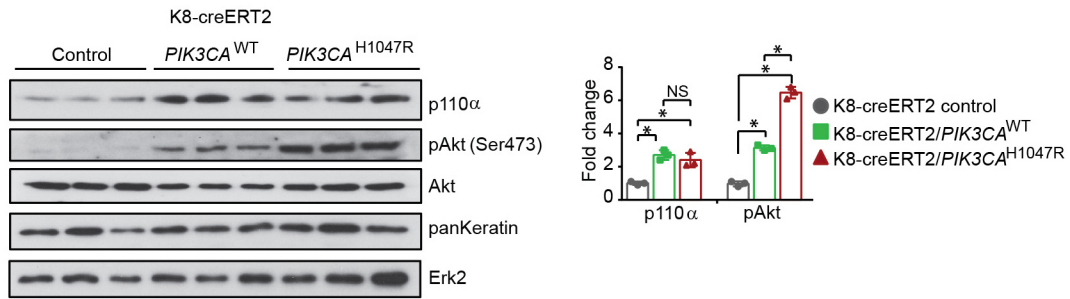
Extended Data Figure 5 | Tracing of GFP-positive mammary subsets.
a, Percentage of GFP-labelled cells in K8-CreER^{T2}/$PIK3CA^{H1047R}$ versus
K8-CreER^{T2}/$PIK3CA^{WT}$ animals 4 days after tamoxifen (24 h after the last
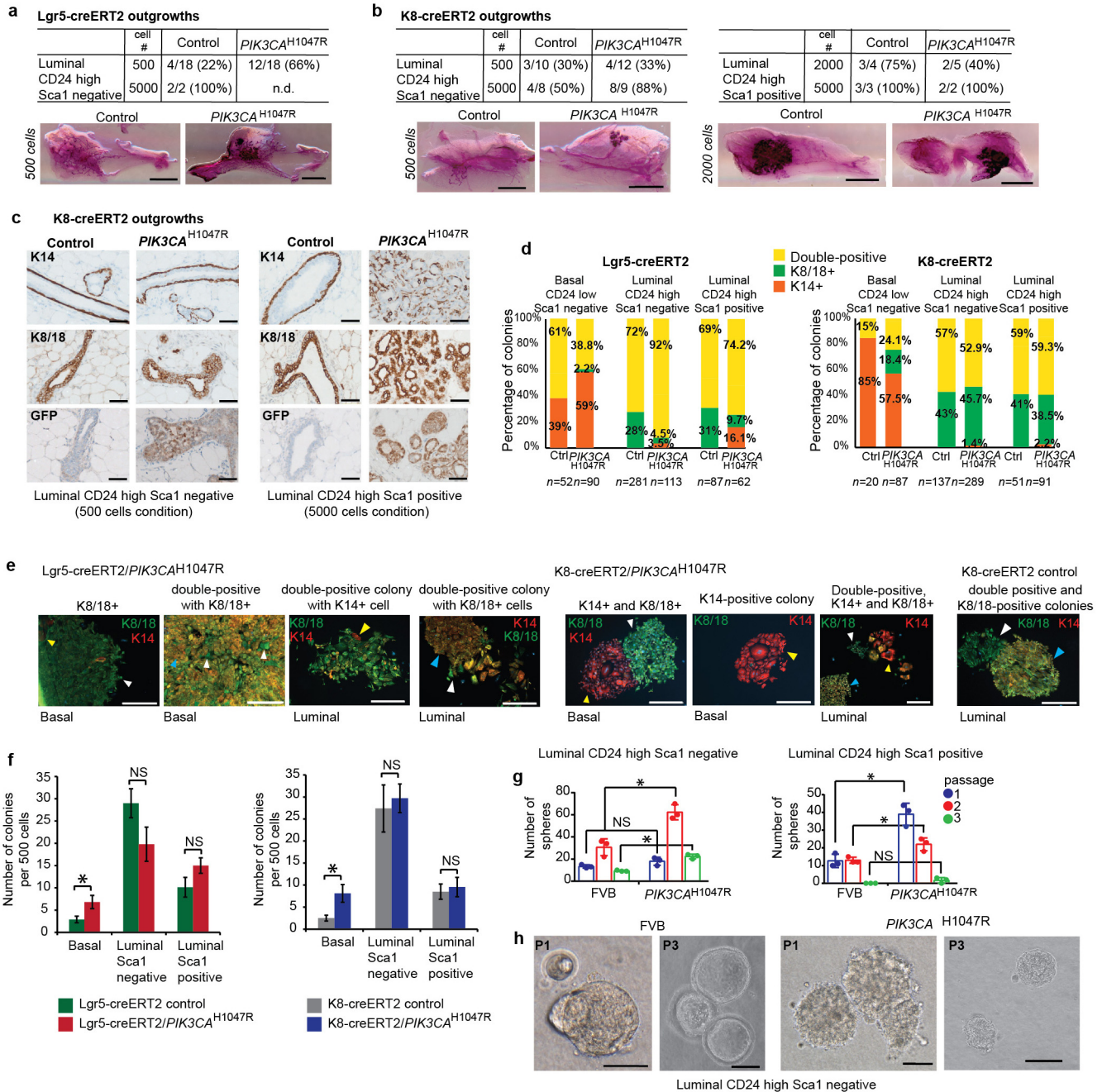tamoxifen injection) ($n = 3$ technical replicates, 2 mice per genotype).
b, **d**, Representative FACS plots and percentages of GFP-positive cells in
mammary gland subsets and total mammary epithelial cells 4 and 8–11 weeks
after tamoxifen. **c**, **e**, Bar graphs showing total numbers of GFP-positive cells
and numbers of GFP-positive cells in basal (CD24^{Lo}Sca1$^-$) and luminal
(CD24^{Hi}Sca1$^-$; CD24^{Hi}Sca1$^+$) subsets of Lgr5-CreER^{T2}/$PIK3CA^{H1047R}$ (**c**) and

K8-CreER^{T2}/$PIK3CA^{H1047R}$ (**e**) mammary epithelial cells. **b**, **c**, 4 weeks:
non-induced control $n = 3$, control $n = 9$, $PIK3CA^{WT}$ $n = 3$, $PIK3CA^{H1047R}$
$n = 9$ sortings with each 1–4 pooled mice; 8–11 weeks: non-induced control
$n = 3$, control $n = 3$, $PIK3CA^{WT}$ $n = 3$, $PIK3CA^{H1047R}$ $n = 4$ sortings with each
1–4 pooled mice. **d**, **e**, 4 weeks: $PIK3CA^{WT}$ and $PIK3CA^{H1047R}$ $n = 3$ sortings
with each 1–5 pooled mice; 8–11 weeks: $PIK3CA^{WT}$ $n = 4$, $PIK3CA^{H1047R}$
$n = 5$–6 sortings with each 1–4 pooled mice. Bar graphs show means ± s.e.m.;
two-sided unpaired Student's t-test; *$P < 0.05$; NS, not significant.

Extended Data Figure 6 | Expression of *PIK3CA*[H1047R] induces Akt phosphorylation. Immunoblot and quantification of lysates from K8-CreER[T2] control, *PIK3CA*[WT] and *PIK3CA*[H1047R] mammary glands 4 weeks after tamoxifen for p110α, pAkt, Akt, pan-keratin and Erk2 (loading control).

$n = 3$ mice per genotype. Protein levels were normalized to pan-keratin for normalization of epithelial content. Bar graphs depict fold change over control lysate. Bar graph shows means ± s.d.; two-sided unpaired Student's t-test; *$P < 0.006$; NS, not significant.

Extended Data Figure 7 | Expression of basal- and luminal-lineage genes in *PIK3CA*[H1047R] subsets. a, Expression heat maps of selected luminal and basal genes (left: Lgr5-CreER[T2]/*PIK3CA*[H1047R] versus control; right: K8-CreER[T2]/*PIK3CA*[H1047R] versus control). LM, mature luminal cells; LP, luminal progenitors; Myo, myoepithelial; SC, stem-cell enriched. **b, c,** Expression profiles of basal- and luminal-lineage genes in mammary subsets of Lgr5-CreER[T2]/*PIK3CA*[H1047R] compared with Lgr5-CreER[T2] control (**b**) and K8-CreER[T2]/*PIK3CA*[H1047R] compared with K8-CreER[T2] control animals (**c**). The qRT–PCR results are representative of 2–3 experiments of 4 pooled animals of each genotype. Bar graphs show means ± s.e.m.; two-sided unpaired Student's *t*-test; *$P < 0.05$; NS, not significant; N.d., not detected.
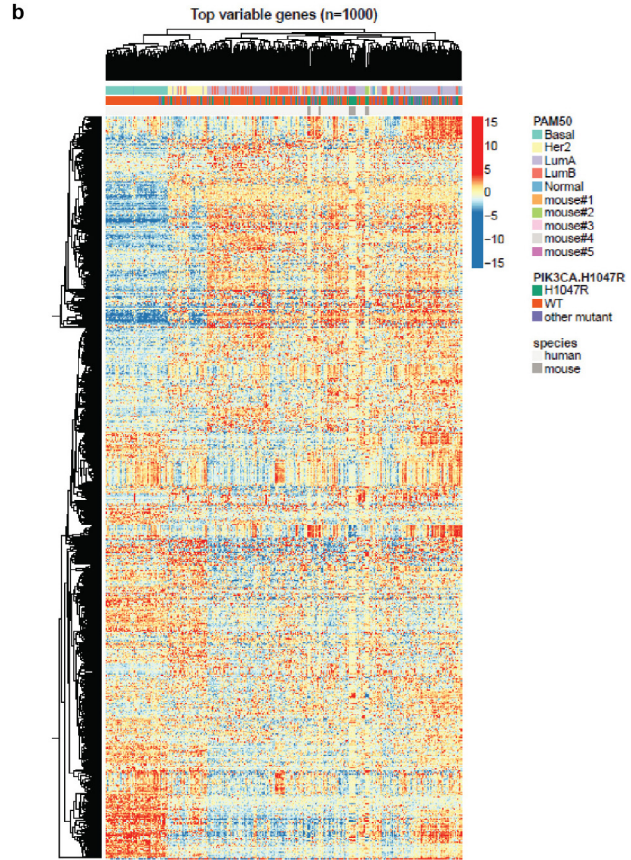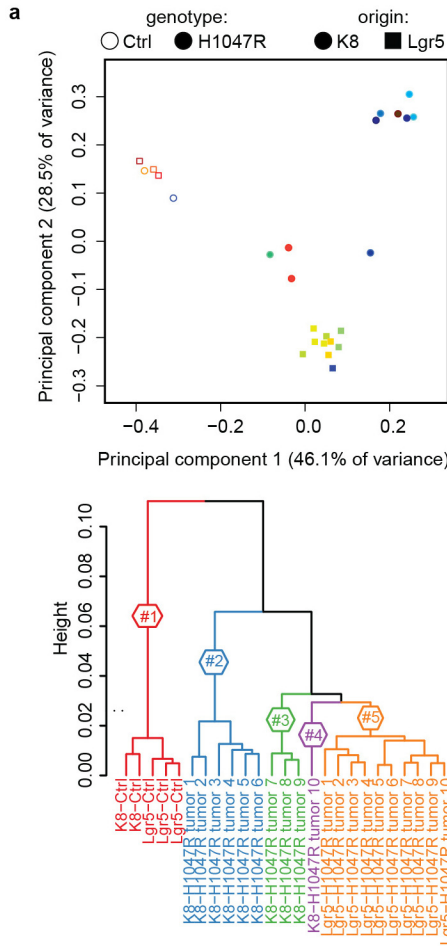
Extended Data Figure 8 | Luminal *PIK3CA*^H1047R cells repopulate a mammary gland. a, **b**, Number of outgrowths in cleared-fat pad transplantation of GFP-negative Lgr5-CreER^T2 control and GFP-positive Lgr5-CreER^T2/*PIK3CA*^H1047R-expressing luminal subsets (CD24^Hi^Sca1^−) (**a**) and GFP-negative K8-CreER^T2 control and GFP-positive K8-CreERT2/*PIK3CA*^H1047R-expressing luminal subsets (left, CD24^Hi^Sca1^−; right, CD24^Hi^Sca1^+) (**b**). Representative carmine-stained whole mounts (bottom). Scale bars, 500 μm. **c**, Representative immunostained sections. Scale bars, 50 μm. **a–c**, Data from three independent experiments. **d**, Percentage of K14-, K8/18- and double-positive (K14/K8/18) colonies derived from Lgr5-CreER^T2/*PIK3CA*^H1047R, Lgr5-CreER^T2 control (left, pooled data from *n* = 4 independent experiments (1–5 pooled mice)), K8-CreER^T2/*PIK3CA*^H1047R and K8-CreER^T2 control subsets (right, pooled data from *n* = 3 independent experiments (1–5 pooled mice)). Total number of quantified colonies is shown. **e**, Representative images of colonies. Arrowheads indicate K8/18- (white), K14- (yellow) and double-positive (blue) colonies. Scale bars, 500 μm. **f**, Number of colonies derived from basal and luminal cells from Lgr5- and K8-CreER^T2/*PIK3CA*^H1047R and control mice. Left, pooled data from three independent

sortings (each 1–5 pooled animals), total *n* = 8 (control), *n* = 10 (mutant) technical replicates for basal subset, *n* = 9 (control), *n* = 5 (mutant) technical replicates for luminal CD24^Hi^Sca1^− subset and *n* = 8 (control), *n* = 4 (mutant) technical replicates for luminal CD24^Hi^Sca1^+ subset. Right, pooled data from two independent sortings (each 1–5 pooled animals), total *n* = 8 (control), *n* = 10 (mutant) technical replicates for basal subset, *n* = 5 (control), *n* = 10 (mutant) technical replicates for luminal CD24^Hi^Sca1^− subset and *n* = 6 (control), *n* = 9 (mutant) technical replicates for luminal CD24^Hi^Sca1^+ subset. Five-hundred cells were seeded for each replicate. A colony was defined as a cell cluster of >5 cells. Bar graphs show means ± s.e.m.; two-sided unpaired Student's *t*-test; *P < 0.05. **g**, Bar graphs showing number of spheres derived from FVB-control and *PIK3CA*^H1047R-expressing luminal (CD24^Hi^Sca1^−/+) mammary cells over three passages. Representative data (three replicates, *n* = 4 mice per genotype) from two independent experiments. Bar graphs show means ± s.d. *P < 0.02, two-sided unpaired Student's *t*-test. **h**, Representative images of spheres derived from CD24^Hi^Sca1^− cells in passage one (P1) and three (P3). Scale bars, 100 μm. N.d., not determined; NS, not significant.
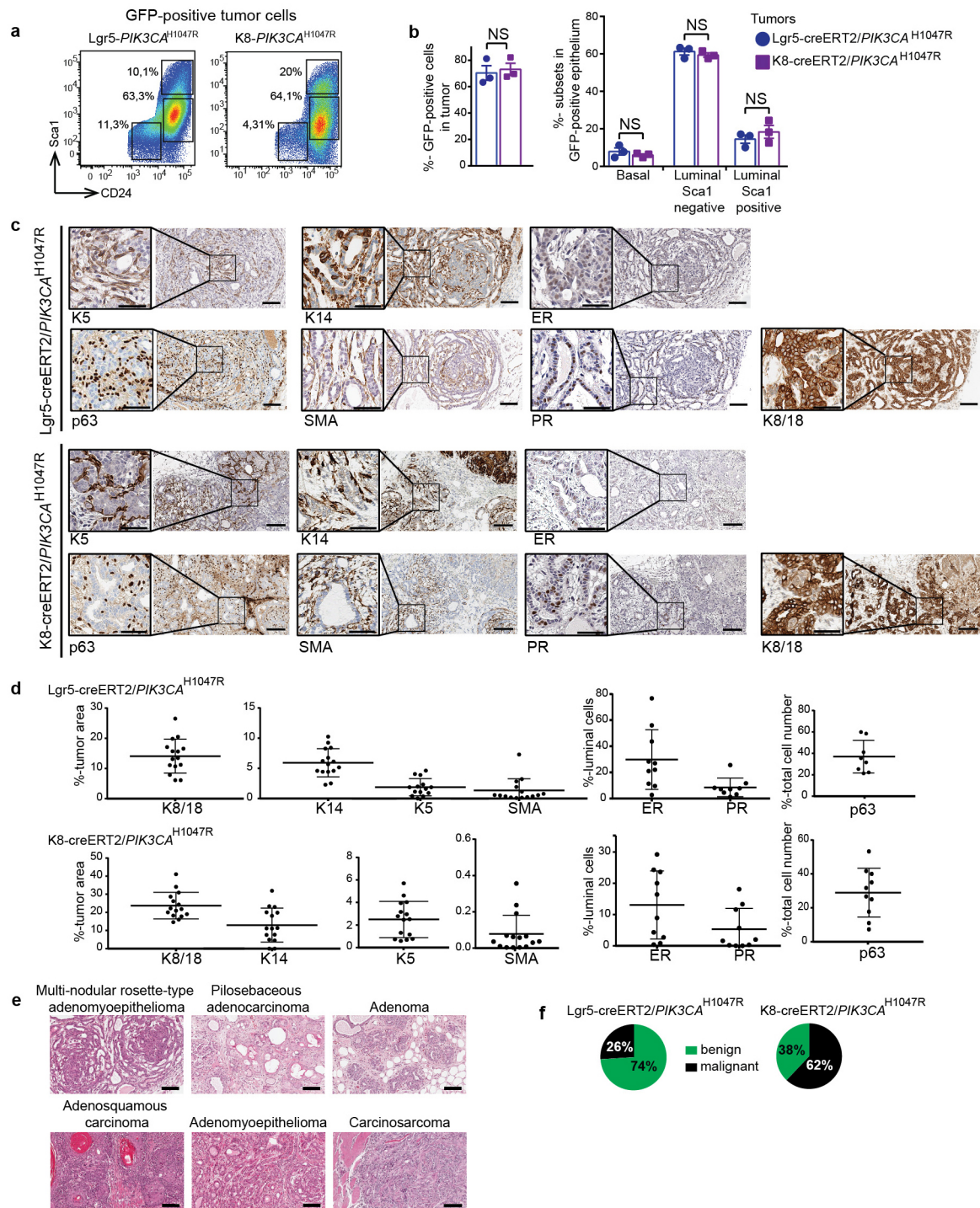
Extended Data Figure 9 | *PIK3CA*[H1047R]-evoked tumours express basal and luminal markers. **a**, Representative FACS plots of Lgr5-CreER[T2]/ *PIK3CA*[H1047R] and K8-CreER[T2]/*PIK3CA*[H1047R] tumours (*n* = 3). **b**, Percentages of total GFP-positive cells and GFP-positive basal (CD24[Lo]Sca1[−]) and luminal (CD24[Hi]Sca1[−/+]) subsets of Lgr5-CreER[T2] and K8-CreER[T2]/ *PIK3CA*[H1047R] tumours (*n* = 3). Bar graphs show means ± s.e.m. NS, not significant; two-sided unpaired Student's *t*-test **c**, Immunostaining for basal and luminal markers on serial sections of a multi-nodular rosette-type adenomyoepithelioma (Lgr5-CreER[T2]/*PIK3CA*[H1047R]) and adenomyoepithelioma (K8-CreER[T2]/*PIK3CA*[H1047R]). Scale bars, 100 μm, 50 μm

(magnifications). **d**, Quantification of basal- and luminal-lineage markers of Lgr5-CreER[T2] and K8-CreER[T2]/*PIK3CA*[H1047R] tumours. Each dot represents one tumour (top: K8/18, K14 and SMA *n* = 15, K5 *n* = 14, ER *n* = 10, PR *n* = 9, p63 *n* = 8; bottom: K8/18, K14, SMA and K5 *n* = 15, ER, PR and p63 *n* = 10). All Lgr5-CreER[T2]/*PIK3CA*[H1047R] tumours and 8/10 and 6/10 of K8-CreER[T2]/*PIK3CA*[H1047R] tumours show more than 1% of ER- and/or PR-positive cells, respectively. Bar graphs show means ± s.d. **e**, Representative haematoxylin and eosin stainings of tumour phenotypes. Scale bars, 100 μm. **f**, Percentage of benign and malignant mammary tumours.

Extended Data Figure 10 | Expression profiling of K8- and Lgr5-CreER^T2/ *PIK3CA*^H1047R mammary tumours. a, Principle component analysis and dendogram of a hierarchical clustering of gene expression profiles from 10 K8- and 10 Lgr5-CreER^T2/*PIK3CA*^H1047R tumours and 2–3 reference mammary glands. Each dot indicates one sample. Circles represent K8-CreER^T2 and squares represent Lgr5-CreER^T2 animals expressing *PIK3CA*^H1047R (filled symbols) or not (open symbols). **b**, Heat map of the top 1,000 genes that vary between K8-CreER^T2/*PIK3CA*^H1047R and Lgr5-CreER^T2/ *PIK3CA*^H1047R tumours and The Cancer Genome Atlas (TCGA) human breast cancer gene signatures. Lum, luminal.